



Enero–Abril 2021, Vol. 28, N° 1, ISSN 1316–6832
Online ISSN: 2610-8240



Revista INGENIERÍA UC

Órgano de Divulgación
Científica y Tecnológica de la Facultad de Ingeniería
de la Universidad de Carabobo

Disponible en:
Revista Ingeniería UC

<http://servicio.bc.uc.edu.ve/ingenieria/revista>
www.redalyc.org/revista.oa?id=707

VALENCIA – VENEZUELA

ISSN 1316–6832
Depósito legal pp 92.0200
Online ISSN: 2610-8240

**Revista
Ingeniería UC**

Revista indizada en:
REVENCYT (Venezuela); Actualidad Iberoamericana (CIT–Chile);
IET/INSPEC (United Kingdom UK); DRJI (The Directory of Research Journal Indexing);
LATINDEX–Catálogo (México); REDALYC (México); Periódica (México)
Emerging Sources Citation Index, Thomson Reuters;
MIAR–Matriz de información para el análisis de revista;
RedIB Red Iberoamericana de Innovación y Conocimiento Científico;
DOAJ–Directory of Open Access Journals.

**Órgano de Divulgación Científica y Tecnológica
Facultad de Ingeniería, Universidad de Carabobo
Valencia–Venezuela**

Editor - Jefe

Prof. Ángel D. Almarza M., Dr. Dirección de Investigación y Producción Intelectual, Facultad de Ingeniería, Universidad de Carabobo, Venezuela. ✉

Editores Honorarios

Prof. Edilberto Guevara, PhD. Vocal of the National Tribunal of Water. National Water Authority, Perú. Instituto Hidrología Aplicada y Medio Ambiente (IHAMA), Lima, Perú. Facultad de Ingeniería, Universidad de Carabobo, Venezuela. Período (1992-2003). ✉

Prof. Francisco Arteaga, PhD. Facultad de Ingeniería, Universidad de Carabobo, Venezuela. Período (2003-2009). ✉

Prof^a. Lisbeth Manganiello, PhD. Facultad de Ingeniería, Universidad de Carabobo, Venezuela. Período (2009-2018). ✉

Editores Asociados

Prof^a. Evelyn Yamal. Dra. Facultad de Ingeniería, Universidad de Carabobo, Venezuela. ✉

Prof^a. Lissette Jiménez. Dra. Facultad de Ingeniería, Universidad de Carabobo, Venezuela. ✉

Prof^a. Beatriz Herrera. Dra. Universidad Católica de Leuven, Bélgica. ✉

Corrector de redacción y estilo

Prof^a. Vanessa Hurtado. M.Sc. Escuela de Ingeniería Mecánica. Facultad de Ingeniería. Universidad de Carabobo. Venezuela. ✉

Junta Editorial

Dr. Adelmo Ortiz Conde
Universidad Simón Bolívar, Venezuela. ✉
Dra. Maura Vázquez de Ramírez
Universidad Central de Venezuela, Venezuela. ✉
Dr. Guillermo Ramírez
Universidad Central de Venezuela, Venezuela. ✉
Dr. Antonino Caralli
Universidad de Carabobo, Venezuela. ✉
Dr. Pablo Baricelli
Universidad de Carabobo, Venezuela. ✉
Dr. Johnny Bullon
Universidad de los Andes, Venezuela. ✉
Dr. Sergio Mora
Universidad de Santiago de Chile, Chile. ✉
Dr. Herbert J. Sira R.
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México. ✉

Dr. Enrique Barbieri
University of Houston, Estados Unidos. ✉
Dr. Ángel Rios Castro
Universidad de Castilla La Mancha, España. ✉
Dr. Merlín Rosales
Universidad del Zulia, Venezuela. ✉
Dr. José A. Díaz
Universidad de Carabobo, Venezuela. ✉
Dr. Carlos Martínez
Universidad de Carabobo, Venezuela. ✉
Dr. Roberto Callarotti
Instituto Venezolano de Investigaciones Científicas, Venezuela. ✉
Dr. Teodoro García
Universidad de Carabobo, Venezuela. ✉

ISSN 1316–6832
Depósito legal pp 92.0200
Online ISSN: 2610-8240

**Revista
Ingeniería UC**

Revista indizada en:
REVENCYT (Venezuela); Actualidad Iberoamericana (CIT–Chile);
IET/INSPEC (Unitec Kingdom UK); DRJI (The Directory of Research Journal Indexing)
LATIDEX–Catálogo (México); REDALYC (México); Periódica (México)
Emerging Sources Citation Index, Thomson Reuters.
MIAR–Matriz de información para el análisis de revista;
RedIB Red Iberoamericana de Innovación y Conocimiento Científico;
DOAJ–Directory of Open Access Journals.

**Órgano de Divulgación Científica y Tecnológica
Facultad de Ingeniería, Universidad de Carabobo
Valencia–Venezuela**

Comité Técnico*

Dr. Luis G. Zárate López, Universidad Popular Autónoma del Estado de Puebla, México
Dr. Enrique Cázares Rivera, Instituto Tecnológico y de Estudios Superiores de Monterrey, México
Dra. Laura Sáenz, Universidad de las Fuerzas Armadas, Sede Latacunga, Ecuador
Ing. Johel Rodríguez, M.Sc., Universitat Politècnica de València, España
Ing. Pedro Vargas Ferrer, M.Sc., Universidad de Talca, Chile
Dr. Eliecer Colina, Universidad de Los Andes, Venezuela
Dr. Francisco García Sánchez, Universidad Simón Bolívar, Venezuela
Dr. Fernando Mora, Universidad Simón Bolívar, Venezuela
Dr. Carlos Hurtado, Universidad de Oriente, Venezuela
Dr. Giovanni De Mercato, Universidad Simón Bolívar, Venezuela
Dr. Gian Franco Passariello, Universidad Simón Bolívar, Venezuela
Dr. William Castro, Instituto Venezolano de Investigaciones Científicas, Venezuela
Dr. Oscar Crescente, Universidad de Oriente, Venezuela
Dr. Luis Mata Guevara, Universidad de Oriente, Venezuela
Dra. Maj Britt Mostue, Universidad de Oriente, Venezuela
*Lista parcial

Dra. Susana Salinas de Romero, Universidad del Zulia, Venezuela
Lic. Hernando Herrera Mata, M.Sc., Universidad de Oriente, Venezuela
Ing. Carlos Lameda, M.Sc., Universidad Nacional Experimental Politécnica Antonio José de Sucre, Venezuela
Ing. Alfonso Alfonsi, Universidad de Oriente, Venezuela
Dr. William C. Henriquez Guzmán, Universidad de Oriente, Venezuela
Dr. Sergio Pérez Pacheco, Universidad Estatal de Bolívar, Ecuador
Dr. César O. Seijas Fossi, Universidad de Carabobo, Venezuela
Dr. Demetrio Rey Lago, Universidad de Carabobo, Venezuela
Dr. Alfonso Zozaya, Universidad de Carabobo, Venezuela
Dra. Damelys Zabala, Universidad de Carabobo, Venezuela
Dra. Lilly Marcano, Universidad de Carabobo, Venezuela
Dr. Antonio Bonsjak, Universidad de Carabobo, Venezuela
Dr. Guillermo Montilla, Universidad de Carabobo, Venezuela
Dr. Luis Vallés, Universidad de Carabobo, Venezuela
Dr. José A. Díaz, Universidad de Carabobo, Venezuela

Frecuencia de publicación

Revista INGENIERÍA UC se publica un volumen anual en tres números: abril, agosto y diciembre

Información de Contacto

<http://servicio.bc.uc.edu.ve/ingenieria/revista/>


<https://www.redalyc.org/revista.oa?id=707>

Correo–e: revistaing@uc.edu.ve

Correo–e: revistaing@gmail.com

Correo–e: adalmarza@uc.edu.ve (A. Almarza)

Redes sociales

 Twitter: @RevistaIngUC



Instagram: revistainguc

Diseño, Diagramación y Montaje en L^AT_EX: **Revista Ingeniería UC**, Dirección de Investigación y Producción Intelectual,
Facultad de Ingeniería, Valencia, Venezuela
Web master: Ing. Francisco Ponte, Biblioteca Central,
Universidad de Carabobo, Valencia, Venezuela

ISSN 1316-6832
Online ISSN: 2610-8240



Revista Ingeniería UC

ÓRGANO DE DIVULGACIÓN CIENTÍFICA Y TECNOLÓGICA

FACULTAD DE INGENIERÍA

UNIVERSIDAD DE CARABOBO

Valencia – Venezuela

Contenido General, (*General content*)

Volumen 28 Número 1	1
Editorial	2
<i>Manuel Jiménez-Bahri, Ángel Almarza-Morales</i>	
1 Dinámica fractal de las partículas PM_{2,5} en la Ciudad de México	
Fractal Dynamic for the particles PM _{2,5} in Mexico City	
<i>Mario Aguilar-Fernández, Sergio Barrientos-Ramírez, Arturo Merlín-Rodríguez, Brenda García-Jarquín, Víctor M. López-Sánchez</i>	
	4
2 Disminución de la actividad del grafito irradiado y los residuos radioactivos líquidos	
Decrease in the activity of irradiated graphite and liquid radioactive waste	
<i>Valerii V. Krymsky, Natalya V. Plotnikova</i>	
	23
3 Modelamiento numérico de un flujo de escombros asociado a una rotura de presa en la subcuenca Quillcay, Áncash, Perú	
Numerical modeling of a debris flow associated with a dam break in the Quillcay sub basin, Ancash, Peru	
<i>Abelardo M. Díaz-Salas, Edilberto Guevara-Pérez, James D. Vidal-Moreno</i>	
	35
4 Fundamentos metodológicos para la creación y desarrollo de clase de sistemas de ingeniería natural en la gestión de recursos hídricos	
Basics of Methodology of Creating and Developing the Class of Natural Engineering Systems in Water Resources Management	
<i>Vladimir L. Bondarenko, Elguja D. Khetsuriani, Allaberdy I. Yliasov, Elena A. Semenova, Valery N. Azarov</i>	
	47
5 Modelación probabilística de la relación entre ENSO y sequías agrícolas en el sur de Perú	
Probabilistic model of the relationship between ENSO and the occurrence of agricultural droughts in southern Peru	
<i>Juan Cabrera, Jesús Mejía, Edilberto Guevara-Pérez</i>	
	59

6 Génesis y Geoquímica de Suelos de Paisajes Urbanísticos en la Costa Rusa del Mar Negro

Genesis and geochemistry of the soils of urban landscapes of the Black Sea coast of Russia

Lalita V. Zakharikhina, Lyudmila S. Malyukova

69

7 Anomalías Tecnogénicas Geoquímicas en Suelos y Sedimentos de Fondo en Voronezh

Technogenic geochemical abnormalities in the soils and bottom sediments of Voronezh

Natalia V. Kaverina, Semen A. Kurolap, Pavel M. Vinogradov, Tatyana V. Dubovitskaya, Natalya N. Nazarenko, Roman A. Kondaurovc

83

8 Desarrollo de modelos estadísticos multinivel para determinar presión de flujo de fondo en operaciones de pozos petroleros comerciales

Developing multilevel statistical models of determining bottom-hole flowing pressure in commercial oil well operations

Inna N. Ponomareva, Vladislav Ign. Galkin, Alexander V. Rastegaev, Sergey Vl. Galkin

97

9 Materiales Compuestos de Boro Producidos por Síntesis Autopropagante a Temperaturas Elevadas

Boron Carbide Composites produced by Self-Propagating High-Temperature Synthesis

Roza G. Abdulkarimova, Aizhan J. Seidualiyeva, Aisulu N. Batkal, Sanat Tolendiuly, Sergey M. Fomenko

111

10 Nota Técnica. Fragilización de aceros inoxidables dúplex SAF 2205 y SAF 2507 sometidos a tratamientos térmicos de envejecimiento artificial

Tech Note. Embrittlement of duplex stainless steels SAF 2205 y SAF 2507 subjected to artificial aging heat treatments

Laura Sáenz, Krishna Aguilera, Reinaldo Heredia, Trino Romero, Caterine Donoso, Marcos Singaña

121

Jornada de Investigación de la Escuela de Ingeniería Eléctrica “Prof. César Rodolfo Ruiz” Octubre, 2020.

129

11 Diseño de línea 34,5 kV para respaldo de energía de los servicios auxiliares de la S/E La Arenosa 765 kV

Design of a 34,5 kV line for backup auxiliary power services of the La Arenosa 765 kV Substation

Santiago Abreu, Antonio Otero, Eva Monagas, César Rodolfo Ruiz ⁽⁺⁾, Juan C. Ataya, Yonder Guevara

130

12 Detector de Neuropatologías en EEG usando Estadísticas de Orden Superior y Aprendizaje Profundo

Neuropathology Detector in EEG using Higher Order Statistics and Deep Learning

César Seijas, Sergio Villazana, Guillermo Montilla, Egilda Pérez, Ricardo Montilla

141

13 Descriptor del Entorno para Personas con Discapacidad Visual Implementado en Raspberry Pi Basado en Redes Neuronales Artificiales Convolucionales y Recurrentes

Environment Descriptor for Visually Impaired People Implemented on Raspberry Pi Based on Convolutional and Recurrent Artificial Neural Networks

Rafael Chourio, Wilmer Sanz

152

14 Detección de fallas incipientes en rodamientos de Generadores Sincrónicos utilizando Máquinas de Vectores de Soporte

Detection of incipient faults of bearings in a Generator Synchronous using Support Vector Machines (SVMs)

Aniello Sparano, Jesús Ramírez, Ledy Guerra, Rubén Terán

165

15 Agrupamiento de Señales EEG con Rasgos Aprendidos Usando Autoencoder Profundo

EEG Signal Clustering With Learned Features Using Deep Autoencoder

Sergio Villazana, César Seijas, Guillermo Montilla, Egilda Pérez

180

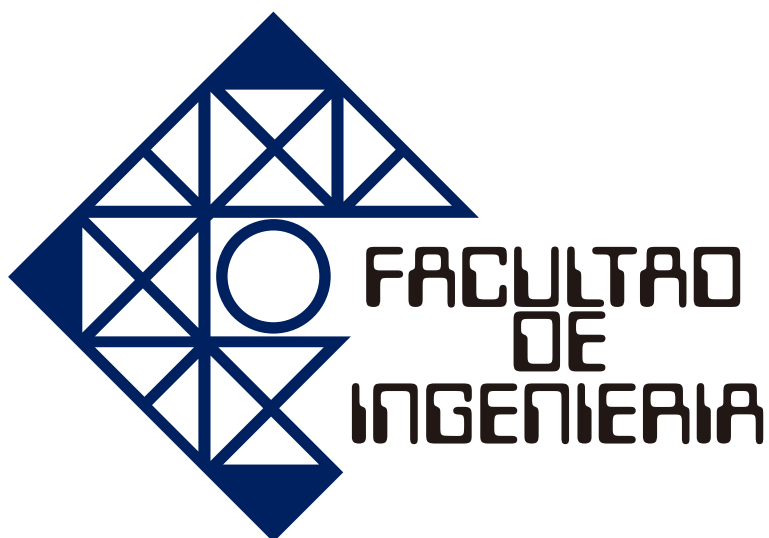
16 Identificación de modelo y evaluación comparativa de métodos de sintonización de controlador PID para proceso de regeneración catalítica en Scilab-XCOS

Model identification and comparative evaluation of PID controller tuning methods for a catalytic regeneration process in Scilab-XCOS

Gerardo Lanza, Francisco Arteaga-Bravo

193

Índice de materias	210
Subjects index	212
Índice de autores (<i>Author index</i>)	214
Apéndices	1
A. Línea editorial	2
B. Editorial line	3
C. Políticas de ética y publicación	4
D. Ethics and Publication Policies	7
E. Normas para la presentación de artículos	10
F. Authors Information	13
G. Carta de compromiso	15
H. Compromise letter	16
I. Índices de Revista Ingeniería UC	17



Volumen 28 Número 1

Editorial

En esta edición del Volumen 28, N° 1 de Revista Ingeniería UC continuamos avanzando con la normalización y visibilidad como órgano editorial en medio de una situación, con dificultades subyacentes del sector universitario a nivel nacional, y la dura realidad que como seres humanos nos ha tocado enfrentar por tiempo de pandemia que genera una crisis a nivel mundial y la lamentable pérdida de seres queridos y allegados.

Agradecemos el trabajo de todo el equipo editorial, técnicos, árbitros y revisores, así como el de todos nuestros autores, quienes han depositado su confianza en nosotros para la publicación de cada una de sus investigaciones.

Iniciamos con la investigación presentada por Aguilar y su grupo, siendo su estudio vinculado al área de ambiente, específicamente analizando la contaminación atmosférica a través de series de tiempo mediante análisis fractal y de rugosidad cinética de las partículas $PM_{2,5}$ en diferentes zonas de Ciudad de México, definiendo así las distribuciones características según la zona analizada. De igual forma Krymsky y Plotnikova estudian la disminución de la actividad del grafito irradiado y residuos líquidos radioactivos por medio de pulsos electromagnéticos.

Por otro lado, Díaz y sus colaboradores presentan los resultados del modelamiento numérico de flujo de escombros, asociado a una posible rotura de la presa Shallap, ubicada en la cabecera de cuenca del río Quillcay, en el departamento de Áncash, Perú. Mientras que, Bondarenko y su grupo de investigadores muestran los fundamentos metodológicos para la creación y desarrollo de clase de sistemas de ingeniería natural en la gestión de recursos hídricos.

Así mismo, Cabrera, Mejía y Guevara analizan la información de precipitaciones en la Provincia de Candarave, ubicada en el Sur del Perú contrastándola con diferentes índices del fenómeno ENSO para definir la existencia de influencia de este fenómeno sobre las temporadas secas en la región, alcanzando la construcción de un modelo probabilístico usando funciones cópula.

Zhakharihina y su grupo, describen geoquímicamente los suelos de paisajes Urbanísticos en la Costa Rusa del Mar Negro, determinando el índice total de contaminación de estos. Paralelamente, Kaverina y su equipo de investigadores, enfocan su estudio en las anomalías tecnogénicas geoquímicas en suelos y sedimentos de fondo de Voronezh.

Ponomareva y su grupo presentan modelos estadísticos multinivel para determinar la presión de flujo de fondo en operación de pozos petroleros. Por su parte, Abdulkarimova y colaboradores estudian las condiciones para la obtención de compuestos de carburo de boro, magnesia y óxido de aluminio por medio de síntesis autopropagante a temperaturas elevadas (SAPTE).

Sáenz y su equipo estudian la fragilización de los aceros inoxidables dúplex SAF 2205 y SAF 2507 sometidos a tratamientos de envejecimiento artificial.

Por otra parte, contamos en este número con los productos seleccionados de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica “Prof. César Rodolfo Ruíz”, evento realizado de forma conjunta entre la citada dependencia y la Dirección de Investigación y Producción Intelectual de la Facultad de Ingeniería UC.

Dentro de este contexto, presentamos el trabajo desarrollado por Abreu y colaboradores, quienes muestran el diseño de línea 34,5 kV para respaldo de energía de los servicios auxiliares de la S/E La Arenosa 765 kV. Por su parte, Seijas y su grupo, presentan un detector de neuropatologías, a partir del electroencefalograma (EEG) del paciente basada en la clasificación de imágenes a partir de análisis

de estadísticas de orden superior derivadas de series de tiempo utilizando aprendizaje profundo y una arquitectura de Red Neuronal Convolutiva pre-entrenada.

Chourio y Sanz presentan un descriptor del entorno para personas con discapacidad visual implementado en Raspberry Pi basado en redes neuronales artificiales convolucionales y recurrentes. Así, Sparano y colaboradores, presentan una metodología para detectar fallas incipientes en rodamientos de un generador sincrónico de prueba a través de máquinas de vectores de soporte.






Villazana y su grupo proponen un algoritmo basado en autoencoders convolucionales como extractor de rasgos no supervisado, para hallar grupos o clusters de señales electroencefalográficas (EEG), como apoyo para el especialista médico para facilitar el diagnóstico de la condición de epilepsia. Por último, Lanza y Arteaga-Bravo abordan los métodos de sintonización de controlador PID para proceso de regeneración catalítica en Scilab-XCOS teniendo como variable controlada la temperatura de la capa de catalizador.

Cerramos este editorial, invitando a publicar en nuestra Revista INGENIERÍA UC a todos aquellos investigadores nacionales e internacionales interesados en confiarnos sus productos científicos en las diferentes áreas de la ingeniería y ciencias afines.

Profesor *Manuel Jiménez-Bahri*
Decano de la Facultad de Ingeniería

Profesor *Ángel Daniel Almarza, Dr.*
Editor – Jefe

Fractal Dynamic for the particles PM_{2,5} in Mexico City

Mario Aguilar-Fernández ^{*,a} , Sergio Barrientos-Ramírez ^b , Arturo Merlín-Rodríguez ^a ,
Brenda García-Jarquín ^c , Víctor M. López-Sánchez ^b 

^aInstituto Politécnico Nacional. UPHICSA. SEPI. Iztacalco, Ciudad de México, México.

^bUniversidad Anáhuac México. Facultad de Ingeniería. Lomas Anáhuac, Huixquilucan, Estado de México. México.

^cInstituto Politécnico Nacional–ESIME-ZACATENCO. Manuel de Anda y Barredo. Unidad Profesional Adolfo López Mateos, Zacatenco, Gustavo A. Madero, Ciudad de México, México.

Abstract.- Air pollution is a relevant problem today due to the risk posed by urban environments. High concentrations of PM_{2,5} particles are extremely harmful to human health and cause high mortality rates in the most vulnerable population. The objective of this work is to characterize the dynamics of PM_{2,5} particles in the north, south, east, west and center of Mexico City by means of fractal analysis and kinetic roughness within the framework of the systemic to find emerging patterns (at different time-space scales), to model relevant aspects for decision-making in environmental policy and control strategies. The results show that PM_{2,5} levels adjust to different distributions, and that the time series of standard deviations display persistent behaviors as the time horizon lengthens. The projections that use the dynamic fractal analysis allow to know the response that the system may have over a time horizon. In conclusion, PM_{2,5} concentration levels vary in each country or region and depend on innumerable factors that satisfy the Family-Vicsek ansatz scaling dynamics.

Keywords: PM_{2,5}; Fractal Analysis; Time Series; Forecasts.

Dinámica fractal de las partículas PM_{2,5} en la Ciudad de México

Resumen.- La contaminación atmosférica es un problema relevante en la actualidad debido al riesgo que representa para los entornos urbanos. Las altas concentraciones de partículas PM_{2,5} son sumamente dañinas para la salud humana y generan altos índices de mortalidad en la población más vulnerable. El objetivo de este trabajo es caracterizar la dinámica de las partículas PM_{2,5} en las zonas norte, sur, oriente, poniente y centro de la Ciudad de México mediante análisis fractal y de rugosidad cinética dentro del marco de la sistémica para hallar patrones emergentes (a diferentes escalas de tiempo-espacio), para modelar los aspectos relevantes para la toma de decisiones en materia de política ambiental y estrategias de control. Los resultados muestran que los niveles de PM_{2,5} se ajustan a diferentes distribuciones, y que las series de tiempo de desviaciones estándar, muestran comportamientos persistentes a medida que se alarga el horizonte de tiempo. Las proyecciones que utilizan el análisis fractal dinámico permiten conocer la respuesta que puede tener el sistema en un horizonte temporal. En conclusión, los niveles de concentración de PM_{2,5} varían en cada país o región y dependen de innumerables factores que satisfacen la dinámica de escalado Family-Vicsek ansatz.

Palabras clave: PM_{2,5}; Análisis fractal; Series de tiempo; Pronósticos.

Recibido: 23 de enero, 2021.

Aceptado: 02 de marzo, 2021.

1. Introducción

La Organización Mundial de la Salud, estima que, una de cada nueve muertes en el mundo, es atribuida a la contaminación atmosférica [1]. A nivel nacional se estiman 26,941 muertes al año a

causa de las partículas PM₁₀ y PM_{2,5}. En la Zona Metropolitana del Valle de México, fallecieron alrededor de 2864 personas (7,8 diariamente, por cáncer de pulmón y situaciones del corazón), a causa de las partículas ultrafinas PM_{2,5} en el año 2017. Los costos sociales llegan a ser de entre el 3% y el 5%, del PIB nacional, relacionados con gases de efecto invernadero, ruido, siniestros de tránsito y congestión [2] lo que se traduce en un grave y creciente problema no solo ambiental, sino también social y económico.

Las partículas suspendidas (PM_{2,5} y PM₁₀), re-

* Autor para correspondencia:

Correo-e: maguilarf@yahoo.com (M. Aguilar-Fernández)

presentan una variable común en la contaminación atmosférica. En específico, las partículas $PM_{2,5}$ pueden llegar a los pulmones, además de ingresar al torrente sanguíneo, contribuyendo al desarrollo de afecciones cardiovasculares, respiratorias y cáncer pulmonar. Solamente, a causa de la exposición a las $PM_{2,5}$, mundialmente se estiman 4,2 millones de decesos prematuros al año [1][3], dato relevante en las estadísticas por muerte antes de nacer.

Para este estudio, la problemática se centra en la Ciudad de México [4], que es reconocida por ser un territorio con mayor vulnerabilidad ambiental en el país y donde urge la elaboración y la aplicación de nuevas prácticas en la gestión ambiental.

Las infructuosas acciones gubernamentales se relacionan con la escasa información y comprensión del problema [5], así como con la falta de competencia de los organismos encargados de atenderlo. Es necesario, que en materia ambiental, se desarrollen e instrumenten estrategias innovadoras [4] que complementen acciones políticas, legislativas y tecnológicas, dirigidas a favorecer los principios de sustentabilidad, en el manejo de los recursos naturales.

A pesar de que la Secretaría del Medio Ambiente del Gobierno de la Ciudad de México, cuenta con un sistema de pronóstico, este sólo estima el estado de la calidad del aire para las siguientes 24 horas, por ello, es importante desarrollar una herramienta útil que favorezca a tomar decisiones informadas, a lo largo de uno o varios horizontes de tiempo, para la protección de la salud y contribuir a mejorar la calidad del aire. Este trabajo se justifica por el hecho de que, el pronóstico en los niveles de partículas $PM_{2,5}$, es un componente de suma importancia en la investigación de la calidad del aire.

Es importante mencionar, que la mayoría de los estudios enfocados en partículas suspendidas de menos de 2,5 micras ($PM_{2,5}$) en la Ciudad de México [6][7], tienen como objetivo, solamente estimar los perfiles de fuentes de riesgo para la salud, de los residentes de dicha zona. Además, de que sólo utilizan modelos clásicos como el análisis factorial, o simples regresiones, para sus pronósticos. En este sentido, el análisis fractal, proporciona herramientas poderosas para comprender la naturaleza compleja, no lineal, de las

series temporales, en diversos campos que estudien sistemas naturales y socioeconómicos, complejos [8]. Los fractales, como su nombre lo indica, son partes fraccionales de un todo, que explican el comportamiento rugoso de una eventualidad o inconsistencia, a través de escalamientos y persistencias. Surgen como solución a los sistemas complejos que imposibilitan la proyección o estimación de valores subsecuentes a través del tiempo, y los generalizan en una dimensión no lineal. A su vez, modela el fenómeno de memoria larga, por medio de la ley de potencias y la enfoca en una distribución de colas pesadas [9] por su alta sensibilidad a estímulos iniciales en el tiempo.

Para determinar si el comportamiento de un sistema complejo, a partir de las series de tiempo auto-afines o estocásticas generadas a partir de dicho sistema, despliega persistencia a diferentes escalas de tiempo, se debe obtener, en principio, el valor del exponente Hurst (H) (o de rugosidad) para las series de tiempo auto-afín, con base en la ecuación (1).

$$F(bx) = aF(x) = bHF(x) \quad (1)$$

Donde el exponente $H = \log a / \log b$, es el exponente de Hurst [10]. La solución de la ecuación (1) es simplemente $F(x) = Ax^H$.

Los valores de $H < 1/2$, indican comportamiento anti-correlacionado (o antipersistente) a largo plazo: si los valores de la variable observable van aumentando, lo más probable es que el siguiente valor sea menor al último, y viceversa. Valores de $H > 1/2$, indican comportamiento correlacionado positivo a largo plazo: si los valores de la variable observable van aumentando, lo más probable es que el siguiente valor sea mayor al último, y viceversa. Por último, para valores de $H = 1/2$, no existen correlaciones, se tiene un comportamiento totalmente aleatorio.

Se realizó una revisión tradicional de literatura, con base en las propuestas de [23, 24, 25, 26], y los criterios de [27]. La exploración de documentos científicos, se produjo en el periodo de 1980 al 2020, en la base de datos *Web of Science*. Se pone especial atención en los artículos publicados con mayor influencia. Una breve descripción y estratificación, de cada uno de ellos, fue efectuada

Tabla 1: Artículos más citados en la *Web of Science* para los trabajos relacionados con los contaminantes PM_{2,5} en el mundo

Referencia	Descripción	Aportación	Metodología	Resultados
[11]	Un modelo híbrido de pronóstico para mapear el PM _{2,5} en Beijing.	Aplicación de redes neuronales artificiales.	Análisis de variables geo-atmosféricas con distribución logística.	Predicciones para PM _{2,5} con 2 días de anticipación identificando los días pico.
[12]	Localización de fuentes contaminantes en 25 ciudades de China.	El método de clustering permitió identificar fuentes potenciales contaminantes.	Clúster jerárquico y análisis de sensibilidad para las concentraciones del PM _{2,5} .	Un modelo de pronósticos estacionales temporales.
[13]	Identificación de las principales regiones contaminadas por PM _{2,5} .	Un modelo que estima las concentraciones de PM _{2,5} para el mundo.	A través de un modelo de investigación de operaciones.	China, EUA y Japón resultaron ser las regiones más contaminadas.
[14]	Un modelo de pronóstico de PM _{2,5} a largo plazo.	A diferencia de los modelos a corto plazo, este predice a largo plazo.	Análisis de series de tiempo considerando temperatura, índice de morbilidad.	Los registros satelitales facilitaron el mapeo en sobrepoblaciones.
[15]	Modelo lineal para la composición del PM _{2,5} en Heze, China.	Los modelos lineales resultaron útiles para el cálculo de impacto.	Modelo de factorización de matrices positivas y correlaciones lineales.	Patrones de estacionalidad en las localidades de la ciudad en estudio.
[16]	Modelo para pronosticar concentraciones de PM _{2,5} en EUA.	Se comprueba que la precisión de los datos de entrada influye en la salida.	Monitoreo satelital para un modelo de interpolación y regresión simplificado en R.	Alta precisión en las predicciones (> 8, 8).
[17]	Proyecto ESCAPE el cual estima PM en 20 regiones de Europa.	LUR que es un sistema de ecuaciones de regresión múltiple y GPS.	Se usó el modelo de regresión de uso del suelo (LUR).	Modelo con precisión relativamente baja (<7).
[18]	Con MAIAC se estimó el PM _{2,5} en el sureste de EUA.	Las concentraciones terrestres son más complejas que las del aire.	El algoritmo MAIAC es de naturaleza no lineal y fue empleado en radios de 1 km.	La precisión del modelo >8, ero se limita a áreas con acceso a esta tecnología.
[19]	Análisis para demostrar los efectos de los árboles con el PM _{2,5} .	Se demostró que la masa vegetativa es reguladora a favor de la salud.	Se modelaron distintas regiones de EUA mediante el algoritmo BenMap.	Una alta reducción de PM _{2,5} en las zonas con más población vegetativa.
[20]	Modelo de predicción de PM _{2,5} en Suiza del 2003 al 2013.	Uso del algoritmo MAIAC desarrollado por la NASA.	El algoritmo MAIAC usado en radios de 1 km.	Alta precisión en los resultados r>8.
[21]	Modelo multifractal asimétrico para PM _{2,5} China.	Modelos fractales con comportamientos multifractales.	Correlaciones asimétricas y series de tiempo.	Resultados con tendencia alcista. Similar a Hong Kong y Shangai.
[22]	Cotejo de 3 métodos para PM _{2,5} en Chile.	La evaluación de métodos de pronóstico tanto lineales como no lineales.	Métodos: estadística lineal, red neuronal multicapa y clustering jerárquico.	El algoritmo de clustering jerárquico resultó ser el pronóstico más preciso.

Fuente: información consultada en <https://www.clarivate.com/> (2020), *Web of Science*.

(los más citados y los relacionados con México). La calidad de la información sobre el tema PM_{2,5}, cumple el atributo de vigencia [28][29]. Se localizaron los artículos más citados, el 60 % enfocados en la contaminación atmosférica por

PM_{2,5} en China [30], según lo indicado en la Tabla 1.

En la Tabla 2 se revisaron los artículos enfocados en análisis de contaminación atmosférica para México con metodologías de orden complejo y

Tabla 2: Artículos dedicados al estudio de los contaminantes $PM_{2,5}$ en México, en la *Web of Science*

Referencia	Descripción	Aportación	Metodología	Resultados
[31]	Análisis de la calidad del aire en de 3 zonas de la ciudad de México.	Zonas: 1) Xalostoc (norte), 2) la merced (centro) y 3) pedregal (sur).	Análisis de muestras en laboratorio y estadística descriptiva.	$R > 0.9$: alta precisión en los resultados. Xalostoc fue la zona más contaminada.
[32]	Análisis de la calidad del aire en de 6 zonas de la ciudad de México.	Zonas: 1) Xalostoc, 2) Merced, 3) pedregal, 4) Nezahualcoyotl, 5) Tlalnepantla y 6) Iztapalapa.	Análisis de muestras en laboratorio y estadística descriptiva.	Nezahualcoyotl, Xalostoc e Iztapalapa fueron las zonas más contaminadas.
[33]	Análisis de la calidad del aire en de 6 zonas de la ciudad de México con estaciones satelitales.	Datos tomados de la estación satelital IMADA-AVER de EUA.	Recolección de datos vía satélite en 24 horas continuas y analizadas con IMADA-AVER.	La zona Nezahualcoyotl resultó ser la zona más contaminada por sulfatos y nitratos.
[34]	Identificación de fuentes y análisis de la composición del aire en la ciudad de Monterrey.	El monitoreo y análisis de contaminantes de $PM_{2,5}$ en la ciudad de Monterrey.	Análisis de cromatografía de 60 muestras del sureste y noreste de la ciudad.	Las principales fuentes son la refinería y zonas industriales. Sobresale el aerosol inorgánico.
[35]	Análisis de exposición al benceno, $PM_{2,5}$ y CO para los pasajeros de transporte público en la ciudad de México.	Conocer las posibles afecciones de salud a las que son expuestas las personas que viajan en transporte público diariamente.	Análisis gravimétrico para descomponer los contaminantes y regresiones para el análisis de datos.	Principales contaminantes: carbono y monóxido de carbono, principalmente en las horas pico.
[36]	Contribución de la quema de basura al $PM_{2,5}$ en la ciudad de México en marzo 2006.	Debido a que la quema de basura es una práctica común en México, la aportación de su análisis es de sumo impacto.	Se usó el modelo WRF-CHEM para simular las emisiones, transportes y compuestos químicos	La combustión por quema de basura constituye de un 3-30 % de los contaminantes de $PM_{2,5}$.
[37]	Evaluación de los riesgos a la salud de los habitantes de la ciudad de México por $PM_{2,5}$.	Este estudio se enfocó en la medición del nivel de riesgo de salud de acuerdo a las muestras poblacionales.	Análisis de partículas PIXE (particle induced X-ray emission).	La población de niños de 6 a 12 años de edad en Naucalpan y Xalostoc mostraron el mayor riesgo.
[38]	Modelo de redes neuronales para $PM_{2,5}$ en la frontera de México con EUA.	Las redes neuronales artificiales permitieron modelar de una manera precisa el comportamiento.	Se modeló con la base de datos de la red de monitoreo de calidad del aire de Texas.	Las redes presentaron coeficientes de correlación cercanos a 8.
[39]	Identificación de componentes orgánicos para el $PM_{2,5}$ en la ciudad de México.	Se utilizó segregación por temporadas en 5 zonas de la Cd. de México.	Análisis gravimétrico, multivarianza, regresión lineal y diagramas de cajas.	El carbón natural fue el mayor factor orgánico en la zona de Tlalnepantla en época de sequía.
[6]	Identificación de contaminantes en las fuentes emisoras en Cd. de México.	Identificación de contaminantes desde las principales fuentes emisoras.	Mediante análisis gravimétrico se tomaron muestras de diversas fuentes.	Presencia de metales en la quema de basura y carbono en las emisiones de vehículos de diésel.
[40]	Identificación de fuentes y distribución de carbono en ambientes de $PM_{2,5}$ y PM_{10} para la Cd. de México.	La mayor aportación reside en la evaluación con isótopos de carbón considerando al carbón como el mayor contaminante de las partículas suspendidas.	Recolección de muestras en 5 zonas de la ciudad de México y la evaluación con isótopos.	La zona norte de la ciudad de México resultó ser la mayor fuente de carbono por las incineradoras de plásticos y caucho.
[41]	Localización geográfica de las distribuciones del $PM_{2,5}$ en Cd. de México.	Recolección de datos por satélite AOD de forma localizada.	Recolección de datos con AOD y modelado en lenguaje R.	Alta precisión en los resultados pero se limita a radios de 1 km.

Fuente: información consultada en <https://www.clarivate.com/> (2020), *Web of Science*.

comparando tanto la Tabla 1 y 2, se tiene que en conclusión, se revisaron 24 artículos de la *Web of Science*, donde los más citados usaron modelos no lineales. En México, sólo se ha realizado 1 no

lineal (redes neuronales) pero no se ha empleado el análisis fractal, he aquí la necesidad de realizar el presente estudio.

2. Metodología

El tipo de estudio es de carácter correlacional, ya que busca determinar si existe, o no, persistencia en las series de tiempo de los niveles de partículas $PM_{2,5}$, y en sus fluctuaciones. El estudio es longitudinal, puesto que se buscan cambios en la relación en determinadas variables a través del tiempo.

Las cifras históricas de los niveles de partículas $PM_{2,5}$, se obtuvieron de la Red Automática de Monitoreo Atmosférico (RAMA), la cual, emplea equipos, de manera continua, para la medición de dióxido de azufre, monóxido de carbono, dióxido de nitrógeno, ozono, PM_{10} y $PM_{2,5}$. La RAMA se forma por 34 estaciones de monitoreo, y un laboratorio para el mantenimiento y calibración de los equipos de monitoreo, a través de los métodos de gravimetría y de atenuación de radicación beta [42].

Las cifras históricas de los niveles de partículas $PM_{2,5}$, se obtuvieron del sitio [42], de la Secretaría del Medio Ambiente de la Ciudad de México. La información que se genera es pública y los usuarios son la población en general.

Dicha Secretaría del Gobierno, calcula el promedio de 24 horas desde la hora 1 a la 24, por lo que el primer dato de publicación, corresponde a la hora 1 del día, es el promedio aritmético de los minutos comprendidos dentro las 00 horas, es decir, se realiza con los datos desde el minuto 0 (00:00) al minuto 59 (00:59). Este promedio aritmético de las 00 horas se publica en la hora siguiente, es decir a la 01:00 hora, y así sucesivamente, para cada hora, hasta llegar al cálculo del último dato del día (hora 24), los cuales son los datos comprendidos dentro de las 23:00 horas (de las 23:00 a las 23:59), y este se publica a las 00:00 horas. Para no tener series de tiempo excesivamente grandes se decidió que de manera aleatoria, se seleccionaron cinco observaciones por día, para generar las series de tiempo originales, para los últimos cinco años (2016-primer semestre del 2020).

Como primer paso, de las series de tiempo originales $z(t)$ de los datos de partículas $PM_{2,5}$ contra tiempo, se construyeron series de tiempo de desviaciones estándar (fluctuaciones) de cada una

de las cinco zonas de la Ciudad de México (norte, sur, poniente, oriente y centro), ya que la desviación estándar, es el parámetro más empleado para determinar la variación, cambio o fluctuación de una variable (en este caso de las partículas $PM_{2,5}$), a través del tiempo. Para caracterizar la dinámica de las series de tiempo $v(t, \tau)$, se consideró la longitud de cada una de las 5 series de tiempo $z(t)$ de $T = 6,650$ registros de partículas $PM_{2,5}$ (μV) versus tiempo (segundos), con una tasa de muestreo $\Delta t = 1$ segundo. Así como también se consideró para las 5 series de tiempo $v(t, \tau)$, un rango del intervalo de tiempo de la muestra de $3 \leq \tau \leq 300$, con ventanas de tiempo de los intervalos de las muestras de $\delta_t = 1$ segundo.

Una vez obtenidos los datos necesarios para la construcción de las series de tiempo de la zona norte, sur, centro, oriente y poniente de la Ciudad de México, se analizó y caracterizó la dinámica de las fluctuaciones de las emisiones de partículas $PM_{2,5}$, dentro de un marco conceptual de la cinética del crecimiento de interfaces rugosas (análisis fractal o de escalamiento). El comportamiento aleatorio de la dinámica espacio-tiempo, puede ser caracterizado por las fluctuaciones de los parámetros observables. De igual forma, la memoria a largo plazo en las series de tiempo de fluctuaciones, se ha analizado frecuentemente, a través de su función de estructura, definida según la ecuación (2).

$$\sigma(\tau, \delta_t) = \langle \overline{[v(t + \delta_t, \tau) - v(t, \tau)]^2} \rangle^{(1/2)} \quad (2)$$

Donde la barra superior, denota el promedio sobre todos los tiempos t en las series de tiempo de longitud $T - \tau$ (con T como la longitud de la serie de tiempo original $z(t)$, y los paréntesis triangulares, indican el promedio sobre diferentes realizaciones de la ventana de tiempo de tamaño δ_t . Normalmente, se espera que la dinámica de escalamiento de invarianza, implique que las fluctuaciones $\sigma(\tau, \delta_t)$ satisfagan la dinámica de escalamiento *ansatz* de Family-Vicsek. Dicha función de estructura de las fluctuaciones, exhibe comúnmente, el comportamiento de ley de potencia, característico de los fractales auto-afines (Balankin, 2007) según la ecuación (3).

$$\sigma \propto (\delta_t)^\zeta \quad (3)$$

El exponente de escala o de rugosidad ζ , mejor conocido como exponente de Hurst, se encarga de caracterizar el grado de intensidad de las correlaciones a largo plazo, en la dinámica de las fluctuaciones.

De acuerdo con Morales [43], el cálculo del coeficiente Hurst, por el método de Rango Reescalado (R/S), se inicia con una serie de tiempo de tamaño M , pero como interesan los rendimientos logarítmicos, entonces la serie original, se reduce a una nueva serie de tiempo de tamaño $N = M - 1$, donde cada rendimiento está definido por la ecuación (4).

$$N_i = \log \left(\frac{M_i + 1}{M_i} \right), \quad i = 1, 2, 3, \dots, N \quad (4)$$

Se divide este periodo de tiempo N en A subperiodos contiguos de longitud n , tal que $A \cdot n = N$. Se nombra cada uno de los subperiodos o subgrupos I_a , con $a = 1, 2, 3, \dots, A$. Y cada elemento es etiquetado por $N_{(k,a)}$ tal que $k = 1, 2, 3, \dots, n$. Y para cada subperiodo I_a de longitud n , el valor promedio está definido, por la ecuación (5).

$$e_a = \frac{1}{n} \sum_{i=1}^n N_{k,a}. \quad (5)$$

La diferencia de cada elemento $N_{(k,a)}$ con respecto a la media e_a para cada subperiodo I_a se van sumando para obtener la serie de tiempo acumulada ($X_{k,a}$), definida según la ecuación (6) como:

$$X_{k,a} = \sum_{i=1}^k (N_{i,a} - e_a), \quad k = 1, 2, 3, \dots, n. \quad (6)$$

El rango R_{I_a} se define como la diferencia entre el valor máximo y el valor mínimo de $X_{k,a}$ para cada subperiodo I_a , de acuerdo con la ecuación (7).

$$R_{I_a} = \text{Max}(X_{k,a}) - \text{Min}(X_{k,a}), \quad (7)$$

donde $1 \leq k \leq n$.

Para el análisis estadístico se siguió la metodología de Morales [43], se utilizó el software @Risk, con la finalidad de ajustar los datos a la mejor

distribución, y posteriormente, se analizaron sus parámetros estadísticos para identificar distribuciones de colas pesadas (comportamiento de leyes de potencia), esto con el objetivo de identificar si persiste algún comportamiento complejo o fractal en la distribución de la contaminación por $\text{PM}_{2.5}$. a través de indicar el rango de valores, el software @Risk, ofrece un listado de posibles distribuciones ordenadas por bondad de ajuste, por medio de los criterios estadísticos de la Chi-cuadrada, Anderson-Darling, y Kolmogorov-Smirnov, lo cual permite comparar los datos en un gráfico.

3. Análisis y discusión de resultados

Para un mayor orden y entendimiento de los resultados obtenidos, estos se muestran divididos en series de tiempo original, análisis estadístico, análisis dinámico fractal y pronóstico. Los hallazgos obtenidos se muestran a continuación.

En las Figuras 1a, 2a, 3a, 4a, 5a se muestran las gráficas de las series de tiempo originales, $z(t)$, de las zonas sur (Figura 1), poniente (Figura 2), centro (Figura 3), oriente (Figura 4) y norte (Figura 5) de la CDMX para intervalos de tiempo $\tau = 0$ s, $\tau = 75$ s, $\tau = 150$ s y $\tau = 300$ s, respectivamente.

Posteriormente, lo anterior sumado al análisis estadístico realizado con el software @Risk citado en la metodología se analizaron los parámetros estadísticos para identificar algún comportamiento fractal y la Figura 6 indica que los datos originales de las series de tiempo de la zona norte, sur, centro, oriente y poniente de la CDMX, tienen un mejor ajuste con las distribuciones Gamma, Log-logística, Log-normal, Pearson 5 y Gamma, respectivamente. También, al analizar los resultados del ajuste de datos, se puede apreciar que existe un exceso de curtosis en todas las series originales $z(t)$, lo cual significa que la distribución de los resultados de los eventos, presenta valores extremos (datos atípicos), causando colas pesadas en la curva de distribución. Cuanto más alto sea el coeficiente de curtosis por encima del «nivel normal», o «cuanto más gruesas» sean las colas en el gráfico de distribución de retornos, más probable será que los retornos futuros sean extremadamente

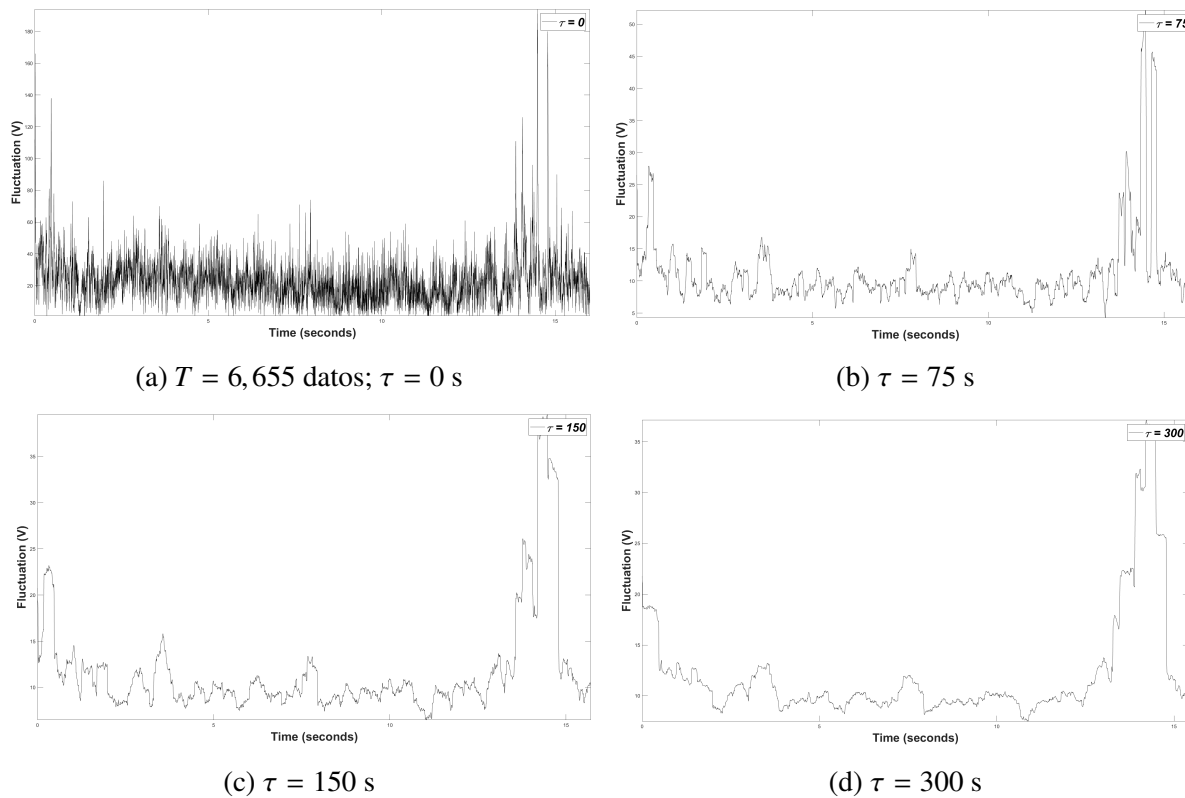


Figura 1: Series de tiempo de fluctuaciones $v(t, \tau)$, del canal F_{p1} con intervalos de tiempo, para la zona sur de la CDMX

grandes o pequeños. Para confirmar la presencia de colas pesadas (es decir, exceso de curtosis >0) en las distribuciones de las distintas zonas analizadas, se realiza una prueba estadística, la cual descartó la presencia de ruido blanco y de normalidad. De hecho, los datos originales de la muestra de cada zona, exhiben un efecto ARCH, ya que los valores cuadrados de la serie de tiempo están autocorrelacionados.

Muchos conjuntos de datos de series temporales muestran una interdependencia temporal entre sus valores, por lo que es necesario detectar y optimizar la calidad de previsión del modelo. A través del complemento NumXL, se puede simplificar el proceso de construcción de los gráficos ACF (Función de Autocorrelación) y PACF (Función de Autocorrelación Parcial), ya que dicho software, admite numerosas funciones. Los resultados sobre las series originales para todas las zonas de la CDMX, muestran la presencia de autocorrelación, tanto en la ACF como en la PACF. También, es

posible ejecutar el análisis de correlograma en las series temporales diferenciadas (es decir, en los rendimientos logarítmicos y en los valores absolutos de estos rendimientos). Los retornos del registro, muestran una fuerte interdependencia, ya que las barras laterales exceden los límites, por lo que existe presencia de autocorrelación para todas las zonas. Sin embargo, la función de autocorrelación de los niveles de partículas $PM_{2,5}$, decrece lentamente. Para cuantificar la intensidad de las correlaciones a largo plazo, se calcula el exponente de Hurst (H), utilizando la metodología de Rango Reescalado (R/S), para las series de tiempo de los niveles de partículas $PM_{2,5}$ en la CDMX.

Como indican los resultados de la Tabla 3 utilizando series de datos de rendimientos logarítmicos $\delta(\tau)$, todas las zonas de la CDMX, muestran un valor de ζ perteneciente al rango $0 < \zeta < 0,5$, lo que implica un comportamiento antipersistente o anticorrelacional. En otras palabras, se esperarí,

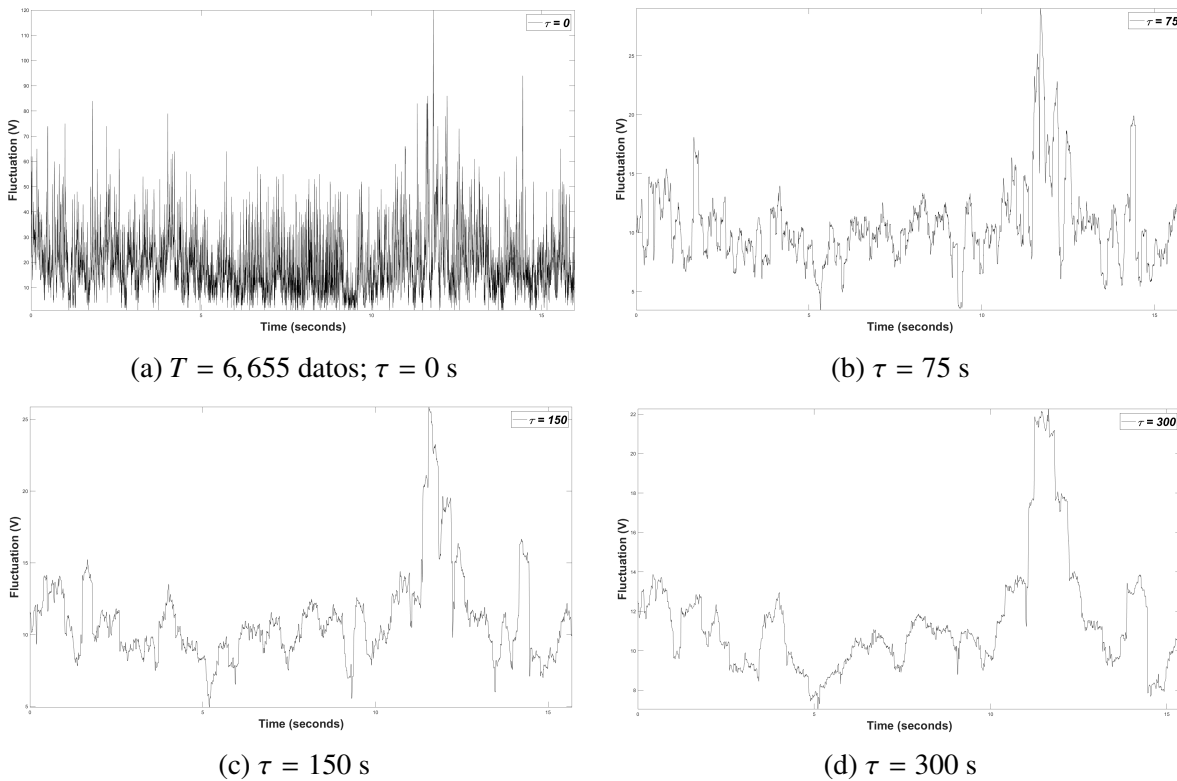


Figura 2: Series de tiempo de fluctuaciones $v(t, \tau)$, del canal F_{p1} con intervalos de tiempo, para la zona poniente de la CDMX

Tabla 3: Coeficiente de Hurst, dimensión fractal y exponente de crecimiento para las diferentes zonas analizadas de la CDMX utilizando el software @Risk

Zona	Coeficiente de Hurst (ζ) utilizando rendimientos logarítmicos $\delta(\tau)$	Coeficiente de Hurst (ζ) utilizando rendimientos logarítmicos absolutos $ \delta(\tau) $	Dimensión fractal $\delta(\tau)$	Dimensión fractal $ \delta(\tau) $	Exponente de Crecimiento (β)
Norte	0,23	0,78	1,77	1,22	0,96
Sur	0,18	0,79	1,82	1,21	0,98
Centro	0,18	0,73	1,82	1,27	0,89
Oriente	0,27	0,75	1,73	1,25	0,93
Poniente	0,16	0,84	1,84	1,16	0,96
CDMX	0,20	0,78	1,80	1,22	0,94

que tras un período de crecimiento en los niveles de partículas $PM_{2,5}$, se continuará con un período de decrecimiento, y viceversa. El haber obtenido este tipo de valores, se indica la presencia de ruido rosa, el cual se caracteriza por su alta frecuencia, y, por tanto, tiende a regresar a su lugar de procedencia, siendo señales muy irregulares. En cambio, utilizando series de tiempo del valor absoluto de los rendimientos logarítmicos $|\delta(\tau)|$, los valores de H , en todas las zonas, se ubican en el

rango $0 < H > 0,5$, lo que significa que las series de tiempo empleadas despliegan comportamiento persistente (correlaciones positivas), con un ajuste de ley de potencia. (Existe invarianza de escala asociada a correlaciones positivas a largo plazo). Hay que aclarar, que el valor absoluto de los rendimientos logarítmicos, refleja el carácter volátil de la serie, al describir la amplitud de la fluctuación, ya que éste siempre es positivo, y no existen tendencias globales que sean obvias [44].

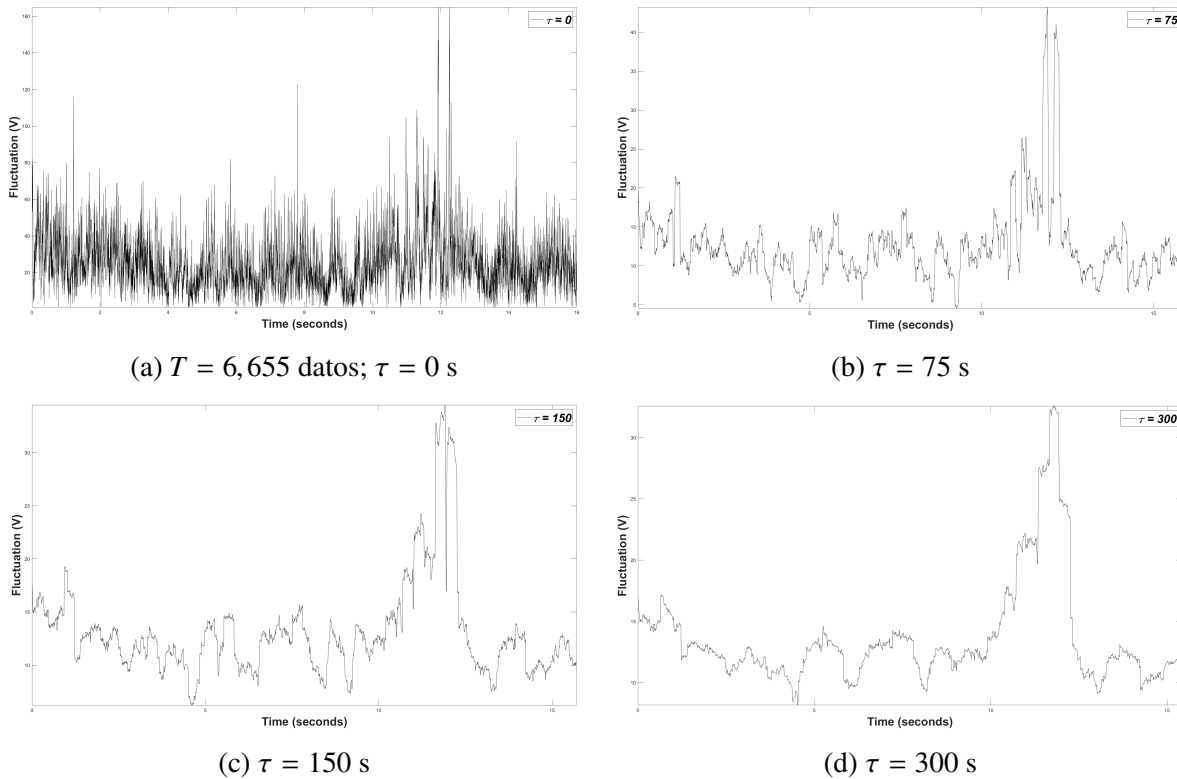


Figura 3: Series de tiempo de fluctuaciones $v(t, \tau)$, del canal F_{p1} con intervalos de tiempo, para la zona centro de la CDMX

Otro rasgo importante en el exponente de Hurst, es que éste se encuentra directamente relacionado con la dimensión fractal D , por la relación, $D = 2 - H$. En los resultados obtenidos, utilizando series de tiempo de datos de rendimientos logarítmicos $\delta(\tau)$, los valores del coeficiente H son bastante bajos, lo cual hace que se obtenga una dimensión fractal mayor, es decir, una traza más “rugosa” de acuerdo con la ecuación anterior. Mientras que los resultados, utilizando series de tiempo con valores absolutos de los rendimientos logarítmicos $|\delta(\tau)|$, muestran valores del coeficiente H bastante altos, con lo cual, la dimensión fractal D , disminuye, y la traza es “suave”, y por tanto menos rugosa. En cuanto a los valores del exponente de crecimiento β , se observa que estos valores son mayores que 0,5 para todas las zonas, sobre todo para la parte sur, norte y poniente y en menor medida para la parte centro y oriente. Lo anterior, indica que las fluctuaciones del sistema continúan creciendo en los diferentes intervalos de tiempo τ , como una ley de potencia, lo que sugiere que el

sistema en estudio, tarda mucho tiempo en asimilar la información que fluye hacia el mismo, pero parece que en algún intervalo de tiempo mayor al considerado, el sistema va a dejar de fluctuar. En la Figura 7, aparecen las series de fluctuaciones para diferentes intervalos de tiempo, en donde se aprecia que la volatilidad cambia día a día, de tal forma que las volatilidades históricas, a diferentes intervalos de tiempo, parecen ser similares. Resalta el hecho de que las fluctuaciones o volatilidades de los niveles de partículas $PM_{2,5}$ en un rango amplio de horizontes, posee una invarianza estadística autoafín, atribuida a los niveles altos de persistencia de su función de autocorrelación.

Las fluctuaciones de los valores absolutos se correlacionan con el tiempo o número de días utilizados en el análisis, dado que las fluctuaciones despliegan comportamientos cada vez más persistentes, conforme aumenta el plazo. Lo dicho hasta aquí, supone que las fluctuaciones a largo plazo, satisfacen la dinámica de escalamiento

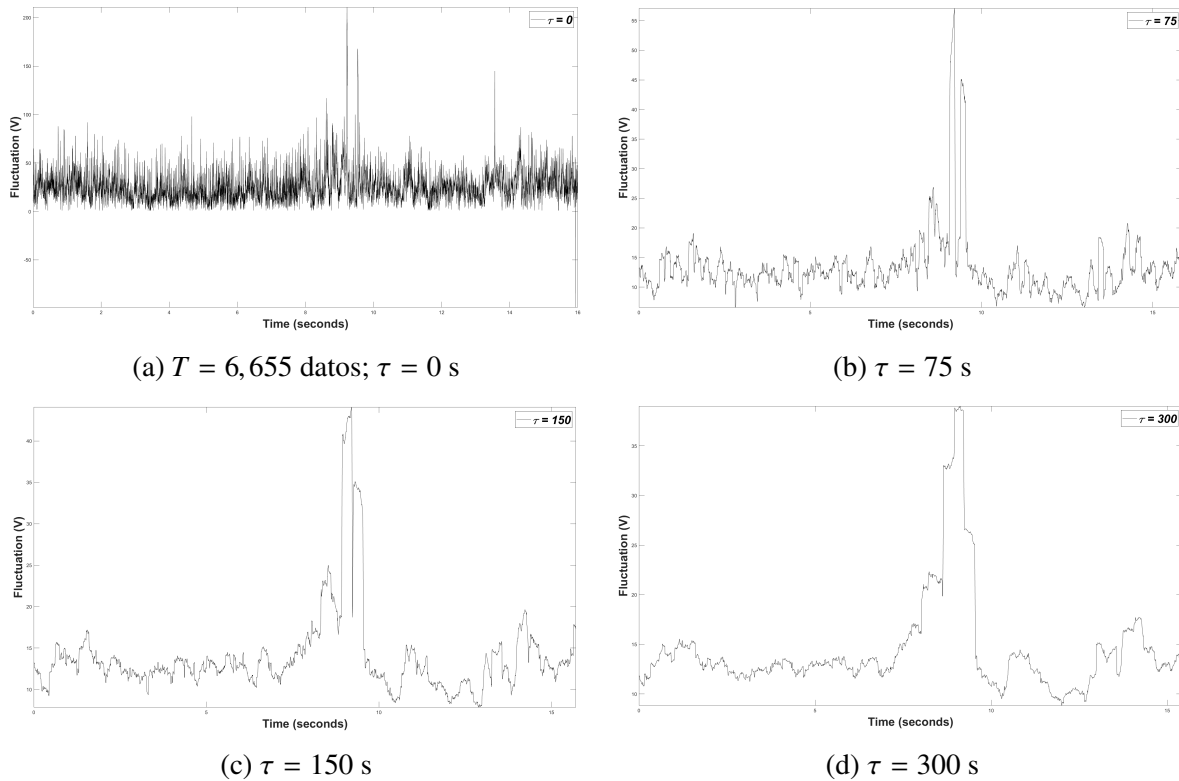


Figura 4: Series de tiempo de fluctuaciones $v(t, \tau)$, del canal F_{p1} con intervalos de tiempo, para la zona oriente de la CDMX

de la célebre ansatz de Family-Vicsek [44][45, 46, 47, 48], el cual se caracteriza por ser un sistema complejo fractal.

Para tener una visión más profunda de la dinámica de la volatilidad histórica en los niveles de $PM_{2.5}$ en la CDMX, también se realiza el análisis estadístico de las fluctuaciones o volatilidades. Por otra parte, también, se investiga la distribución de probabilidad de las series de fluctuaciones, en donde para pequeños intervalos de tiempo, los datos se ajustan mejor a las distribuciones Pearson5, valores extremos, Beta General, triangular y Gamma. En tanto, para fluctuaciones a largo plazo, los datos se ajustan a las distribuciones de valores extremos, Beta General y Gamma.

Como se puede observar en la Tabla 4, para la zona norte, la distribución de valores extremos, es la que más se ajusta a las fluctuaciones de corto y largo plazo, mientras que para las otras cuatro zonas, la distribución Beta General y Gamma, son las que mejor se ajustan a los diferentes plazos de

sus fluctuaciones.

Para predecir la evolución del sistema, se utilizó el software NumXL, el cual permite construir un modelo, con base al análisis de las series de tiempo y de las volatilidades o fluctuaciones. Dentro, del análisis de las series de tiempo, se calcula el promedio de los niveles de partículas $PM_{2.5}$ (alturas de los perfiles), en cada zona. Como se nota en la Figura 8, la zona norte es la que tarda más en mantenerse constante, ya que lo hace, a partir del promedio número 1,007, seguido de la zona oriente y poniente (ambas a partir del promedio número 860). En cambio, la zona centro y sur, a partir de la observación 810 y 734, alcanzan un estado constante en sus series.

Por otra parte, se determinó también, el momento a partir del cual, el promedio de $H|\delta(\tau)|$ que representa el valor absoluto de los rendimientos logarítmicos del exponente de Hurst de los niveles de partículas $PM_{2.5}$, se vuelve prácticamente constante. Para ello, se calculó primero, el valor absoluto de los rendimientos logarítmicos

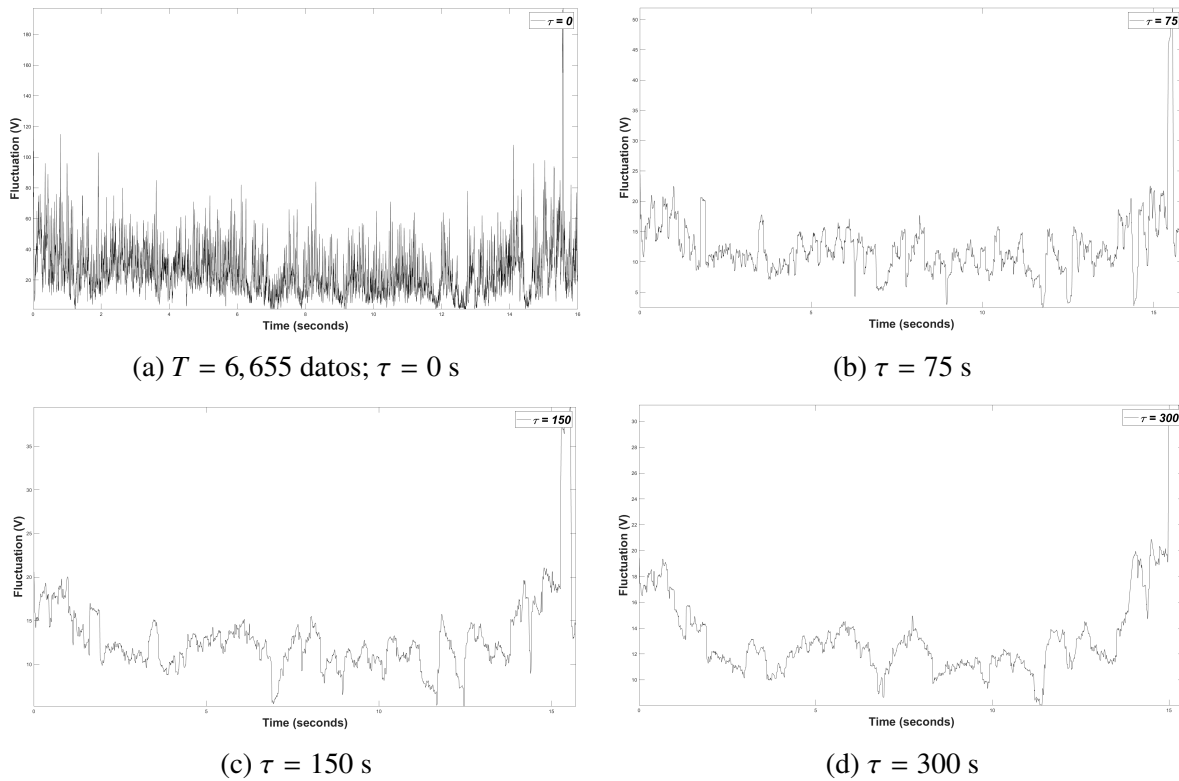


Figura 5: Series de tiempo de fluctuaciones $v(t, \tau)$, del canal F_{p1} con intervalos de tiempo, para la zona norte de la CDMX

Tabla 4: Distribuciones probabilísticas de las fluctuaciones de las series de tiempo de $PM_{2.5}$ en la CDMX, utilizando el software @Risk

	Norte	Sur	Centro	Oriente	Poniente
Fluctuaciones a corto plazo	Pearson 5 y Valores extremos	Beta General	Beta General y triangular	Gamma	Gamma y Beta General
Fluctuaciones a largo plazo	Valores extremos	Beta General	Beta General	Beta General	Gamma

$|\delta(\tau)|$, para diferentes horizontes de estudio $n = 2, 3, 4, \dots, 101$, y posteriormente, con el método de Rango Reescalado, se obtuvo $H|\delta(\tau)|$. En la Figura 8, se aprecia que el promedio de $H|\delta(\tau)|$, alcanza un estado estacionario distinto en cada zona. En la zona norte, se alcanza dicho estado, aproximadamente, en el dato 18, mientras que en el sur, se logra en el dato 38. En el caso de la zona centro, se da en el 26, en el oriente en el 52, y por último, para el poniente, en el 17. Esto se puede interpretar como los días que tardan las fluctuaciones en estabilizarse. Por tanto, el oriente, sur y centro, son las zonas que más tardan, para que su sistema alcance un estado

estacionario, en cambio en el poniente y norte, lo hacen en un periodo bastante corto. Los promedios del coeficiente de Hurst, desde el dato en que se satura el sistema, para la zona norte, sur, centro, oriente y poniente es de 0,73; 0,64; 0,72; 0,72 y 0,37; respectivamente.

Las proyecciones de los niveles de partículas $PM_{2.5}$ para un año (1,800 pasos), muestran que cada zona tendrá un comportamiento diferente. Sin embargo, tienen en común, que el sistema en cierto punto, se estabiliza. En el caso de la zona norte, centro y oriente, el comportamiento de sus fluctuaciones es muy similar, con una tendencia a la baja. En tanto, el sur, muestra

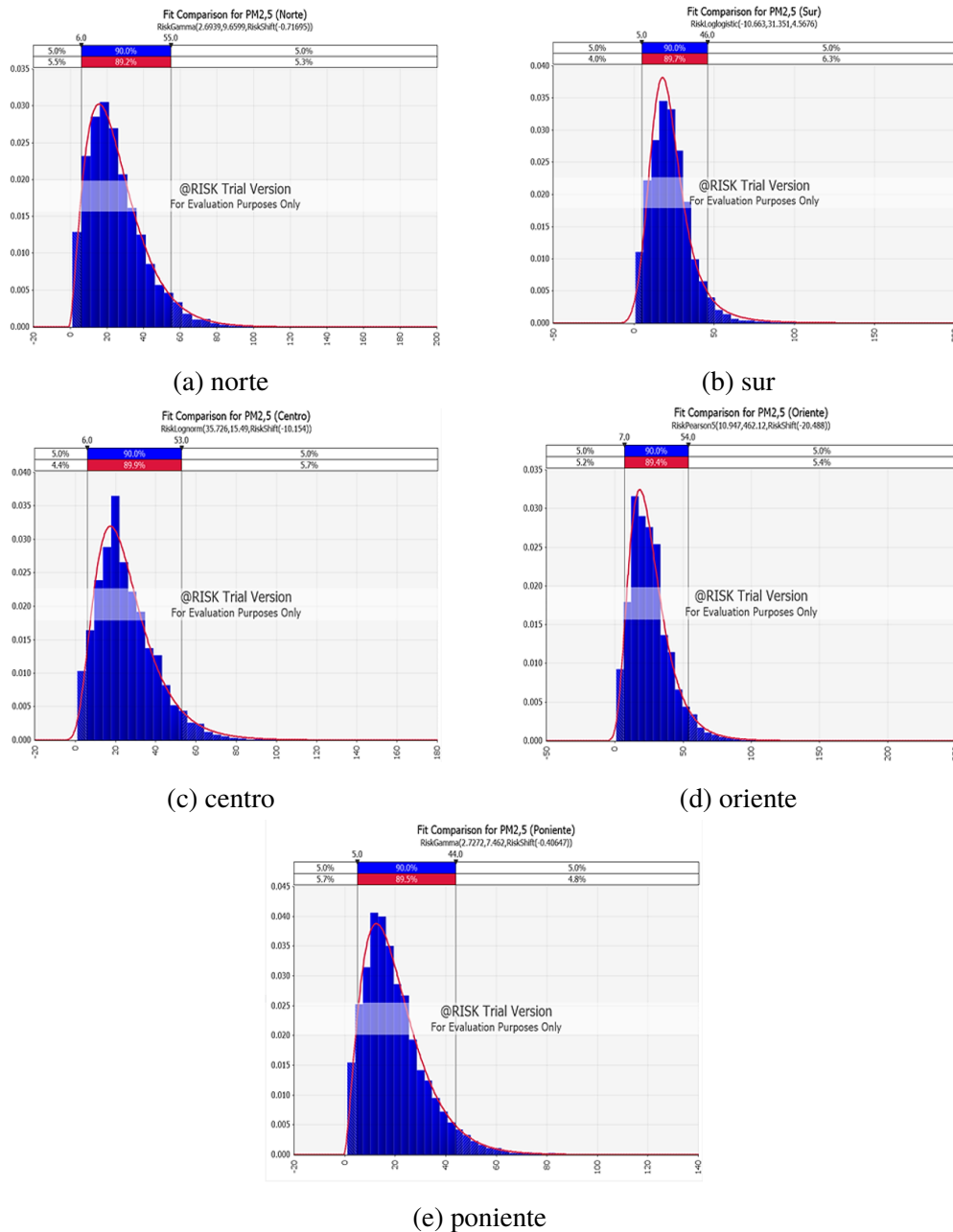


Figura 6: Ajuste de las distribuciones probabilísticas más adecuadas a los datos históricos para el nivel de PM_{2.5} en la CDMX utilizando el software @Risk

un decrecimiento exponencial, y el poniente un crecimiento temprano seguido de periodos de estabilidad (Figura 10).

Los resultados muestran que los niveles de PM_{2.5}, varían significativamente cuando se utilizan rendimientos logarítmicos, y cuando se utilizan este tipo de rendimientos, en sus valores absolutos. En el primer caso, el comportamiento de la serie es persistente (el sistema bajo estudio tiene media

estable), mientras que en el segundo caso, el comportamiento cambia a persistente, es decir, caracterizadas por efectos de memoria de largo plazo. Por su parte, los datos originales de las series de tiempo de las distintas zonas de la CDMX, se ajustan a diferentes distribuciones, además de que presentan un exceso de curtosis, lo que hace más probable que los retornos futuros sean extremadamente grandes o pequeños.

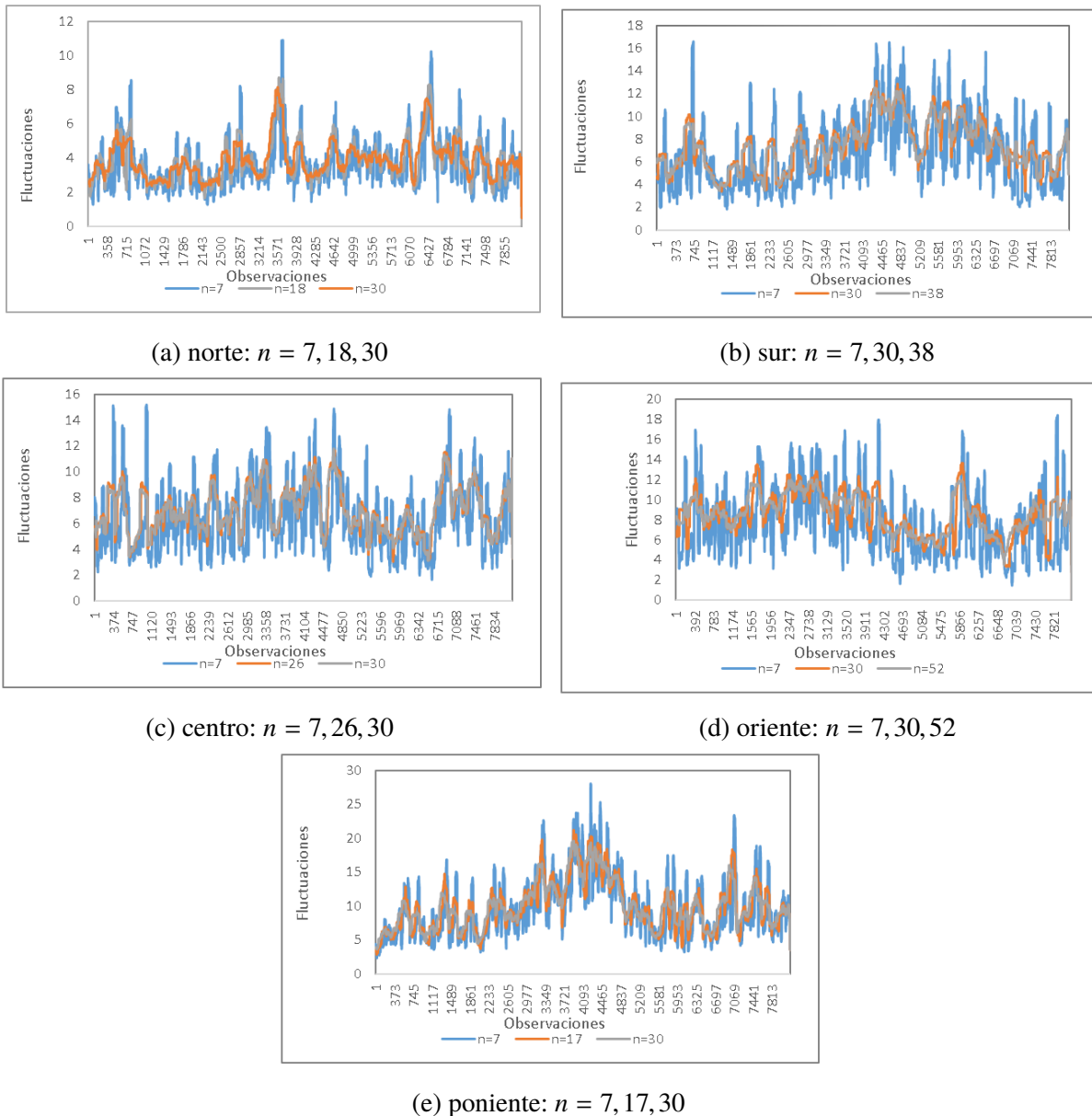


Figura 7: Volatilidades históricas para diferentes horizontes de tiempo de cada zona de estudio

En cuanto a las series de tiempo de fluctuaciones o volatilidades, estas despliegan comportamientos cada vez más persistentes conforme se alarga el horizonte de tiempo, dado los niveles significativos de persistencia en su función de autocorrelación. De igual forma, el análisis estadístico mostró que las fluctuaciones o volatilidades a diferentes intervalos, se ajustan a distribuciones de valores extremos para el caso del norte, mientras que para las zonas restantes, se ajustan mejor a la distribución Beta General y Gamma.

PM_{2.5}, utilizando el análisis de dinámica fractal, permiten conocer mejor la respuesta que puede tener el sistema, a lo largo de un horizonte de tiempo. En otras palabras, al conocer el momento en el que el sistema colapsa, nos brinda un periodo más exacto para pronosticar el comportamiento del sistema, y de esta manera, se puedan implementar a tiempo, mejores estrategias y controles para la disminución de dicho contaminante, sin tener que esperar a que se dé una contingencia ambiental para actuar.

Las proyecciones de los niveles de partículas

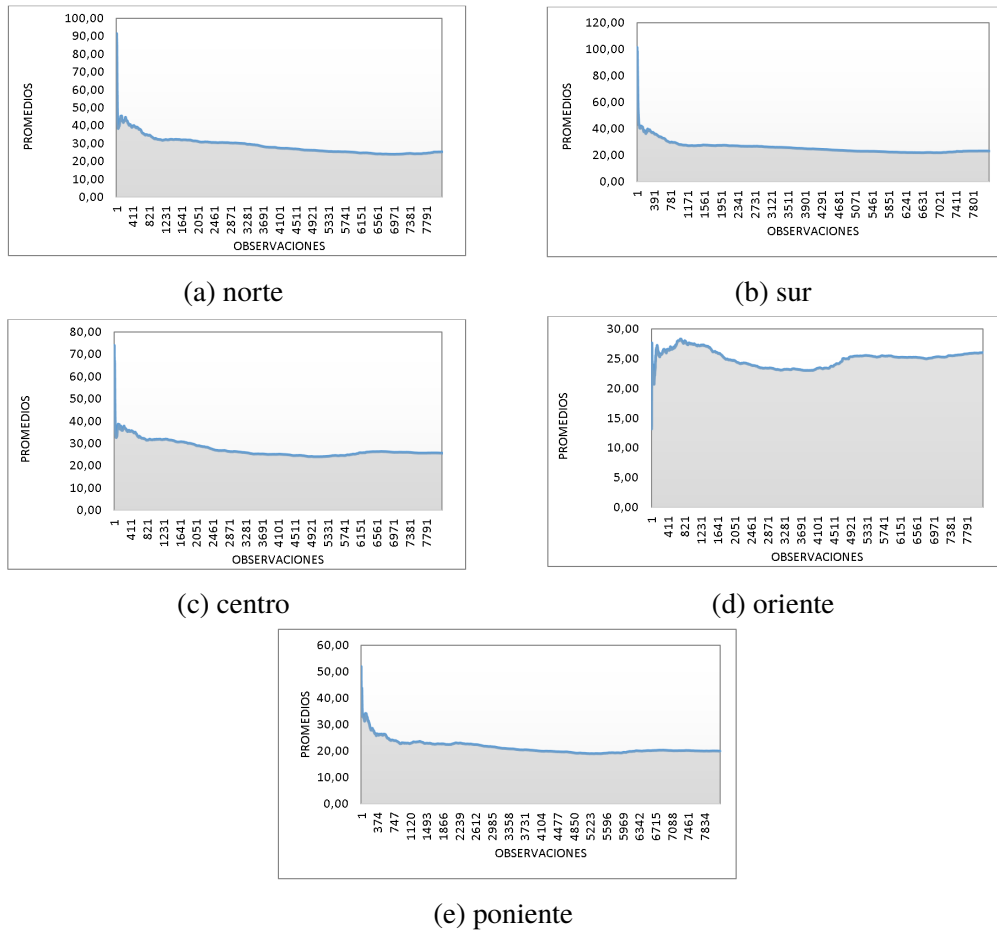


Figura 8: Promedio de los niveles de partículas $PM_{2,5}$ (alturas de los perfiles) en cada zona analizada en la CDMX

4. Conclusiones

Todo parece indicar, que los resultados de las diferentes zonas analizadas, muestran comportamientos muy similares en los niveles de partículas $PM_{2,5}$, por lo que la ciudad parece ser muy homogénea, ya que no hay un área que se diferencie demasiado de las demás, lo cual da la posibilidad de instrumentar de manera integrada, acciones tendientes a la gestión de riesgos a la salud por emisión de estas partículas.

El estudio de $PM_{2,5}$ implica una gran complejidad, ya que sus niveles de concentración varían en cada país o región. Hay una infinidad de variables que afectan las fuentes de aumento de las concentraciones de $PM_{2,5}$, y por tanto, la dinámica de sus series de tiempo, como lo son: las geográficas, climáticas, antropogénicas, meteorológicos, etc. Esto hace que los modelos de predicción de

las partículas $PM_{2,5}$, no aplique el supuesto de normalidad. Dada esta infinidad de fuentes no controladas, generan un comportamiento inusual en el sistema.

En el caso de la Ciudad de México, esto no es diferente, ya que los datos originales y las fluctuaciones de las series de tiempo de las distintas áreas analizadas, no se ajustan a una sola distribución, y muestran valores extremos, causados probablemente por todas estas fuentes externas y de las cuales no se puede tener ningún control, pero que al tener conocimiento del comportamiento de dicho sistema a lo largo del tiempo, se puede mitigar los riesgos a la salud, por emisión de estas partículas.

El caso específico de la Ciudad de México, es difícil comparar los resultados obtenidos, con otros estudios que abordan la misma problemática

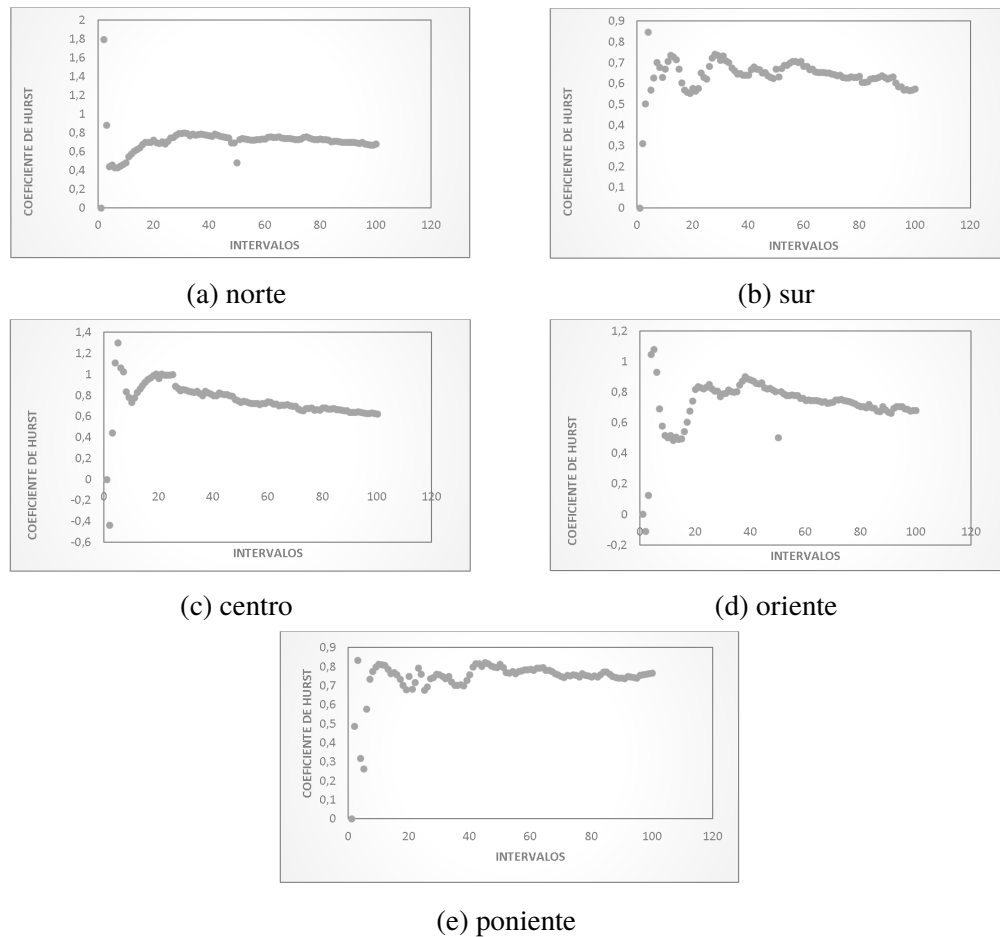


Figura 9: Promedios globales $H_{|\delta(\tau)|}$ del nivel de $PM_{2,5}$ para las diferentes zonas analizadas en la CDMX de cada zona de estudio, obtenidas con el software NumXL

de $PM_{2,5}$, ya que cada investigador, utiliza diferentes teorías, metodologías, modelos, métodos y técnicas. Asimismo, la gran mayoría de las investigaciones, se centran en caracterizar la composición química de dichas partículas, y así identificar las principales fuentes, que contribuyen con efectos adversos para la salud.

A pesar de lo anterior, trabajos como el de [39], que se centran en el comportamiento espacial y temporal, encuentran que hay una distribución bastante uniforme de $PM_{2,5}$ en el valle de México, lo cual, concuerda con los resultados, de la presente investigación. Asimismo, este autor encuentra la presencia de valores extremos ocasionados por hogueras y fuegos artificiales. Si bien, dentro de nuestros resultados se encontró la presencia de valores atípicos en las series analizadas, no se tuvo como objetivo descubrir, si estos hallazgos

se debían a fuentes biogénicas o antropogénicas. También, algo que coincidió con investigaciones, como la de Zhang, Ni & Ni (2016), es que las asimetrías se vuelven cada vez más débiles con el aumento del tamaño de las ventanas, lo cual indica cambios dinámicos en la correlación multiescala para series temporales $PM_{2,5}$. Al igual que [49], coincidimos en que el análisis fractal, es capaz de detectar la complejidad del conjunto de datos al escalar datos empíricos, ya que las series temporales por dimensión fractal, permiten definir los niveles de episodios de contaminación.

Al momento de construir las distintas series de tiempo, se descubrió que existen grandes diferencias diarias en la concentración media de $PM_{2,5}$ en el área de la Ciudad de México, lo cual coincide con lo expuesto por [41]. Sin embargo, estos autores incorporan predictores espaciales y

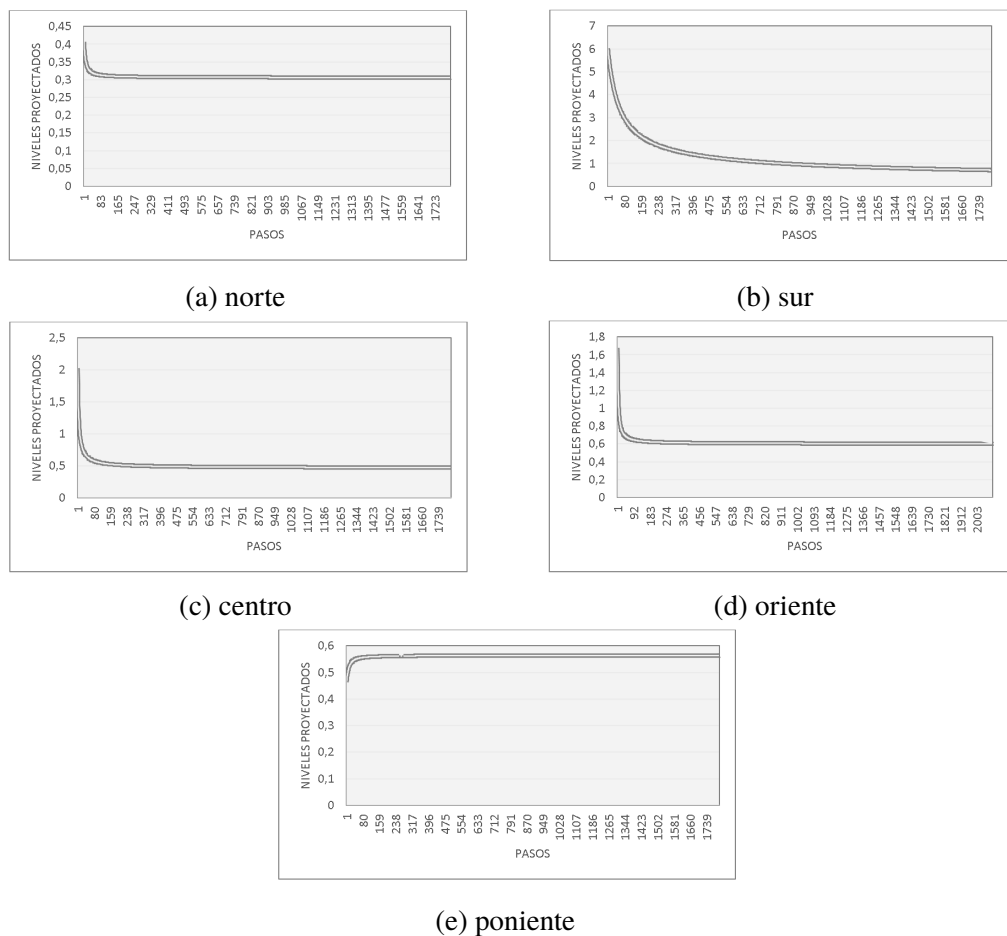


Figura 10: Proyecciones utilizando las volatilidades del nivel de $PM_{2.5}$ para las diferentes zonas analizadas en la CDMX, obtenidas con uso del software NumXL

temporales, junto con efectos aleatorios específicos del día, para tener en cuenta estas variaciones temporales. Por lo tanto, para posibles nuevas investigaciones, es importante tener en cuenta estas cuestiones, ya que como muestra [40], en sus hallazgos, las fuentes estacionarias podrían desempeñar un papel fundamental en el aumento de las concentraciones de $PM_{2.5}$.

La aportación más importante de este estudio es que se cumplió con el objetivo de caracterizar la dinámica de las partículas $PM_{2.5}$, en diversas zonas de la Ciudad de México, hallando distintos patrones, a diferentes escalas de tiempo-espacio, y sobre todo un comportamiento complejo en los niveles de $PM_{2.5}$. Lo antes descrito, permitió generar proyecciones a largo plazo relevantes, para la toma de decisiones en materia de política pública ambiental, además, el diseño de medidas

estratégicas de control. Algunas de ellas podrían ser educación (desde los niños) y la generación y uso de conocimiento científico (mayor inversión como porcentaje del PIB), fundamentalmente. También, se debe reconocer que es imperativo aprender a usar los recursos limitados del planeta. Esto último, ya se estimaba y advertía, desde inicios de los años 70, en el libro “Limits to Growth” [50].

La generalidad de literatura que aborda la problemática sobre los niveles de $PM_{2.5}$, sólo se concentra en caracterizar la composición química de dichas partículas, para así, identificar las principales fuentes que contribuyen efectos adversos para la salud. También, la mayoría utilizan modelos clásicos, y son pocos aquellos que se arriesgan a utilizar herramientas más sofisticadas y complejas. Por lo cual, en futuras investigaciones, sería interesante crear un modelo más robusto

que abarque, tanto un análisis gravimétrico y químico, como un análisis factorial, con series de tiempo, e incluso utilizar otras herramientas, como la simulación de valores extremos, o modelos evolutivos, como las redes neuronales y algoritmos genéticos.

5. Referencias

- [1] Organización Mundial de la Salud, “Calidad del aire y salud,” Organización Mundial de la Salud, Nota descriptiva, 2020.
- [2] Secretaría del Medio Ambiente y Recursos Naturales, “Estimación de impactos en salud por contaminación atmosférica en el región centro del país y alternativas de control,” Instituto Nacional de Salud Pública, México, Informe técnico, 2016.
- [3] Organización Panamericana de la Salud, “Calidad del aire,” Organización Panamericana de la Salud, Nota descriptiva, 2020.
- [4] M. Perevochtchikova, “La situación actual del sistema de monitoreo ambiental en la Zona Metropolitana de la Ciudad de México,” *Estudios demográficos y urbanos*, vol. 24, no. 3, pp. 513–547, 2009.
- [5] A. Navarro-Arredondo, “Control de la contaminación atmosférica en la Zona Metropolitana del Valle de México,” *Estudios demográficos y urbanos*, vol. 34, no. 3, pp. 631–663, 2019.
- [6] V. Mugica, F. Mugica, M. Torres, and J. Figueroa, “PM_{2.5} Emission Elemental Composition from Diverse Combustion Sources in the Metropolitan Area of Mexico City,” *The Scientific World Journal*, vol. 8, pp. 275–286, 2008.
- [7] R. D. Godoy, “Sources of Risk to Health by Exposure to PM_{2.5} in Mexico City,” *Epidemiology*, vol. 19, no. 1, pp. S223–S224, 2008.
- [8] Z.-Q. Jiang, W.-J. Xie, W.-X. Zhou, and D. Sornette, “Multifractal analysis of financial markets,” *Reports on Progress in Physics*, vol. 82, pp. 1–145, 2019.
- [9] F. Ramírez, “Comparación entre modelo browniano, a través de la metodología Box-JenKins y el análisis multifractal: Aplicación al sector financiero,” Tesis de Licenciatura, Universidad Autónoma del Estado de México, Estado de México, 2018.
- [10] B. Mandelbrot, *The Fractal Geometry of Nature*. New York, USA: Times Books, 1982.
- [11] X. Feng, Q. Li, Y. Zhu, J. Hou, L. Jin, and J. Wang, “Artificial neural networks forecasting of PM_{2.5} pollution using air mass trajectory based geographic model and wavelet transformation,” *Atmospheric Environment*, vol. 107, pp. 118–128, 2015.
- [12] X. Qiao, Q. Ying, X. Lid, H. Zhang, J. Hu, Y. Tang, and X. Chen, “Source apportionment of PM_{2.5} for 25 Chinese provincial capitals and municipalities using a source-oriented Community Multiscale Air Quality model,” *The Science of the total environment*, vol. 612, pp. 462–471, 2018.
- [13] L. Tang, T. Nagashima, K. Hasegawa, T. Ohara, K. Sudo, and N. Itsubo, “Development of human health damage factors for PM_{2.5} based on a global chemical transport model,” *The International Journal of Life Cycle Assessment*, vol. 23, pp. 2300–2310, 2018.
- [14] I. Kloog, B. Ridgway, P. Koutrakis, B. A. Coull, and J. D. Schwartz, “Long- and short-term exposure to PM_{2.5} and mortality: using novel exposure models,” *Epidemiology*, vol. 24, no. 4, pp. 555–561, 2013.
- [15] B. Liu, J. Wua, J. Zhang, L. Wang, J. Yang, D. Liang, Q. Dai, X. Bi, Y. Feng, Y. Zhang, and Q. Zhang, “Characterization and source apportionment of PM_{2.5} based on error estimation from EPA PMF 5.0 model at a medium city in China,” *Environmental Pollution*, vol. 222, pp. 10–22, 2017.
- [16] I. Kloog, A. A. Chudnovsky, A. C. Just, F. Nordio, P. Koutrakis, B. A. Coull, A. Lyapustin, Y. Wang, and J. Schwartz, “A new hybrid spatio-temporal model for estimating daily multi-year PM_{2.5} concentrations across northeastern USA using high resolution aerosol optical depth data,” *Atmospheric Environment*, vol. 95, pp. 581–590, Oct. 2014.
- [17] M. Eeftens, R. Beelen, K. de Hoogh, T. Bellander, G. Cesaroni, M. Cirach, C. Declercq, A. Dédèlè, E. Dons, A. de Nazelle, K. Dimakopoulou, K. Eriksen, G. Falq, P. Fischer, C. Galassi, R. Gražulevičienė, J. Heinrich, B. Hoffmann, M. Jerrett, D. Keidel, M. Korek, T. Lanki, S. Lindley, C. Madsen, A. Mölter, G. Nádor, M. Nieuwenhuijsen, M. Nonnemacher, X. Pedeli, O. Raaschou-Nielsen, E. Patelarou, U. Quass, A. Ranzi, C. Schindler, M. Stempfelet, E. Stephanou, D. Sugiri, M.-Y. Tsai, T. Yli-Tuomi, M. J. Varró, D. Vienneau, S. v. Klot, K. Wolf, B. Brunekreef, and G. Hoek, “Development of Land Use Regression Models for PM_{2.5}, PM_{2.5} Absorbance, PM₁₀ and PM_{coarse} in 20 European Study Areas; Results of the ESCAPE Project,” *Environmental Science & Technology*, vol. 46, pp. 11 195–11 205, 2012.
- [18] X. Hu, L. A. Waller, A. Lyapustin, Y. Wang, M. Z. Al-Hamdan, W. L. Crosson, M. G. Estes Jr, S. M. Estes, D. A. Quattrochi, S. J. Puttaswamy, and Y. Liu, “Estimating ground-level PM_{2.5} concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model,” *Remote Sensing of Environment*, vol. 140, pp. 220–232, 2014.
- [19] D. J. Nowak, S. Hirabayashi, A. Bodine, and R. Hoehn, “Modeled PM_{2.5} removal by trees in ten U.S. cities and associated health effects,” *Environmental Pollution*, vol. 178, pp. 395–402, 2013.
- [20] K. de Hoogh, H. Hérítier, M. Stafoggia, N. Künzli, and I. Kloog, “Modelling daily PM_{2.5} concentrations at high spatio-temporal resolution across Switzerland,” *Environmental Pollution*, vol. 233, pp. 1147–1154, 2018.

- [21] C. Zhang, Z. Ni, L. Ni, J. Li, and L. Zhou, "Asymmetric multifractal detrending moving average analysis in time series of PM_{2,5} concentration," *Physica A: Statistical Mechanics and its Applications*, vol. 457, pp. 322–330, 2016.
- [22] P. Perez and G. Salini, "PM_{2,5} forecasting in a large city: Comparison of three methods," *Atmospheric Environment*, vol. 42, no. 35, pp. 8219–8224, 2008.
- [23] C. Hart, *Doing a Literature Review: Releasing the Research Imagination*. University of Chester, UK: Sage Publications Ltd, 2018.
- [24] J. Jesson, L. Matheson, and F. M. Lacey, *Doing Your Literature Review: Traditional and Systematic Techniques*. University of Chester, UK: SAGE Publications, 2011.
- [25] L. A. Machi and B. T. McEvoy, *The Literature Review: Six Steps to Success*. University of Chester, UK: SAGE Publications, 2009.
- [26] D. Ridley, *The Literature Review: A Step-by-Step Guide for Students*. University of Chester, UK: SAGE Publications, 2012.
- [27] A. Serenko and N. Bontis, "Investigating the current state and impact of the intellectual capital academic discipline," *Journal of Intellectual Capital*, vol. 14, no. 4, pp. 476–500, 2013.
- [28] P. Creme and M. R. Lea, *Escribir en la universidad CDMX*. México: Gedisa Mexicana, 2017.
- [29] M. Walker, *Cómo escribir trabajos de investigación*. México: Gedisa Maexicana, 2007.
- [30] X. Liu, J. Li, Y. Qu, T. Han, L. Hou, J. Gu, C. Chen, Y. Yang, X. Liu, T. Yang, Y. Zhang, H. Tian, and M. Hu, "Formation and evolution mechanism of regional haze: a case study in the megacity Beijing, China," *Atmospheric Chemistry and Physics*, vol. 13, no. 9, pp. 4501–4514, 2013.
- [31] E. Vega, E. Reyes, H. Ruiz, J. García, G. Sánchez, G. Martínez-Villa, U. González, J. C. Chow, and J. G. Watson, "Analysis of PM_{2,5} and PM₁₀ in the Atmosphere of Mexico City during 2000-2002," *Journal of the Air & Waste Management Association*, vol. 54, no. 7, pp. 786–798, 2004.
- [32] E. Vega, E. Reyes, G. Sánchez, E. Ortiz, M. Ruiz, J. Chow, J. Watson, and S. Edgerton, "Basic statistics of PM_{2.5} and PM₁₀ in the atmosphere of Mexico City," *Sci Total Environ*, vol. 287, no. 3, pp. 167–176, Mar. 2002.
- [33] J. C. Chow, J. G. Watson, S. A. Edgerton, and E. Vega, "Chemical composition of PM_{2,5} and PM₁₀ in Mexico City during winter 1997," *Science of The Total Environment*, vol. 287, no. 3, pp. 177–201, Mar. 2002.
- [34] M. A. Martínez, P. Caballero, O. Carrillo, A. Mendoza, and G. M. Mejía, "Chemical characterization and factor analysis of PM_{2.5} in two sites of Monterrey, Mexico," *Journal of the Air & Waste Management Association*, vol. 62, no. 7, pp. 817–827, 2012.
- [35] J. E. Gómez-Perales, R. N. Colvile, M. J. Nieuwenhuijsen, A. Fernández-Bremauntz, V. J. Gutiérrez-Avedoy, V. H. Páramo-Figueroa, S. Blanco-Jiménez, E. Bueno-López, F. Mandujano, R. Bernabé-Cabanillas, and E. Ortiz-Segovia, "Commuters' exposure to PM_{2,5}, CO, and benzene in public transport in the metropolitan area of Mexico City," *Atmospheric Environment*, vol. 38, no. 8, pp. 1219–1229, 2004.
- [36] G. Li, W. Lei, N. Bei, and L. T. Molina, "Contribution of garbage burning to chloride and PM_{2,5} in Mexico City," *Atmospheric Chemistry and Physics*, vol. 12, no. 18, pp. 8751–8761, 2012.
- [37] R. V. Díaz and E. R. Dominguez, "Health risk by inhalation of PM_{2,5} in the metropolitan zone of the City of Mexico," *Ecotoxicol Environ Saf.*, vol. 72, no. 3, pp. 866–871, 2009.
- [38] J. B. Ordieres, E. P. Vergara, R. S. Capuz, and R. E. Salazar, "Neural network prediction model for fine particulate matter (PM_{2,5}) on the US–Mexico border in El Paso (Texas) and Ciudad Juárez (Chihuahua)," *Environmental Modelling & Software*, vol. 20, no. 5, pp. 547–559, May 2005.
- [39] O. Amador-Muñoz, R. Villalobos-Pietrini, J. Miranda, and L. E. Vera-Avila, "Organic compounds of PM_{2,5} in Mexico Valley: Spatial and temporal patterns, behavior and sources," *Science of The Total Environment*, vol. 409, no. 8, pp. 1453–1465, 2011.
- [40] D. López-Veneroni, "The stable carbon isotope composition of PM_{2,5} and PM₁₀ in Mexico City Metropolitan Area air," *Atmospheric Environment*, vol. 43, no. 29, pp. 4491–4502, 2009.
- [41] A. C. Just, R. O. Wright, J. Schwartz, B. A. Coull, A. A. Baccarelli, M. M. Tellez-Rojo, E. Moody, Y. Wang, A. Lyapustin, and I. Kloog, "Using High-Resolution Satellite Aerosol Optical Depth To Estimate Daily PM_{2,5} Geographical Distribution in Mexico City," *Environmental Science & Technology*, vol. 49, no. 14, pp. 8576–8584, 2015.
- [42] Gobierno de la Ciudad de México, "Calidad del aire. Partículas suspendidas," Secretaría del Medio Ambiente. Dirección General de Calidad del Aire, México, Reporte técnico, 2019.
- [43] O. Morales, "Modelos mecánicos de la dinámica fractal del mercado petrolero," Tesis doctoral, Instituto Politécnico Nacional-Escuela Superior de Ingeniería Macánica y Eléctrica, Ciudad de México, México, 2004.
- [44] O. Morales, A. Balankin, and L. M. Hernández, "Metodología de predicción de precios del petróleo basada en dinámica fractal," *Científica*, vol. 9, pp. 3–11, 2005.
- [45] F. Family and T. Vicsek, "Scaling of the active zone in the Eden process on percolation networks and the ballistic deposition model," *Journal of Physics A: Mathematical and General*, vol. 18, pp. L75–L81, 1985.
- [46] T. Vicsek and F. Family, "Dynamic Scaling for

- Aggregation of Clusters,” *Physical Review Letters*, vol. 52, pp. 1669–1672, 1984.
- [47] T. Vicsek, *Fractal Growth Phenomena*. Singapur: World Scientific Publishing Company, 1992.
- [48] F. Family and T. Vicsek, *Dynamics of Fractal Surfaces*. Singapur: World Scientific Publishing Company, 1991.
- [49] V. Evagelopoulos, S. Zoras, A. Triantafyllou, and T. Albanis, “PM₁₀ -PM_{2,5} time series and fractal analysis,” *Global NEST Journal*, vol. 8, no. 3, pp. 234–240, 2006.
- [50] D. H. Meadows, J. Randers, and D. L. Meadows, *Limits to Growth: The 30-Year Update White River Junction*. Vermont, USA: Chelsea Green Publishing, 2004.

Decrease in the activity of irradiated graphite and liquid radioactive waste

Valerii V. Krymsky *  , Natalya V. Plotnikova 

South Ural State University, Chelyabinsk, Russian Federation.

Abstract.- Nowadays, the problem of nuclear waste deactivation is very urgent. The method proposed consists of exposing the aqueous solution of radionuclide to powerful nanosecond electromagnetic pulses by placing the treated solution between two electrodes connected to the outlets of the generator. Single-pole current pulses with a length of 1 ns, an amplitude of more than 5 kV, and a repetition frequency of 1 kHz are used. Such pulses can be obtained from the generator units with an amplitude of 5 kV to 15 kV. They have a pulse capacity of 1 to 4 MW, and the mains capacity is below 50 W. The pulse action leads to radiolysis of water and accelerated decay of radionuclides. As a result of radiolysis, hydrated electrons, hydrogen atoms, and various radicals are formed that causes precipitation of radionuclides, such as ^{137}Cs and ^{90}Sr . The description of several experiments and results of experimental investigations of the influence of powerful nanosecond electromagnetic pulses on the properties of water solutions with radioactive nuclides ^{137}Cs and ^{90}Sr and on items of irradiated graphite is presented. Under the influence of pulses on real liquid waste and aqueous solutions, radioactive nuclides in water are reduced by 5 to 50 times during 15 minutes. The effect described is preserved for 20 days. The reduction of beta activity twice during 25 min is observed in big blocks of the irradiated graphite. Flow-through unit for processing of liquid waste is developed, tested, and patented (RU 2726145. MPK G21F 9/28, July 7, 2020). This technology can be used to treat emergency solutions with tritium at Fukushima NPP. of the irradiated graphite.

Keywords: nanosecond electromagnetic pulses; graphite; radiation; radioactive liquid.

Disminución de la actividad del grafito irradiado y los residuos radioactivos líquidos

Resumen.- Actualmente, el problema de la desactivación de residuos radioactivos es muy urgente. El método propuesto consiste en exponer la solución acuosa de radionúclido a potentes pulsos electromagnéticos de nanosegundos colocando la solución tratada entre dos electrodos conectados a las terminales del generador. Se aplican pulsos de corriente de polo único con una longitud de 1 ns, amplitud de más de 5 kV y frecuencia de repetición de 1 kHz. Dichos pulsos se pueden obtener de las unidades generadoras con una amplitud de 5 a 15 kV. Tienen una capacidad de pulso de 1 a 4 MW y la capacidad de la red debajo de 50 W. Los pulsos causan la radiólisis del agua y la descomposición rápida de radionúclidos. La radiólisis, causa la generación de electrones hidratados, átomos de hidrógeno y varios radicales que provocan la precipitación de radionúclidos, como ^{137}Cs y ^{90}Sr . Se describen las pruebas y se presentan los resultados experimentales de la influencia de fuertes pulsos electromagnéticos de nanosegundos en las propiedades de las soluciones acuosas con los núclidos radioactivos de ^{137}Cs y ^{90}Sr y en piezas de grafito irradiado. La influencia por 15 minutos de los pulsos en los residuos líquidos reales y soluciones acuosas reduce el tamaño de los núclidos en el agua en 5 a 50 veces. El efecto descrito permanece por veinte días. Se observa el doble de la disminución de actividad beta por 25 minutos en trozos grandes de grafito irradiado. Se desarrolla, ensaya y patenta la unidad de flujo (RU 2726145. MPK G21F 9/28, 2020). Se puede usar esta tecnología para tratar soluciones de emergencia con tritio en la central nuclear de Fukushima.

Palabras clave: pulsos electromagnéticos de nanosegundos; grafito; radiación; líquido radioactivo.

Received: January 23, 2021.

Accepted: April 01, 2021.

1. Introduction

Nowadays, a promising trend in nuclear waste deactivation is the use of powerful energy depositions. Urutskoev *et al.* [1] is worth highlighting as significant work on the subject. The

*Correspondence author:
*e-mail:*krymskiivv@susu.ru (V. Krymsky)

influence of high pulse current upon the electric explosion of titanium foil in distilled water is described: a capacitor bank with charging voltage $U = 4,8$ kV and energy $W = 50$ kJ is discharged on a sheet of titanium foil. The foil is placed in a polyethylene explosion chamber shaped like a torus with openings evenly bored circumferentially, where 30 to 50 mL of liquid was poured. An increase followed the explosion in the content of Mn, Fe, Ni, Cu, Zn in the samples; plus, such foreign elements were found in the samples as Na, Al, K, Si, Ca, Cl, Mo, Ag, In, Sn, Sb.

The absolute number of atoms of the elements generated in the samples after the electric explosions ranged from 10^{15} to 10^{18} .

The check tests conducted at the Joint Institute for Nuclear Research (Dubna) Kuznetsov *et al.*, [2] confirmed the main test findings.

The electric explosion of titanium foil in uranyl sulfate solution distorts the initial isotopic ratio of U. The maximal distortion of isotopic composition is observed in a vapor gas release in favor of enriching ^{235}U .

Nigmatulin *et al.* [3] describes the joint action of ultrasound oscillations and a powerful flow of neutrons on the acetone from deuterium. The test results can be rendered as the conversion of deuterium to tritium under the action of cavitation.

A. F. Kladov [4] proposes a method of making radioactive nuclides less active by mechanical activation. The required oscillation intensity must exceed 1 MW/m^2 . The laboratory setup consisted of a rotary hydrodynamic generator of sound and was powered by a 15 kW motor. The setup's oscillation intensity is $5,9 \times 10^3$ Hz, its operational pressure is 1 MPa, and the active treatment zone size is 250 cm^3 . A detailed description of the activator's design is provided in [4].

This setup was used to test the chloride solution of ^{137}Cs . In the initial test phase, the setup's sorption characteristic was checked. The contribution of the sorption component did not exceed 1 %. The decrease in the solution's activity for the 20 hours of the setup's operation exceeded 60 %.

Jung *et al.* [5] presents the test results of the beta decay under the influence of the external fields of

absolutely stable atoms ^{163}Dy and ^{193}Ir , the half-life of which is also determined.

Andreev and Shafeev [6] presents a detailed description of the effect of laser beams on solutions of radioactive nuclides. The method uses a laser beam to irradiate an absorbing target placed in a transparent water solution with a radioactive nuclide. The gold nanoparticles pushed out of the metal target migrate to the water and speed up the decay of radioactive nuclides. The experiment involved studying the ^{137}Cs and ^{238}U radioactive nuclides, and a golden plate was used as the target. The laser's energy was 2 mJ per pulse, the pulse duration was 10 ns, the repeating frequency was 10 kHz, and the exposure time was 10 hours. The laser beam was used to scan across the tray with the solution of ^{137}Cs .

In the time of exposure, the concentration of ^{137}Cs goes down by 5 % and remains unchanged for 15 hours. The reduction in the concentration for exposure time is formally equal to the 1 200-fold acceleration of decay. This method can be used only with small quantities of solution.

The provided examples show that the isotopic properties of elements can change under an external energy influence.

Nowadays, NPP units with uranium-graphite channel-type reactors are being removed from operation all around the world. The issue of discarding irradiated graphite complicates this process. The amount of irradiated graphite in one Russian uranium-graphite channel-type reactor reaches 1.798 tones. The total amount of irradiated graphite in Russia and worldwide is approximately 60.000 and 250.000 tons, respectively [7]. Irradiated graphite is referred to as class 2 and 3 radioactive waste (RW). The induced activity of irradiated graphite is determined mainly by ^{14}C , ^{36}Cl , ^{60}Co , and ^3H . The share of activity of carbon ^{14}C in the overall activity is 95 %. The intensity of its activity in irradiated graphite reaches 108 - 109 Bq/kg. This radioactive nuclide is a source of beta-radiation, and its half-life is 5.700 years.

In addition to the foregoing radioactive nuclides, irradiated graphite also contains small quantities of ^{235}U (fuel), such elements of its transmutation

as ^{238}Pu , ^{241}Am , and such products its fission as ^{134}Cs , ^{137}Cs . An element especially hazardous in the burial of waste is ^{94}Nb . It follows from the foregoing that the weakening in irradiated graphite activity is a relevant task at the global level.

Many patents describe the existing processing methods spent on graphite materials from nuclear reactors ([8, 9, 10, 11]. The main weaknesses of these methods are complex equipment, high energy costs, processing duration, absence of concrete results.

There are other non-electric methods [12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23] for reducing the content of radioactive cesium and strontium in aqueous solutions.

All these methods involve the use of additional chemicals. For example, the effectiveness of potable water treatment processes that consist of the stages of coagulation-flocculation-decantation using iron-based coagulants was studied [12]. It was found that the resulting decontamination depends on the chemical behavior of each of the radionuclides considered, on the pH at which the process of coagulation is carried out, and on the concentration of the other stable cations present.

It has been determined that radioactive ions in water can be readily immobilized onto the getter surface by forming precipitates due to the high local concentrations of precipitants on the getters [13]. The secured ionic pollutants can be subsequently removed from water by filtration or sedimentation for safe disposal.

Granular and powdered activated carbon was used to remove radioactive iodine and cesium [22]. The efficiency was approximately 30 %–40 % in some cases.

The methods described in [24, 25] allow for efficient external influence on radioactive nuclides but are characterized by high energy costs and processing periods.

The work examines the influence of nanosecond electromagnetic pulses (NSEMPs) on the activity of water solutions of radionuclides and irradiated reactor graphite without the use of additional reagents.

2. Materials and methods

2.1. Equipment used

Method of NSEMPs exposing on solutions consisted in placing processed solution between two electrodes, which were connected to the outlets of the generator. The construction of the unit is shown in Figure 1.

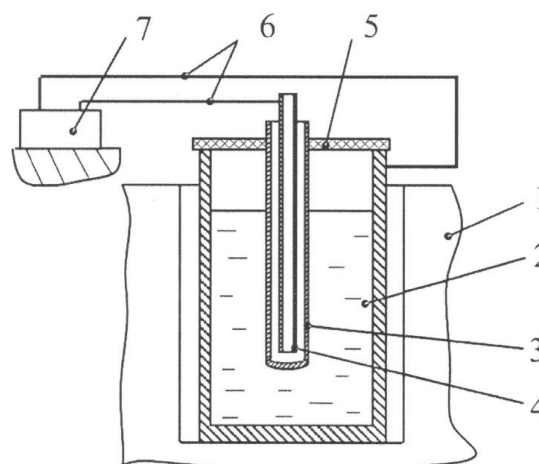


Figure 1: Unit construction: 1 - housing for glass; 2 - current-conducting glass with solution; 3 - insulating tube; 4 - rod; 5 - cover, 6 - supply cable; 7 - pulse generator

The emitters used to test the liquids are described below. The rod-in-the-glass emitter had a round copper rod or a rod with a silver coating. This rod was submerged right in the jar filled with liquid or in an isolated test tube. The rod was connected to the outlet of the center core of the generator's cable. In the case of a metal jar, the cable braiding outlet was connected right to the jar; in case of a glass jar, it was placed into a copper foil shield. The shield was connected to the cable braiding outlet. To ensure more efficient performance, the rod's and jar's diameters had to be correlated in a particular proportion depending on the properties of the irradiated liquid. This emitter was similar to a coaxial line.

Large quantities of liquid were exposed to radiation on a flow-through unit with nozzles. Similarly to the previous one, this unit was shaped as a coaxial line and had the following dimensions:

the inside rod diameter was 150 mm; the outside jar diameter was 400 mm; the unit height was 2000 mm. The generator's outlets were connected to the rod and the jar. The four nozzles were located right between them. The processed solution with a radioactive nuclide was pumped bottom-up to the nozzles under pressure.

When the generator was switched on, an electromagnetic pulse field was generated between the jar and the rod and affected the pulverized liquid. The unit required thoroughly purifying the water solution of solids for ensuring the sustainable performance of the nozzles.

The flow-through unit for processing water with electric pulses is described in RF patent 2531814 and consists of two coaxial copper cylinders. The nonconductive tube (hose), along which the radioactive solution flows, is coiled around the inner cylinder. Each of the cylinders is connected to one of the generator's outlets. The solution feed rate is adjusted using an external pump or flow control valve. This unit complies with the maintenance of staff safety requirements. The construction of the unit is shown in Figure 2.

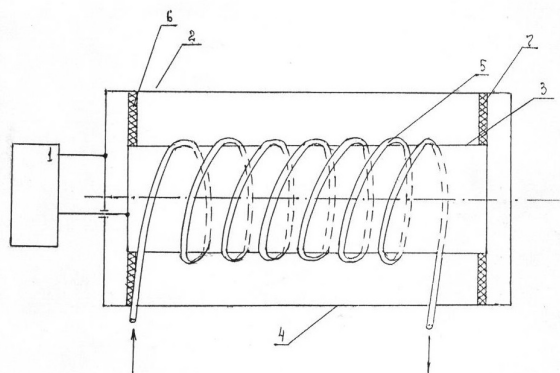


Figure 2: Construction of the flow-through unit: 1 - the pulse generator; 2 - solution exposure element in assembly ; 3 and 4 - coaxial copper cylinders; 5 - dielectric tube (hose); 6 and 7 - dielectric rings

The unit contains shut-off and control valves. The structure of the installation is shown in Figures 3 and 4. An important feature of research was the use of single-pole positive or negative current pulses. A current pulse was shaped like a triangle with a short forefront.

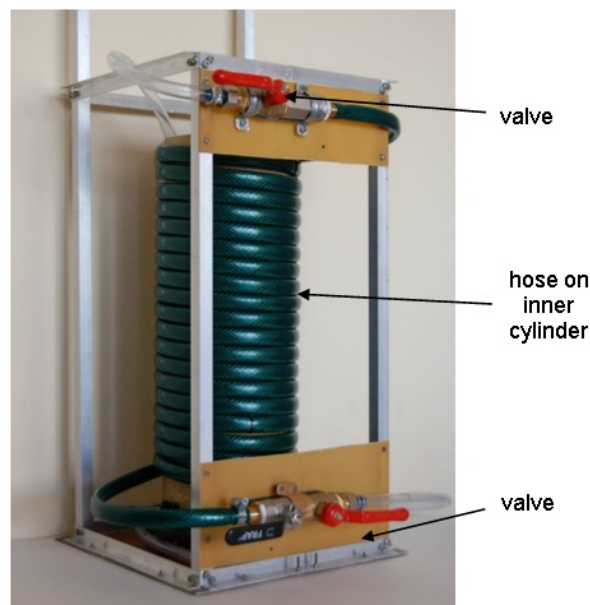


Figure 3: Unit without external cylinder

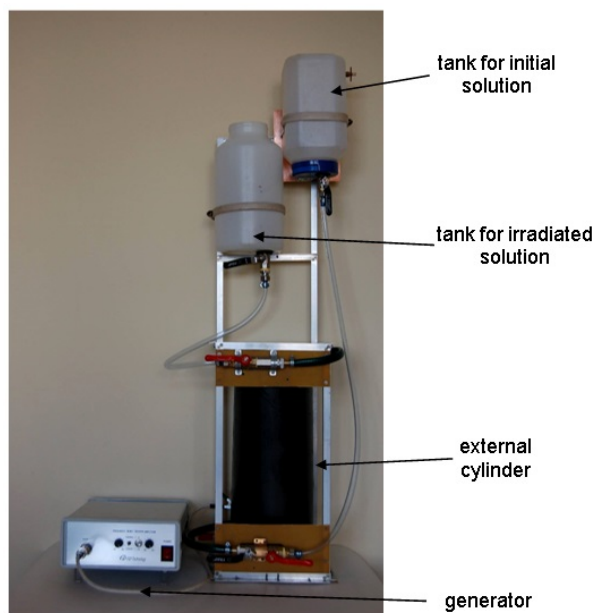


Figure 4: Experimental unit assembly

The principle of operation of the generators used is described in the patent [26]. The generators used in these tests had single-pulse energy of up to 10^{-2} J. That was why the exposure to NSEMPs was not followed by such energy effects as a significant rise in temperature. However, the pulse capacities were even higher than 10 MW.

The main generators used in the tests were the two generator units at the Budker Institute

of Nuclear Physics: GNP with an amplitude of 5 kV and GNI (LEGN) with an amplitude of 6 kV. Other generators used were the ones made by Feed Tekhnologiya company from Saint-Petersburg; they produced pulses with an amplitude of 10 or 15 kV, length of 1 ns, the forefront of 0,1 ns, and repeating frequency of 1 kHz. A distinct feature of all these generator units was their low mains power of up to 50 W.

The main condition for the repeatability of experiments was the use of a similar generator.

2.2. Materials used

The model solutions were prepared using standard elements ^{90}Sr and ^{137}Cs , which were dissolved in various acids. The amount (concentration) of the elements to be dissolved was selected based on the ability to measure their activity most accurately. Real radioactive waste was simply poured into a glass.

In the reactor graphite processing, the test object was a $120 \times 195 \times 325$ mm block of graphite taken from an IVV-2M reactor, graphite of VVG grade. The block's butt surfaces were connected to two copper foil electrodes. The electrodes were connected to the outlet of the generator. When the generator was turned on, a very large pulse current passed through the unit since graphite has a small resistance.

The materials used, the initial and the experimental conditions are given below for each specific test.

2.3. Conditions of experiments

The tests for exposing the radioactive nuclides to NSEMPs were conducted in the radiation control lab of the Chelyabinsk branch of the Ural Territorial District division of the RosRAO FSUE (RosAtom). The lab is certified in the Rosakkreditatsiya system. The solution activity was measured using a GAMMA-1P spectrometrical unit with a DGDK-150V germanium diffusion drift detector, a SEA-13P alpha spectrometrical system with a D30A silicon detector, and a UMF-2000 low-background radiometrical unit. The sampling was made according to the standard documents in effect [27].

The characteristic features of taking the samples and the procedures of measuring their activity will be indicated for each of the conducted tests. The general scheme of the experiments included the following phases: preparation of initial solutions, measurement of their activity, pulse processing, repeated measurement of solution activity.

3. Results and discussion

3.1. Processing solutions with ^{90}Sr

The results of processing the model solutions with ^{90}Sr are presented below. The ^{90}Sr element is a radioactive nuclide with the main beta decay and half-life of 28 years. It is always detected together with ^{90}Y ; to record it in a simpler way; however, we shall more frequently denote it as only ^{90}Sr .

Test 1. The model solutions of ^{90}Sr were bombarded with NSEMPs at various pH in a plastic vessel. The pulses were emitted using the rod-in-the-jar unit. The jar diameter and height were 105 and 130 mm, respectively, the irradiated solution volume was 0,5 L. The rod was a CuZn tube of 6 mm in diameter and 140 mm in length. The generator unit used was GNP, the pulse amplitude was 5 kV, and the pulse repeating frequency was 1 kHz. The exposure continued for 20 minutes. The pH value was set at the beginning of the test. The pH measurement was performed using instrument PH150MI. The pH value was reduced and increased by adding HNO_3 and, respectively, NaOH.

The solution activity was measured by the standard technique. The sample activity measurement results are shown in Table 1. It follows from Table 1 that the specific activity of the solution of ^{90}Sr is much weaker at high pH.

Test 2. The solutions were irradiated in a flow-through unit. Three and a half liters of the strontium nitrate solution with pH = 6 were prepared. The solution was poured into a consumption tank, and then the generator was switched on. The generator used was the one made by Feed Tekhnologiya company and had an amplitude of 15 kV, a length of 0,5 ns; and a repeating frequency of 1 kHz.

The shutoff valve was opened on the consumption tank, and the solution was supplied

Table 1: The specific activity of ^{90}Sr , $N-N_{\text{bckgrn}}$, cps

Solution	Background	Initial	Exposure pH=1	Exposure pH=4	Exposure pH=8	Exposure pH=12
β	0,0655	10,053 \pm 0,279	10,223 \pm 0,281	8,042 \pm 0,279	8,393 \pm 0,255	3,855 \pm 0,174
Change in level of β %	-	-	+1,7	-20	-20	-62

to the irradiation unit. After all the solution passed through, the generator switched off, and the shutoff valve closed. Solution treatment rate was 2,1 L/min.

The solution's activity was measured by the standard procedure. A sample of 1,5 mL in volume was taken from the analyzed solution utilizing a measuring graduated cylinder. The sample was placed on a substrate and dried. The substrate was then placed in the measuring chamber of the instrument UMF-2000. A quadruple measurement of the count rate of β -activity was carried out within 1000 seconds. At the same time, the background activity was measured. The initial sample was taken from the prepared solution. The irradiated sample was taken from the receptacle.

The repeated measurements of the activity of irradiated solutions after 1 and 6 days were performed by taking 50 mL samples and placing them in receptacles with a strained plug. Samples of 1,5 mL were taken from these vessels to measure the activity. The results of measuring the activity of the samples of the initial and irradiated solutions are shown in Table 2. It follows from this table that the solution's activity was weakened by 20 % after one run within 6 days after exposure.

Table 2: Specific activity of strontium, $N-N_{\text{bckgrn}}$, cps

Sample, time	Source	1 run	Change, %
0	7,45 \pm 0,12	8,70 \pm 0,13	+16
1 days		6,702 \pm 0,113	-10
6 days		5,98 \pm 0,11	-20

The comparison of all of the experiments with the radionuclide ^{90}Sr allows making the following conclusions: the shape of the vessels and the type of irradiator affected the weakening in the activity of the solution; in all of the experiments, the sorption component was significantly less than the

change in the activity of the solution; the influence of pulse amplitude was insignificant, whereas the influence of the pH value was significant. After treating the solutions with the average pH value by NSEMPs, a reduction in the activity of the solution by 15-20 % can be obtained without taking into account the average sorption component \approx 5 %. The conclusions on the sorption component do not apply to flow-through units since no such studies were carried out there. At high pH, the activity of ^{90}Sr can be reduced by 60-80 %.

3.2. Irradiation of solutions with ^{137}Cs

The ^{137}Cs element is a radioactive nuclide with the main beta decay (95 %), a low level of gamma radiation, and a half-life of 30 years.

Test 3. One hundred milliliters of the mixed solution of ^{137}Cs were poured into a 64 mm vessel diameter and 94 mm in height. An irradiation source shaped like a copper rod of 6 mm in diameter and 100 mm in height was inserted into the vessel. The end of the rod was 14 mm away from the vessel bottom. The generator unit used had a pulse amplitude of 15 kV and a repetition frequency of 1 kHz. The initial pH value was 5,72.

The activity of the solutions was measured on a UMF-2000 low-background unit: two parallel samples of 0,1 mL each were taken, dried under the bulb, and the count pulses were measured thrice for 100 s. The activity was measured right after and within one day of the exposure. The solution activity measurement results are presented below:

- The initial solution activity was 61 kBq/L;
- After two minutes of exposure, the activity went down by 36 % to reach 39,0 \pm 1,8 kBq/L;
- After five minutes of exposure, the activity went down by 31 % to reach 42,0 \pm 0,5 kBq/L;

- After ten minutes of exposure, the activity went down by 31 % to reach $42,0 \pm 1,1$ kBq/L;
- After one day, the solution irradiated for ten minutes and poured into a plastic vessel showed an activity of $39,0 \pm 2,6$ kBq/L, a decrease of 36 %.

Test 4. The solutions were irradiated in a flow-through unit with coaxial cylinders. Three and a half liters of the cesium chloride solution with $\text{pH} = 6$ were made. The solution was poured into a consumption tank, and then the generator was switched on. The generator unit used a pulse amplitude of 15 kV and a repetition frequency of 1 kHz. The shutoff valve was opened on the consumption tank, and the solution was supplied to the irradiation unit. After all the solution passed through, the generator switched off, and the shutoff valve closed.

The solution activity was measured by the standard technique on a GAMMA-1P gamma spectrometer. The vessel with 50 mL of the source solution of ^{137}Cs was placed right to the spectrometer's measuring chamber. Then the solution was irradiated, and the 50 mL of it were poured from the receiving tank into a vessel of the same kind that was placed in the measuring chamber. The source and the irradiated solution samples were left in those same vessels tightly closed with a cap after being taken. The activity measurements were repeated within 1, 3, and 6 days. The activity levels of the source and the irradiated solution are shown in Table 3. The repeated runs were performed similarly to the maiden run after taking a sample for activity measurements.

It follows from Table 3 that, after the maiden run, the solution became less active by 40 %. There was no weakening in the solution's activity after the second and third run. The activity level decreased after the maiden run remains unchanged after 1, 3, and 6 days.

In sum, the parameters tracked during the irradiation of the solutions with radioactive nuclides were the influences of the vessel shape, the type of irradiation units, the pulse amplitude and repetition frequency, the solution's pH, and the

effect of various added salts. All in all, more than 50 tests were conducted. The sorption component was not examined because it was measured most frequently in the same vessel before and after irradiating the solution.

The most interesting results were obtained on the flow-through electropulse machining unit. The passage of the solution through the unit was attended by a decrease in the solution's activity by 40 %. A decrease in activity by 50 % (half-life period) occurs for thirty years in natural conditions. It is seen that this period shortened to minutes. The essentially new result obtained on that unit was that the change in the solution's activity had its limit, irrespectively of the number of runs through the unit. The solution's activity remains decreased for a long time.

3.3. Irradiating actual radioactive waste

This subsection describes the tests aimed at exposing actual liquid radioactive solutions to NSEMPs. The research subject was the water solution of various radioactive nuclides of complex chemical composition.

Test 5. The pulses were emitted using the rod-in-the-jar unit (description was given in section 2).

This test was conducted using the jar made as a glass bottle of 75 mm in diameter, 140 mm in height, and 600 mL in volume. The bottle was wrapped on the bottom and sides in the foil of 100 μm in thickness. The rod was an M0 copper bar of 6 mm in diameter, 195 mm in length, and the rod's tip was 40 mm away from the bottle bottom.

The generator produced pulses of 0,6 ns in length and 15 kV in amplitude with a repetition frequency of 1 kHz. The pulse power was 4,5 MW. The energy of one pulse was $\approx 10^{-2}$ J. The generator's dimensions were $320 \times 210 \times 80$ mm, its main power was below 50 W, and it was lighter than 5 kg. The processing time was 15 minutes.

At the time of the test, the solution of complex chemical composition had $T = 18^\circ\text{C}$ and $\text{pH} = 5$. The bombardment of the solution with NSEMPs left large flakes on its surface, the test results are shown in Table 4.

The specific activity reduction coefficient was determined as the ratio of the specific radioactive

Table 3: Activity of caesium solutions, kBq/L

Sample, time	Source	1 run	2 runs	Decrease, %
0	37,0 ± 3,7	22,3 ± 3,1	23,0 ± 2,3	-
1 days	37,0 ± 3,3	23,8 ± 2,4	25,1 ± 2,7	35
3 days	38,0 ± 3,7	23,3 ± 2,6	25,3 ± 2,6	35
6 days	39,0 ± 3,2	22,8 ± 2,3	24,8 ± 2,7	33

Table 4: Specific activity reduction coefficient after test 5

Radioactive nuclide	Reduction coefficient
⁵⁴ Mn	8,75
⁶⁰ Co	10
⁹⁴ Nb	9,5
¹²⁵ Sb	3,38
¹³³ Ba	4,02
¹³⁴ Cs	4,6
¹³⁷ Cs	4,46

Table 5: Specific activity reduction coefficient in the solution after test 6

Radioactive nuclide	Reduction coefficient
⁵⁴ Mn	58
⁶⁰ Co	20
⁹⁴ Nb	46
¹²⁵ Sb	3,2
¹³³ Ba	-
¹³⁴ Cs	7,7
¹³⁷ Cs	8,6

activities of a radioactive nuclide before and after processing the solution with NSEMPs.

It follows from Table 4 that the specific activity of all the radioactive nuclides became much weaker. The higher was the radioactive nuclide's atomic number, and the lower was the reduction coefficient.

Test 6. This test differed from test 5 by adding NaOH alkali to the solution with a pH of 10. A glass bottle was replaced with a steel vessel of 100 mm in diameter, 120 in height, and 800 mL in volume. The rod was an M0 copper bar of 6 mm in diameter, 195 mm in length, and the rod's tip was 40 mm away from the bottle bottom.

The 15-minute bombardment with NSEMPs left dark flakes in the solution and on the rod that became black. The processed solution also had many deposits.

The deposit activity was analyzed separately. The sorption effect was measured by soaking a gauze piece in water and using it to pick up the flakes and particles from the vessel's surface. Table 5 provides the values of the specific activity reduction coefficient for the water solution after test 6 compared to the initial solution.

It follows from Table 5 that the specific activity of some of the radioactive nuclides became much weaker. The reduction coefficient of the specific activity decreased with an increase in the sequence

number of the element.

Table 6 provides the values of the specific radioactive nuclide activity rise coefficient for the deposit after test 6 compared to the initial solution. It follows from Table 6 that the specific activity of all the radioactive nuclides in the deposit became much higher. Table 7 provides the values of

Table 6: Specific activity rise coefficient in the deposit after test 6

Radioactive nuclide	Reduction coefficient
⁵⁴ Mn	6,7
⁶⁰ Co	5,3
⁹⁴ Nb	5
¹²⁵ Sb	6,8
¹³³ Ba	7,9
¹³⁴ Cs	5,6
¹³⁷ Cs	7,2

the coefficient of changes in the activity of the radioactive nuclides on the vessel walls after test 6 in comparison with the initial solution. It follows from Table 7 that most of the radioactive nuclides were by two to three orders of magnitude as active on the vessel walls as in the initial solution. As shown by the tests, the bombardment of the water solutions of ⁹⁰Sr, ¹³⁷Cs, and of the actual waste with powerful nanosecond electromagnetic pulses allowed making the solutions much less active

Table 7: Coefficient of changes in the specific activity on the walls after test 6

Radioactive nuclide	Reduction coefficient
^{54}Mn	$0,5 \times 10^{-3}$
^{60}Co	$0,9 \times 10^{-3}$
^{94}Nb	$7,1 \times 10^{-2}$
^{125}Sb	$0,8 \times 10^{-3}$
^{133}Ba	$1,5 \times 10^{-2}$
^{134}Cs	$2,5 \times 10^{-3}$
^{137}Cs	2×10^{-2}

for a short time. The proposed unit designs make possible a simple increase in the amount of processed solution. For the technical reasons for waiting in the queue, the activity measurements were carried out 20 days after processing. This period can be considered as the time of minor changes in the measured activity level.

3.4. Processing irradiated graphite

The test object was a $120 \times 195 \times 325$ mm block of graphite taken from an IVV- 2M reactor. The block's 120×195 mm bottom butt surface had an indentation of 65 mm in diameter and 7 mm in depth. The upper butt surface had a projection of 65 mm in diameter and 7 mm in depth. The block's butt surfaces were connected to two copper foil electrodes. The electrodes were connected to the generator of pulses of 1 ns in length, 15 kV in amplitude, and a repetition frequency of 1 kHz. The generator's pulse power was 4,5 MW; and the consumed power was 35 W. The periods it took to process the block after three consecutive tests were 5, 15 (5+10), and 25 (5+10+10) minutes, respectively.

The graphite samples were taken from the block's 195×325 mm side surface before the test session and after each test. The sampling from the block for analysis was made in three points by drilling the bottom part of the block (30 mm from the bottom edge).

The drilling was carried out using bore bits of two diameters. First of all, the graphite was drilled with a bore bit of 12 mm in diameter to a depth of 10 mm. After the bored graphite was removed, an aperture of 8 mm in diameter was drilled to a depth of 60 mm. The extracted

graphite powder was collected after each drilling into numbered canisters. The powder was then stirred, and samples of 0,1 g were collected from the canisters for activity measurements.

The measurements were made on a UMF-2000 alpha-beta radiometer. Each counted sample was exposed to five measurements of 100 s each. The measurement results are shown in Table 8. The maximal power of a gamma radiation dose from the block's surface before the processing was 2 mSvph. Its small changes were revealed by the gamma spectrometric analysis conducted before and after the processing. Patent [28] has been obtained for a processing method of irradiated graphite blocks and a unit for applying it.

The processing of a graphite block of such dimensions without its destruction was carried out for the first time in the world. There is no information in the literature about similar experiments.

3.5. Comparison of results

The effect on ^{137}Cs can make the best comparison of results. In Kladov [4](2019), a hydrodynamic generator of mechanical vibrations of the rotary type was powered by a 15 kW engine. The volume of the processing zone was 250 cm^3 . The decrease in the activity of the solution was more than 60 % during the 20 hours of unit operation.

In [6], a cuvette of 5 mL in volume with the solution of ^{137}Cs was scanned for 10 hours with a laser with an energy of 2 mJ per pulse. During the irradiation, the concentration of ^{137}Cs went down by 5 %.

In Filippov [29], the theoretical justification is provided for an increase in the probability of the $^{137}\text{Cs} \rightarrow ^{137}\text{Ba}$ disintegration by 2,9 times under the influence of the high-power field of a sinusoidal electromagnetic wave.

In the flow-through unit (test 4), the solution of ^{137}Cs became less active by 40 % during the passage of the solution through the unit for about 2 min. The consumed power of the pulse generator was 35 W. Unfortunately, and there is no such data for other radionuclides in the available literature.

Table 8: Activity of aggregate beta-emitting radioactive nuclides in graphite

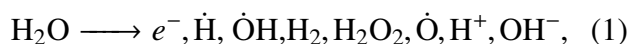
Sampling area	Sampling depth (mm)	Processing time (min)	Specific activity Bq/kg × 10 ⁷	Fold change	Specific activity after 6 days Bq/kg × 10 ⁷	Fold change
Block bottom	10	initial	3,5	-	3,8	-
		5	3,1	-1,1	2,8	-1,4
		15	4,3	+1,2	4,6	+1,2
		25	3,4	-1,02	2,5	-1,52
	60	initial	2	-	1,8	-
		5	9,3	+4,15	10	+5,5
		15	2	1	1,9	+1,05
		25	0,74	-2,85	0,82	-2,2

Thus, the suggested flow-through technology has advantages in operation, processing speed, and energy consumption.

According to existing terminology, it can be attributed to the acceleration of the decay of radionuclides under the influence of high-power energy fields.

3.6. Radiolysis of water

Deposit in test 6 was caused by a change in the chemical composition of the solution. Comparing the liquid irradiation tests with the literary data from [30, 31] allows assuming that the observed effects were similar to the action of radioactive emissions. In this case, the close analogs are X- and γ - radiation. These are defined as sine wave oscillations with a frequency from 3×10^{16} to 3×10^{22} Hz. The frequency spectrum of these pulses is 10^9 Hz. The radiolysis of water followed the bombardment with NSEMPs. The radiolysis of water can be recorded as equation (1)



where e^- is the hydrated electron, $\dot{O}H$ is the radical, \dot{H} is the atom of hydrogen, H_2 is the molecular product, H_2O_2 is hydrogen peroxide H^+ , OH^- are ions. The radiation chemical output of \dot{O} is very low and can be neglected.

The average radiation chemical output values of the components of radiolysis found by experimentation at various pH are enumerated below: $g(e^-)$ was 2,6 to 3, $g(\dot{H})$ was 0,55 to 0,6; $g(\dot{O}H)$ was 2,7 to 2,95; $g(H_2O_2)$ was 0,68 to 0,8; $g(H_2)$ was 0,4 to 0,45.

A hydrated electron is one of the main products of radiolysis of water and water solutions. The

physical properties of this electron are taken from [30]:

- the oxidation and reduction potential is 2,7 V;
- the average charge distribution radius is 0,28 nm;
- the theoretical origination time is 10^{-11} s; as found by test, the hydrated electron is formed in less than 0,02 ns; the equivalent conductivity is $185 \text{ Cm} \cdot \text{cm}^2$.

The hydrated electron is the simplest and a very powerful reducer rapidly reacting with many substances, including most water radiolysis products. The groups in which the reactions with e^- are divided are

1. The addition of e^- to ions.
2. Joining neutral molecules.
3. Dissociative addition.

The H atom is important to the radiolysis of acid water solutions. The H atom is a weaker reducer than e^- and reduces the ions of metals with oxidation-reduction potential that are less negative than -2,3V (Fe_3^+ , Cu_2^+ , Ce_4^+ , and others).

The molecular products of the radiolysis of water are H_2 and H_2O_2 . As a rule, molecular hydrogen does not interact with dissolved particles.

Peroxide can be either a reducer or an oxidizer, which depends on the dissolved substance's essence and conditions; the oxidizing and reducing properties of hydrogen peroxide are more pronounced in acid and alkaline media, respectively.

The bombardment of double-distilled water with NSEMPs made its electrical resistance much weaker, which meant that new particles had

been generated called hydrated electrons. i.e., the bombardment with NSEMPs produced the radiolysis of water. Thus the radiolysis of water significantly modified its composition and physical properties and resulted in precipitation of dissolved substances.

4. Conclusion

The tests carried out have shown that the exposure of powerful NSEMPs on the water solutions of radionuclides leads to a decrease in their activity. In a flow-through unit, the solution of ^{90}Sr is reduced by 20 %, the activity of the solution of ^{137}Cs is reduced by 35 % for 2 minutes of processing time. The measured activity value was maintained for 6 days.

The processing of real radioactive waste showed a decrease in the activity of various radionuclides by 5-50 times. The measured decrease in activity was maintained for 20 days.

The results of the pulsed processing of the irradiated graphite block indicate that the proposed processing method is promising.

The decrease in the solution activity is due to two processes: acceleration of the disintegration during pulse action and radionuclides precipitation due to reactions with water radiolysis products.

There is the repeatability of the results, which is determined and due to the characteristics of the generators used.

Acknowledgments




The research was funded by RFBR and Chelyabinsk Region, project number 20-43-740025.

5. References

- [1] L. I. Urutskoev, V. I. Liksonov, and V. G. Tsinoev, "Experimental Detection of Strange Radiation and Transformation of Chemical Elements [Eksperimental'noye obnaruzheniye "strannogo" izlucheniya i transformatsii khimicheskikh elementov]," *Appl Phys*, no. 4, pp. 83–100, 2000.
- [2] V. D. Kuznetsov, G. V. Myshinskiy, V. I. Zemennik, and V. O. Arbuzov, "Proverochnyye eksperimenty ponablyudeniyu effekta kholodnoy transmutatsii elementov. [Check Tests for Observing Cold Transmutation of Elements]," in *Proceedings of the 8th Russian conference on cold transmutation of chemical element cores and Moscow, Russia*, Moscow, Ed. FIAN, 2001, pp. 308–332.
- [3] R. I. Nigmatulin, J. R. T. Lahi, R. P. Taleyarhan, K. D. West, and R. S. Block, "O termoyadernykh pritsessakh v kaviruyuschikh. [On thermonuclear processes in cavitating bubbles]," *Physics Uspekhi*, vol. 57, no. 9, p. 877–890, 2014.
- [4] A. F. Kladov, "Inventor, Ultrasonic Activator," Russia Russian Federation Patent 2 085 273, N° 95 109 892/25, Application of 1995 Apr 20, 1995 and published on 1997 Jul 27.
- [5] M. Jung, F. Bosch, K. Beckert, H. Eickhoff, H. Folger, B. Franzke, A. Gruber, P. Kienle, O. Klepper, W. Koenig, C. Kozuharov, R. Mann, R. Moshhammer, F. Nolden, U. Schaaf, G. Soff, P. Spädtke, M. Steck, T. Stöhlker, and K. Sümmerer, "First observation of bound-state β -decay," *Physical Review Letters*, vol. 69, no. 15, pp. 2164–2167, 1992.
- [6] S. N. Andreev and G. A. Shafeev, "Nonlinear Quenching of the Radioactivity of Aqueous Solutions of Salts of Nuclides by Laser Ablation of Nanoparticles of Metals [Nelineynoye tusheniye radioaktivnosti vodnykh rastvorov soley nuklidov pri lazernoy ablyatsii nanochastits metallov]," *Radioelectronics. Nanosystems. Information Technologies, (RENSIT)*, vol. 9, no. 1, pp. 65–73, 2017.
- [7] M. A. Tukhtarov, L. A. Andreeva, and A. A. Romanenkov, "Konditsionirovaniye reaktornogo grafita vyvodimykh iz ekspluatatsii uran-grafitovykh reaktorov dlya tseley zakhroneniya. [Conditioning Reactor Graphite of Uranium-Graphite Channel-Type Reactors Put Out of Action for Burial]," *Atomic Energy*, vol. 8, no. 6, 2016.
- [8] V. P. Alexandrov, "Method of radioactive graphite waste treatment," Russia Patent RU 2 273 068, MPK G21F 9/28, published March 23, 2006.
- [9] Y. A. Pohitonov and M. U. Kirshin, "Method of processing reactor graphite wastes," Russia RF Patent RU 2 624 270. MPK G21F 9/28, published July 7, 2017.
- [10] E. V. Bsepela, A. O. Pavluk, A. M. Izmestiev, S. G. Kotlyarevskiy, and A. M. Mikhailets, "Method of processing irradiated reactor graphite," Russia RF Patent RU 2 580 818. MPK G21F 9/28, published on April 10, 2016.
- [11] V. F. Myshkin, A. O. Pavluk, and S. Kotlyarevskiy, "Method of cleaning irradiated graphite bushings of uranium-graphite reactor and device for its implementation," Russia RF Patent RU 2 603 015. MPK G21F 9/28, published on November 20, 2016.
- [12] A. Baeza, M. Fernandez, M. Herranz, L. F., M. C., and S. A., "Elimination of man-made radionuclides from natural waters by applying a

- standard coagulation–flocculation process,” *Journal of Radioanalytical and Nuclear Chemistry*, vol. 260, no. 2, pp. 321–326, 2004.
- [13] A. Bo, S. Sarina, H. Liu, Z. Zheng, Q. Xiao, Y. Gu, G. Ayoko, and H. Zhu, “Efficient removal of cationic and anionic radioactive pollutants from water using hydrotalcite-based getters,” *ACS Applied Materials & Interfaces*, vol. 8, no. 25, pp. 16 503–16 510, 2016.
- [14] A. G. Chmielewski, M. Harasimowicz, B. Tyminski, and G. Zakrzewska-Trznadel, “Concentration of low- and medium-level radioactive wastes with three-stage reverse osmosis pilot plant,” *Separation Science and Technology*, vol. 36, no. 5–6, pp. 1117–1127, 2001.
- [15] S. Y. Ding, Y. Yang, H. O. Huang, H. C. Liu, and L. A. Hou, “Effects of feed solution chemistry on low-pressure reverse osmosis filtration of cesium and strontium,” *Journal of Hazardous Materials*, vol. 294, pp. 27–34, 2015.
- [16] P. A. Haas, “A review of information on ferrocyanide solids for removal of cesium from solutions,” *Separation Science and Technology*, vol. 28, no. 17–18, pp. 2479–2506, 1993.
- [17] K. S. Hwang, C. W. Park, K.-W. Lee, S.-J. Park, and H.-M. Yang, “Highly efficient removal of radioactive cesium by sodium-copper hexacyanoferrate-modified magnetic nanoparticles,” *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 516, pp. 375–382, 2017.
- [18] F. Jia and J. L. Wang, “Separation of cesium ions from aqueous solution by vacuum membrane distillation process,” *Progress in Nuclear Energy*, vol. 98, pp. 293–300, 2017.
- [19] F. Jia, J. F. Li, J. L. Wang, and Y. L. Sun, “Removal of strontium ions from simulated radioactive wastewater by vacuum membrane distillation,” *Annals of Nuclear Energy*, 2017.
- [20] P. Kaewmee, J. Manyam, P. Opaprakasit, G. Le, N. Chanlek, and P. Sreearunothai, “Effective removal of cesium by pristine graphene oxide: performance, characterizations, and mechanisms,” *RSC Advances*, vol. 7, no. 61, pp. 38 747–38 756, 2017.
- [21] Y. Kim, T. Kim, Y. Kim, D. Harbottle, and J. Lee, “Highly effective Cs⁺ removal by turbidity-free potassium copper hexacyanoferrate-immobilized magnetic hydrogels,” *Journal of Hazardous Materials*, vol. 340, pp. 130–139, 2017.
- [22] K. Kosaka, M. Asami, N. Kobashigawa, K. Ohkubo, H. Terada, N. Kishida, and M. Akiba, “Removal of radioactive iodine and cesium in water purification processes after an explosion at a nuclear power plant due to the Great East Japan Earthquake,” *Water Research*, vol. 46, no. 14, pp. 4397–4404, 2012.
- [23] S. Kwon, J. Choi, S. Cho, H. Lee, W. Oh, and S.-J. Choi, “A novel method for separating Cs⁺ from liquid radioactive waste using ionic liquids and a selective extractant,” *Journal of Radioanalytical and Nuclear Chemistry*, vol. 311, no. 3, pp. 1605–1611, 2016.
- [24] J. M. Arnal, M. Sancho, B. García-Fayos, G. Verdú, C. Serrano, and J. T. Ruiz-Martínez, “Declassification of radioactive water from a pool type reactor after nuclear facility dismantling,” *Radiation Physics and Chemistry*, vol. 138, pp. 72–74, Sep. 2017.
- [25] N. J. Hartley, J. Grenzer, W. Lu, L. Huang, Y. Inubushi, N. Kamimura, K. Katagir, R. Kodama, A. Kon, V. Lipp, M. Makita, T. Matsuoka, N. Medvedev, S. Nakajima, N. Ozaki, T. Pikuz, A. V. Rode, K. Rohatsch, D. Sagae, A. K. Schuster, K. Tono, J. Vorberger, T. Yabuuchi, and D. Kraus, “Ultrafast anisotropic disordering in graphite driven by intense hard X-ray pulses,” *High Energy Density Phys*, vol. 32, pp. 63–69, 2019.
- [26] V. S. Belkin and G. I. Shulzhenko, “Nanosecond pulse generator,” Russia RF Patent RU N2 004 064. MPK H03K3/33, H03K3/45, published on June 05, 1991.
- [27] State All-union Standard Russia 51592-2000, “Water. General sampling requirements,” Technical Committee for Standardization TC 343 “Water Quality”, Russia, Tech. Rep., 2000.
- [28] V. V. Krymsky, L. P. Sinelnikov, I. V. Novoselov, and V. N. Ivanov, “Method of reducing radioactivity of spent graphite blocks and apparatus for implementation thereof,” Russia RF Patent RU 2 726 145. MPK G21F 9/28, published on July 7, 2020.
- [29] D. V. Filippov, *Vliyaniye ionizatsii i vzbuzhdeniya atomov elektromagnitnym polem na usloviya stabil’nosti yader i protsessy radioaktivnogo raspada. [Influence of Ionization and Excitation of Atoms by Electromagnetic Field on Nuclei Stability Conditions and Radioactive Disintegration Processes]*. Moscow: A. M. Prokhorov General Physics Institute of the Russian Academy of Sciences Publ, 2008.
- [30] E. J. Hart and M. Anbar, *The Hydrated Electron*. New York: Wiley Inter Science, 1970.
- [31] A. K. Pikaev, *Modern Radiation Chemistry. Sovremennaya radiatsionnaya khimiya. Radioliz gazov i zhidkostey. [Radiolysis of Gases and Liquids]*. Moscow: Nauka Publ, 1986.

Numerical modeling of a debris flow associated with a dam break in the Quillcay sub basin, Ancash, Peru

Abelardo M. Díaz-Salas ^{*,a} , Edilberto Guevara-Pérez ^b , James D. Vidal-Moreno ^a 

^aUniversidad Nacional Santiago Antúnez de Mayolo, Ancash, Perú

^bUniversidad Nacional Agraria La Molina, Lima, Perú

Abstract.- The objective of the research was to show results of a numerical modeling of debris flow, associated with a possible rupture of the Shallap dam, located at the head of the Quillcay river basin, in the department of Ancash, Peru. The modeling used is based on a sequence of chain events, which allowed coupling different numerical models, which simulated a complex process, and then apply methodologies that allow the elaboration of hazard maps. The results obtained show affected areas that vary from 0,82 km² to 0,87 km² for each level of risk, most of which are in low to intermediate categories, and are located on the left bank of the Quillcay River, representing 55 % on average of the total estimated area of impact. The results could allow authorities to develop risk management tools associated with the implementation of early warning systems to protect the most vulnerable areas.

Keywords: debris flow; flow modeling; RAMMS; IBER; dam breaking.

Modelamiento numérico de un flujo de escombros asociado a una rotura de presa en la subcuenca Quillcay, Áncash, Perú

Resumen.- La investigación tuvo como objetivo, mostrar resultados de un modelado numérico de flujo de escombros, asociado a una posible rotura de la presa Shallap, ubicada en la cabecera de cuenca del río Quillcay, en el departamento de Áncash, Perú. La modelización utilizada se basa en una secuencia de eventos en cadena, que permitió acoplar diferentes modelos numéricos, los cuales simulaban un proceso complejo, para luego aplicar metodologías que permiten elaborar mapas de peligros. Los resultados obtenidos muestran áreas afectadas que varían de 0,82 km² a 0,87 km² para cada nivel de peligro, la mayoría de las cuales están en categorías bajo a intermedio, y se ubican en la margen izquierda del río Quillcay, representando un 55 % en promedio del área total estimada de afectación. Los resultados podrían permitir a las autoridades, desarrollar herramientas de gestión de riesgos asociados con la implementación de sistemas de alerta temprana para proteger a las áreas más vulnerables.

Palabras clave: flujo de escombros; modelamiento de flujos; RAMMS; IBER; rotura de presas.

Recibido: 13 de octubre, 2020.

Aceptado: 06 de marzo, 2021.

1. Introducción

A lo largo de la historia, en diferentes partes del planeta, han ocurrido muchos tipos de desastres naturales. Dadas sus características topográficas y su ubicación geográfica, la región de los andes peruanos no es la excepción y está propensa a sufrir el impacto de terremotos, inundaciones por el fenómeno de El Niño, huaycos, avalanchas de hielo

y aluviones por ruptura de lagos de origen glaciar. En lo que se refiere a inundaciones o aluviones producidos por la ruptura de lagos, el departamento peruano de Áncash ha experimentado una trágica historia que se remonta al año 1941 con la ruptura del dique morrénico de la laguna Palcacocha, evento que produjo el mayor aluvión de origen glaciar que se conozca en el pasado, inundando la zona urbana de la ciudad de Huaraz [1].

Debido al peligro que representan los eventos naturales mencionados, es muy importante estudiar sus efectos potenciales para aplicar estrategias que permitan reducir los riesgos, sobre todo en poblaciones vulnerables como las que se asientan en la región de los andes. Los resultados de las

* Autor para correspondencia:

Correo-e: abelardo_mad@hotmail.com (A. Díaz-Salas)

investigaciones deben coadyuvar a la zonificación de las áreas de riesgo y el ordenamiento territorial para el uso de las tierras sobre una base científica confiable, independientemente del hecho de que esta tarea es intrínsecamente difícil en áreas de montaña alta como la que corresponde al presente estudio [2].

Por lo expuesto, se ha considerado llevar a cabo esta investigación sobre modelación numérica del flujo de escombros asociado a una posible rotura de la presa Shallap en la cabecera de cuenca del río Quillcay en el departamento de Áncash, Perú. Se aplican procedimientos establecidos en países de condiciones similares para generar mapas de amenazas de flujos, tanto en forma retrospectiva para eventos documentados; como prospectiva, para posibles escenarios de amenazas [3], [4].

2. Fundamentación teórica

Los flujos de escombros en las áreas montañosas representan una gran amenaza para las personas y las propiedades y han sido una gran preocupación en todo el mundo; se deben a la ocurrencia de las lluvias extremas y al efecto de la gravedad que generalmente acelera el descenso; siguen las zanjas empinadas de las montañas hasta los abanicos o llanuras aluviales donde normalmente se asientan las poblaciones que a menudo resultan afectadas por estos eventos catastróficos [5].

Un flujo de escombros, es un deslizamiento totalmente saturado de arcilla sensible, grava o arena, desencadenado por un exceso de presión de poros o el efecto de licuefacción del material deslizante [6]. Los factores climatológicos en las zonas montañosas afectan la magnitud y frecuencia de los procesos geomorfológicos, entre los que se incluyen los movimientos en masa o flujos de escombros [7]. La precipitación es uno de los factores principales que desencadena un flujo de escombros; además es el más activo cuando se producen estos tipos de flujos [8], cuya formación y características ha sido motivo de investigaciones utilizando diferentes modelos y aplicaciones en diferentes países. En la ocurrencia de este tipo de flujo, además de las lluvias extremas de periodos de retorno elevados, juega un papel muy importante

el deshielo de nevados y movimientos sísmicos, especialmente en zonas de alta montaña.

La denominación de estos flujos varía con las regiones y los países. En Perú se conocen como *Huaicos* y se asocian con flujos rápidos de aguas turbias y turbulentas de corta duración cargados de rocas y material sólido de diferentes tipos y tamaños. En Indonesia se denominan *Lahares* a los flujos de lodos de origen volcánico; en Bolivia se conocen como *Riadas* a los flujos torrenciales cargados de sedimentos. Hay otras denominaciones, como flujo de tierra, de lodo, de lodo rocoso, de escombros, aluviones, flujo de residuos, flujos hiperconcentrados de sedimentos, flujos torrenciales, entre otros, todos dentro de la denominación anglosajona de *Debris Flows* [9].

Se han desarrollado muchos modelos numéricos para dilucidar el proceso de movimiento del flujo de escombros, predecir las distancias de movimiento y el rango del efecto de estos. El modelo de fricción de *Voellmy* es capaz de simular con precisión el movimiento remoto del flujo de escombros; el de dos parámetros describe el comportamiento de fricción entre los escombros que fluyen [5]; fue propuesto originalmente para avalanchas de nieve, usando la suposición del fluido como un modelo monofásico que puede caracterizarse por dos tipos de parámetros; es decir, altura del fluido, $H(x, y, t)$ (m) y media de la velocidad $U(x, y, t)$ (m/s); la expresión matemática es como sigue en la ecuación (1).

$$U(x, y, t) = [U_x(x, y, t)U_y(x, y, t)]^T, \quad (1)$$

donde U_x , denota la velocidad en el eje X , U_y , es la velocidad en la dirección Y , y T es el símbolo de la transposición de la velocidad media. La magnitud de la velocidad también se puede definir como en la ecuación (2):

$$\|U\| = \sqrt{U_x^2 + U_y^2}. \quad (2)$$

$\|U\|$, significa que se toma el valor promedio absoluto de U , lo que asegura que U alcance una velocidad estrictamente positiva en el espacio vectorial; la dirección de la velocidad del fluido está definida por el vector unitario (n_u), cuya denotación es: $n_u = \frac{1}{\|U\|} (U_x U_y)^T$

El modelo reológico de *Voellmy* se expresa mediante la ecuación (3) y constituye el balance de masa:

$$\partial_t H + \partial_x(HU_x) + \partial_y(HU_y) = Q(x, y, t), \quad (3)$$

donde $H(x, y, t)$, denota la altura del flujo (m) y $Q(x, y, t)$, es la fuente de producción en masa en término (m/s), cuando $Q > 0$, se denomina tasa de arrastre; cuando $Q = 0$, no hay erosión o deposición del material; cuando $Q < 0$, denota la tasa de deposición [3].

Existen otros factores geomorfológicos, los derrumbes o deslizamientos masivos, especialmente en áreas deleznable muy accidentadas que producen las inundaciones siendo los efectos directos la erosión y la generación de los flujos de escombros [10]. Para comprender y predecir el comportamiento de los flujos de escombros, se han utilizado modelos numéricos y herramientas analíticas; además existe una gran variedad de modelos, tanto en el enfoque de modelado, como en el tipo de las ecuaciones que se usan [11]. En este contexto, es importante diferenciar los flujos de escombros de las inundaciones; los primeros pueden contener concentraciones en peso de sedimentos entre el 70 y 90 %; mientras que las concentraciones de sólidos en el agua de inundación no pasan del 5 %. Otra diferencia se relaciona con el área de deposición, la cual ocurre para el caso del flujo de escombros en abanicos y diques; mientras que en las inundaciones el material se deposita en capas horizontales. Los flujos de escombros se pueden mover mucho más rápido en canales o cauces de alta pendiente y mucho más lento, cuando las pendientes son reducidas [9]; probablemente debido al incremento del componente de fuerza de la masa del material en el sentido de la pendiente. Esta diferencia entre ambos tipos de flujos hace que la modelación para su tratamiento sea también un tanto diferente; aunque en ambos casos se parte de una fundamentación teórica común.

En el presente estudio se usa el módulo hidrodinámico IBER para resolver las ecuaciones de *Saint Venant*, bidimensionales, (4), (5) y (6), incorporando los efectos de la turbulencia y

rozamiento superficial por viento [12].

$$\frac{\partial h}{\partial t} + \frac{\partial h U_x}{\partial x} + \frac{\partial h U_y}{\partial y} = 0, \quad (4)$$

$$\begin{aligned} \frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x} \left(hu^2 + g \frac{h^2}{2} \right) \\ + \frac{\partial}{\partial y} (huhv) = gh(S_{0x} - S_{fx}), \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{\partial}{\partial t}(hv) + \frac{\partial}{\partial x} \left(hv^2 + g \frac{h^2}{2} \right) \\ + \frac{\partial}{\partial y} (huhv) = gh(S_{0y} - S_{fy}). \end{aligned} \quad (6)$$

Donde h , es el calado, U_x , U_y , son las velocidades horizontales promediadas en profundidad, g , es la aceleración de la gravedad, ρ , es la densidad del agua. En base a la hipótesis que se está siguiendo en el presente trabajo de investigación sobre la generación de flujo de escombros por efecto de la rotura de la presa, se debe precisar que la erosión de presas naturales y artificiales ha ido evolucionando desde principios de la década de 1980, cuando se desarrollaron modelos simples unidimensionales basados en análisis empíricos y paramétricos para procesos de ruptura de presas. Estos modelos son computacionalmente eficientes, pero dependen en gran medida del juicio de ingeniería y el análisis de casos históricos de fallas [13].

La propagación de las ondas del flujo hacia aguas abajo se puede simular mediante modelos que van desde simples modelos empíricos hasta los físicos, como es el caso de los modelos hidráulicos en los cuales se resuelven las ecuaciones de continuidad (conservación de masa o volumen) y de cantidad de movimiento. En el presente estudio se utilizó el modelamiento bidimensional IBER, el cual resuelve volúmenes finitos sobre un sistema de mallas estructuradas y no estructuradas [14]. Una posible aplicación de los denominados modelos bidimensionales de aguas someras vendría a ser la determinación de las zonas afectadas por la rotura total o parcial de una presa. Los esquemas numéricos implementados en el modelo IBER (del tipo Goduv), son especialmente adecuados

para el cálculo de este tipo de flujos, ya que se ha establecido que dichos modelos proporcionan, en los casos indicados, una buena relación entre precisión en los resultados y la complejidad del modelo de cálculo [12].

Las presas desempeñan un papel esencial para satisfacer las demandas de suministro de agua de las ciudades, la agricultura y la industria, o la generación de energía, así como la gestión de inundaciones; sin embargo, también representan un riesgo potencial de avería, que puede provocar graves daños materiales y la muerte [15]. Varias fallas de presas documentadas han sido asociadas con flujos de escombros, que transportan volúmenes significativos de sedimentos y causan cambios geomorfológicos severos. Según datos volumétricos, en muchos casos, el volumen de sedimentos transportado puede ser del mismo orden de magnitud que el volumen del agua drenado del depósito [16]; es por ello que los flujos de escombros se consideran como uno de los fenómenos más peligrosos en las áreas montañosas del mundo [17]. Los estudios de fallas de presas y el mapeo de inundaciones son vitales para establecer planes de emergencia que permiten a las autoridades competentes dar una respuesta rápida y eficaz [15]. En la Cordillera Blanca, la población local es altamente vulnerable a las cadenas de procesos de alta montaña [18]; la ciudad de Huaraz, alrededores y zonas aledañas se encuentran en una expansión urbana de alto índice de ocupación de lugares tradicionalmente agrícolas, ganaderos; así mismo, estos lugares son espacios naturales de sedimentación fluvio-glaciar y de movimiento de masas, frente a eventos extremos de precipitación, aluviones, terremotos [19].

3. Metodología

3.1. Consideraciones generales

Existen diversos métodos para el estudio de la formación y propagación de las ondas de rotura de presas, pero debido a las condiciones orográficas y fisiográficas (altas pendientes, rugosidades y velocidades de propagación) se ha optado por utilizar el método mixto hidrológico-hidráulico [20] y con ello la utilización de modelos

numéricos, que comprenden modelos hidrológicos e hidráulicos.

Para el presente trabajo de investigación se tomó en cuenta una cadena de eventos o procesos, las cuales podrían dar lugar a un aluvión y su posterior determinación de zonas de peligro potenciales en la ciudad de Huaraz [20]. El presunto factor desencadenante es la rotura de la presa Shallap por factores extremos de precipitaciones máximas; la alta dinámica del relieve en relación con los temblores sísmicos y las fuertes lluvias locales son los principales factores desencadenantes [21]; la precipitación es también un parámetro importante para estudios climáticos, cuya variabilidad espacial y temporal pueden impactar en las actividades humanas durante eventos hidroclimáticos extremos como sequías e inundaciones [22].

Los trabajos de campo que se realizaron en la obra de seguridad de la laguna Shallap, estarían presentando deficiencias en temas de mantenimiento y limpieza, como son el talud; esto incluye aguas abajo del dique, así como los canales de entrada, canales de salida, y finalmente el aliviadero, pues, esta construcción se remonta a los años de 1970: además, otro punto desfavorable, es que la presa Shallap carece de un plan de emergencia que evalúe las consecuencias de una hipotética falla, y delimite las posibles inundaciones aguas abajo. La cadena de procesos supuestos desde la rotura de la presa hasta la determinación de zonas de peligro, fueron simuladas utilizando diferentes modelos numéricos. Para la determinación de las tormentas máximas se utilizó el modelo hidrológico HEC-HMS; se simula el hidrograma de una cuenca asociado con datos físicos, y con ello, se estiman los hidrogramas de salida; es decir, caudales máximos y tiempos picos; todo ello partiendo de condiciones extremas de tormentas [23]. Se asume la rotura de la presa usando el modelo numérico IBER, cuyo enfoque de modelado, se basa en un modelo de agua poco profundo de volúmenes finitos [15]. Para modelar la dinámica de la avalancha o flujo de escombros y el posterior depósito de material para las propiedades del canal o cauce existente [2] se utilizó el modelo numérico RAMMS, debido a que el flujo de escombros se moviliza repentinamente;

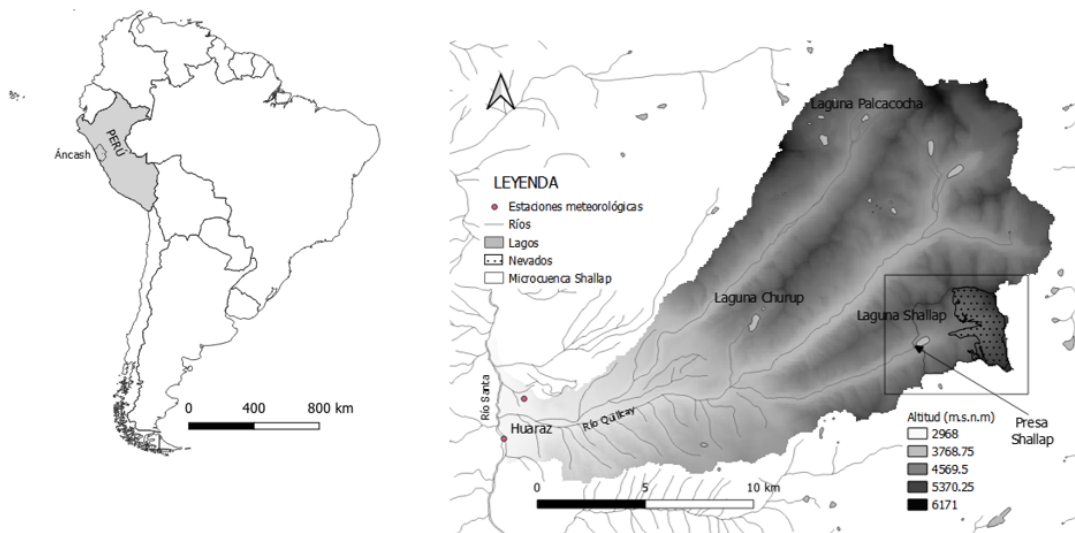


Figura 1: Ubicación de la cadena de eventos para el modelamiento numérico

además el modelo predice trayectorias de flujo, velocidades, flujo de alturas y presiones de impacto [17].

En la Figura 1 se muestra la ubicación de la zona de investigación donde se realizó el modelamiento numérico de la cadena de eventos o procesos.

3.2. Recopilación de información

Información meteorológica

Se recopiló información de cuatro (04) estaciones meteorológicas pertenecientes al Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI, CORPAC y ELECTRO PERÚ). Las estaciones meteorológicas se ubican en las cercanías de la microcuenca Shallow; poseen diferentes períodos de registro, entre 1950 y 2017. Para la investigación se llevó a cabo un análisis regional de las precipitaciones máximas anuales, previa evaluación de control de calidad de los datos mediante métodos y técnicas cualitativas. En la Tabla 1 se presentan las características de las estaciones pluviométricas y de la información utilizada.

Información de campo

A fin de conocer las características técnicas principales y la batimetría, para el uso del modelamiento numérico con IBER, en el área de la presa Shallow se llevó a cabo un levantamiento topográfico del vaso de la laguna, recabando las

Tabla 1: Estaciones meteorológicas cercanas a la zona de estudio

Estación	Latitud S	Longitud W	Altitud (m s.n.m)	Período
Anta	9°21'	77°36'	2748	1977-2001
Huaraz	9°32'	77°32'	3052	1950-2008
Querococha	9°41'27"	77°21'6"	3955	1975-1995
Unasam	9°30'59,5"	77°31'29,5"	3079	1998-2017

características técnicas, como corona, canales de salida, canales de entrada, aliviadero, espejo de agua, borde libre, taludes aguas arriba y taludes aguas abajo.

Trabajo de gabinete

Se recurrió a diversos autores como [13] y [24] de donde se extrajo información de trabajos realizados en escenarios similares de procesos de cadena referente a GLOF. Se recabó información sobre usos y texturas de suelo, parámetros hidráulicos de simulación de flujo de escombros. Todos estos parámetros fueron utilizados para la generación del proceso de cadena bajo un enfoque de modelamiento numérico en la presa Shallow.

3.3. Procesamiento de la información

Regionalización de la precipitación

Para el análisis de las precipitaciones máximas se trabajó con los modelos probabilísticos normal,

log normal 2p, log normal 3p, log pearson tipo III, gumbel y log gumbel. Mediante la prueba de bondad de ajuste se establecieron las precipitaciones máximas de 24 horas en base a información de las cuatro (04) estaciones meteorológicas para los períodos de retorno de 100, 150 y 200 años. La caracterización local de la precipitación se efectuó mediante la regionalización usando la relación dada en la ecuación (7) entre la altitud Z (ms.n.m) y la variable precipitación P (mm) como sigue:

$$P = a + b \cdot Z. \quad (7)$$

Donde, a y b son parámetros de la correlación lineal. Para la estimación de las tormentas de precipitación máximas de 24 horas, se distribuyó la precipitación mediante el método del servicio de conservación de recursos naturales de Estados Unidos (NRCS), con la distribución temporal de lluvia extrema más frecuente Gumbel tipo I, debido a características similares con la zona de estudio.

Parámetros geomorfológicos de la microcuenca Shallap

Se usó una combinación de la plataforma GIS y un modelo de elevación (DEM) digital, para determinar los parámetros geomorfológicos, así como muestreo de suelo, texturas de suelo, grupos hidrológicos, usos y aprovechamientos del suelo, para su posterior determinación del número de curva (CN) en la microcuenca de estudio y el coeficiente de rugosidad de Manning, el cual se utilizó en el modelamiento numérico con IBER [13], [24].

Parámetros hidráulicos para IBER

IBER es un modelo numérico distribuido para la simulación de flujo inestable en superficie libre; se ha utilizado con éxito en una amplia gama de aplicaciones, entre las cuales se incluye la modelización del caudal y la calidad de los ríos, así como los impactos geomorfológicos de la rotura de la presa. El modelo IBER se basa en la Guía de clasificación de presas en función de su riesgo potencial, es por ello que utiliza una fórmula empírica para estimar las características de la rotura [15]. Se supone que el crecimiento de la brecha

es un proceso lineal y la forma de la brecha es trapezoidal con pendientes laterales de 1H:1V, el ancho promedio final de la brecha b (m) y el tiempo de desarrollo de la brecha T (h); estos parámetros se estiman con las ecuaciones (8) y (9).

$$T = \frac{4,8 \cdot V^{0,5}}{h}, \quad (8)$$

$$b = 20(V \cdot h)^{0,25}. \quad (9)$$

Donde V es el volumen almacenado en (hm^3) y h es la altura de la presa (m), además es preciso indicar que este modelo está en sintonía con los modelos de formación de brechas paramétricos ampliamente utilizados, los cuales, como cualquier modelo, tienen incertidumbres relacionadas con mala documentación de eventos históricos [15]. Una discusión más detallada del tema está fuera del alcance de este trabajo por lo que las estimaciones proporcionadas por el modelo se tratan como determinísticas en este trabajo de investigación.

Parámetros hidráulicos en el flujo de escombros

El modelo numérico RAMMS usa ecuaciones bidimensionales de masa y momento que promedia los resultados para el flujo granular en un espacio tridimensional utilizando el método de volumen finito [3]. Las entradas incluyen datos del terreno (DEM), área de liberación del flujo y los parámetros de fricción [13]. Para el dominio de cálculo se usa la información de investigaciones relacionadas a la zona de estudio, tal como la proveniente de la referencia [24], por utilizar condiciones similares a la microcuenca Shallap en lo que se refiere al recorrido del flujo de detritos (objetivo de la investigación) y los parámetros de fricción μ y ξ . El modelo RAMMS usa la relación de fricción de *Voellmy* para describir la fricción de flujo, dada en la ecuación (10).

$$S_f = \mu \rho h g \cos \phi + \frac{\rho g V^2}{\xi}, \quad (10)$$

donde S_f es la pendiente de fricción, ρ es la densidad del fluido (kg/m^3), g es la aceleración gravitacional (m/s^2), ϕ es la pendiente, h es la altura de la avalancha (m) normal a la base y V es

Tabla 2: Niveles de peligro de un flujo de escombros [13]

Nivel de peligro	Descripción	Profundidad del flujo (m)	Descarga por ancho de unidad $v \cdot h(m^2/s)$	
Bajo	Personas en bajo riesgo, daños menores en edificios	$0,2 < h < 1,0$	y	$v \cdot h < 0,2$
Intermedio	Personas en peligro fuera de sus casas. Daño estructural y posible destrucción	$0,2 < h < 1,0$	y	$0,2 < v \cdot h < 1,0$
Alto	Personas en peligro dentro y fuera de sus casas. Estructuras destruidas	$h > 1,0$	o	$v \cdot h > 1,0$

la velocidad del flujo (m/s). En la Ecuación (10) se observa que el parámetro de fricción de Coulomb μ (fricción de superficie seca, en el primer término del lado derecho de la ecuación), domina el valor total de la fricción cuando ésta es relativamente lenta; mientras que ξ (el parámetro de fricción de turbulencia en el segundo término de la ecuación), lo hace cuando el flujo es rápido como el de escombros; aspecto que se está considerando en este trabajo de investigación. En el modelo *Debris flow* de RAMMS, se usaron valores de $\xi = 1000 \text{ m/s}^2$; $\mu = 0,12$; y ; $\rho = 1000 \text{ kg/m}^3$; valores muy similares a los referidos en [24].

En los modelos numéricos se suelen usar herramientas computacionales para la evaluación de los peligros naturales; en este estudio se utilizó el modelo numérico RAMMS, que resuelve ecuaciones bidimensionales de aguas poco profundas para flujos granulares aplicando la relación de fricción de *Voellmy*. Este enfoque es ampliamente utilizado y probado con éxito para avalanchas de nieve; muchos autores enfatizan la similitud entre el comportamiento de flujo de escombros y la aplicación de parámetros de fricción para modelar un flujo de estos [25]. En el estudio se utilizaron parámetros de campo encontrados en referencias de casos similares, como [13] y [24].

3.4. Criterios para la determinación de zonas de peligro

Los mapas de inundaciones se pueden preparar utilizando varios modelos y los datos resultantes se pueden utilizar para diversos propósitos [26]. En el presente caso, para la determinación de zonas de peligro en la ciudad de Huaraz, se utiliza la profundidad h (m) y la velocidad v (m/s), resultados que se obtuvieron con el modelado numérico RAMMS. En un segundo paso las intensidades

de flujos se convierten en niveles de peligro siguiendo los criterios presentados en la Tabla 2 creando un proceso semiautomático implementado en el entorno GIS, el cual convierte la altura y velocidades de flujo generado por RAMMS celda por celda en intensidades y niveles de peligro. Para esta clasificación se utilizó el criterio propuesto en Suiza y Venezuela, donde se asume que la infraestructura en Suiza y Venezuela es menos resistente que la infraestructura en Austria, y esto es similar a la que se tiene en la ciudad de Huaraz [13].

4. Resultados y discusión

Características geomorfológicas de la zona de estudio

En la Tabla 3 se presentan los parámetros geomorfológicos principales calculados para la microcuenca Shallap. Es notoria la altitud media, superando los 5200 ms.n.m. La pendiente es mayor que 30 % y se puede describir como fuertemente accidentada, favoreciendo así las contribuciones al escurrimiento de las precipitaciones que ocurren en temporadas de lluvias en los meses de enero y abril en esta parte de la cuenca.

Tabla 3: Características geomorfológicas de la microcuenca Shallap

Descripción	Microcuenca Shallap
Área (km ²)	13,28
Perímetro (km)	18,70
Altitud media (m s.n.m.)	5213,85
Pendiente (%)	31,70
Número de Curva	90,66
Grupo hidrológico del suelo	Nevado, roca y franco arenoso

4.1. Caracterización de las precipitaciones máximas

En la Tabla 4 se presentan los resultados del análisis de regionalización pluviométrica usando la ecuación (7). Aplicando la prueba de ajuste t-Student se encontró que todos los ajustes son positivos al 95 % de probabilidad. De este modo, los modelos de regionalización muestran un grado de asociación muy bueno, con coeficientes de correlación mayores que 0,84, lo cual permitió usar el modelo de regresión o regionalización para la obtención de las precipitaciones máximas de 24 horas en la zona de estudio (microcuenca Shallap).

Tabla 4: Ecuaciones de regionalización para la microcuenca Shallap

TR	Ecuación (7): $(P = f(Z))$	R^2
100	$P = 0,0227 \cdot Z - 12,543$	0,8499*
150	$P = 0,0243 \cdot Z - 14,826$	0,8451*
200	$P = 0,0254 \cdot Z - 16,451$	0,8417*

*: Correlación muy buena

En la Figura 2 se presentan los hietogramas adimensionales resultantes para la microcuenca Shallap. Para determinar las tormentas de diseño se tuvo en cuenta, el tipo de precipitaciones que se dan en la zona; es decir, precipitaciones máximas para alturas medias de la microcuenca mayores a los 5200 m s.n.m (Tabla 3)

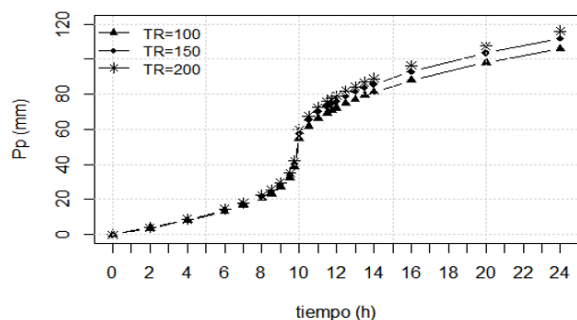


Figura 2: Caracterización de las precipitaciones máximas en la microcuenca Shallap

Caracterización de la rotura de la presa Shallap

En la Tabla 5 se dan las características técnicas de la presa Shallap, las que se usan como datos de entrada del modelo IBER para el modelamiento numérico de la rotura de la presa.

Tabla 5: Parámetros técnicos de la presa Shallap

Descripción	Valor
Cota de corona (m s.n.m)	4273,45
Ancho de corona (m)	9,03
Largo de la corona (m)	24,21
Cota de captación (m s.n.m)	4260
Cota de aliviadero (m s.n.m)	4267,9
Altura de la presa (m)	13,45
Volumen del embalse (m ³)	3467585,29

Los caudales de entrada para el modelo IBER se obtuvieron mediante los modelos hidrológicos HEC-HMS. En la Figura 3 se presentan los hidrogramas de estos caudales de entrada; su comportamiento muestra picos de hasta 60 m³/s a las 15 horas y de hasta 120 m³/s a las 24 horas.

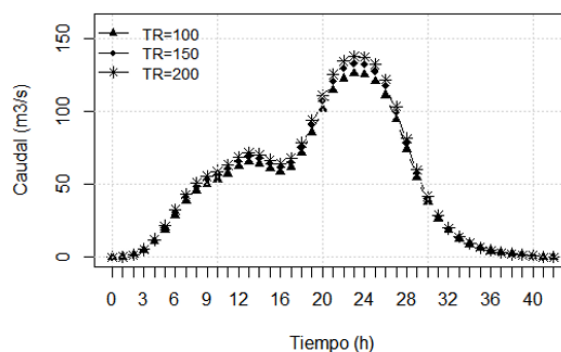


Figura 3: Caracterización de los caudales de ingreso en la microcuenca Shallap

Caracterización del flujo de escombros

En la Figura 4 se presentan los hidrogramas de salida del modelo IBER, así como los hidrogramas de rotura de la presa para períodos de retorno de 100, 150 y 200 años, observándose caudales máximos de 359 m³/s, 424 m³/s y 474 m³/s, respectivamente. Estos valores máximos se utilizaron para la simulación de los niveles de riesgo y peligro con el programa RAMMS, clasificados bajo, intermedio y alto.

Caracterización de las zonas de peligro

Para la caracterización de las zonas de peligro, se utilizó la información dada en la Tabla 2, y la metodología propuesta en la referencia [13] para estimar los diferentes niveles de peligro en la ciudad de Huaraz. Como se ha mencionado,

Tabla 6: Áreas de afectación por cada nivel de peligro en la ciudad de Huaraz

Clasificación del nivel de peligro	Área total de afectación (km ²)	Área de afectación (margen izquierda del río Quillcay) (km ²)	Porcentaje de afectación (%)
Bajo (1)	0,82	0,43	52,43
Intermedio (2)	0,85	0,46	54,11
Alto (3)	0,95	0,56	58,94

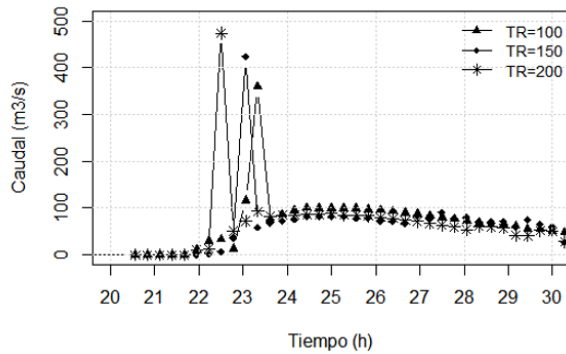


Figura 4: Hidrograma de rotura de la presa Shallap para diferentes períodos de retorno

los resultados provienen de metodologías para estimar flujos de escombros en casos similares de Suiza y Venezuela. Igualmente se consideraron las experiencias de investigaciones realizadas en la ciudad de Huaraz y Carhuaz [24] y [2]. En la Figura 5 se presentan los mapas de peligro obtenidos en la investigación mediante el modelamiento numérico con RAMMS para tres escenarios: bajo (TR = 100), intermedio (TR = 150), alto (TR = 200); provenientes de una cadena de procesos de movimiento masivo, considerando parámetros de fricción [24].

Los criterios de clasificación que se muestran en la Tabla 2 son un caso más probable de ocurrencia cuando se habla de flujo de escombros. Los impactos en volumen de flujo de escombros que ocurren bajo condiciones de los tres escenarios arrojan valores de $4,78 \times 10^5 \text{ m}^3$; $4,81 \times 10^5 \text{ m}^3$ y $4,95 \times 10^5 \text{ m}^3$, para períodos de retorno de 100, 150 y 200 años respectivamente. Las áreas totales de afectación correspondientes son de $0,82 \text{ km}^2$, $0,85 \text{ km}^2$ y $0,95 \text{ km}^2$ respectivamente, para cada nivel de peligro, y están ubicadas mayormente en la margen izquierda del río Quillcay, lo que representa en promedio un 55 % del área total estimada en

la simulación numérica, como se puede observar en la Tabla 6. Los resultados son estimaciones empíricas sujetas a incertidumbres, pero ayudarían en gran medida a conocer el comportamiento de un flujo de escombros bajo los supuestos escenarios modelados para los diferentes periodos de retorno.

El análisis de cada escenario siempre incluye cierto nivel de incertidumbre de estimación, debido a que se basan en relaciones de magnitud-frecuencia, que son difíciles de evaluar y pueden diferir para cualquier región geográfica donde se realice algún trabajo relacionado con modelos numéricos de simulación.

El enfoque que se ha presentado en este trabajo de investigación es de tipo hidrológico e hidráulico, es decir supuestos escenarios de precipitaciones máximas extremas que pueden poner en peligro la seguridad de la presa, y ocasionar la rotura con la subsecuente generación de un flujo de escombros debido a la pendiente que se presenta a lo largo del tramo de río analizado. Como perspectiva para escenarios futuros de modelamiento en quebradas similares se podría incluir deslizamientos simultáneos o caídas de rocas a lagos de origen glaciar, que ocasionan eventos combinados de oleaje que podría representar algún peligro de desbordes para diques morrénicos, tanto artificiales como naturales, eventos desencadenantes éstos que pudieran afectar a poblaciones vulnerables ubicadas aguas abajo.

5. Conclusiones

Se realizó el modelamiento numérico de una cadena de procesos, partiendo de un evento desencadenante como son las precipitaciones máximas en la microcuenca Shallap, para luego realizar la determinación de los niveles de peligro en la ciudad de Huaraz. Para ello se llevaron a cabo

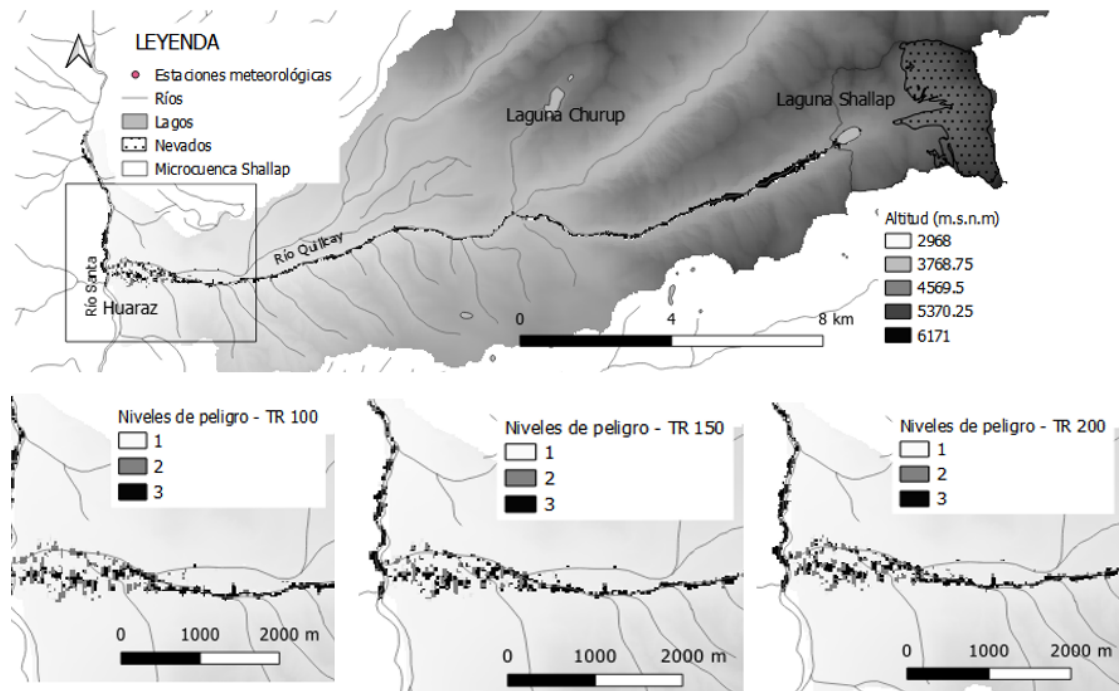


Figura 5: Niveles de peligro determinados en la ciudad de Huaraz, por efecto de la rotura de la presa Shallap

simulaciones con los modelos numéricos HEC-HMS, IBER y RAMMS, así como, experiencias en otros puntos similares a la zona de estudio.

Los caudales que se obtuvieron por efecto de la rotura de la presa Shallap fueron $358,64 \text{ m}^3/\text{s}$, $423,96 \text{ m}^3/\text{s}$ y $473,75 \text{ m}^3/\text{s}$ para periodos de retorno de, 100, 150 y 200 años respectivamente.

El modelamiento que se realizó con el modelo numérico RAMMS, se basó en experiencias trabajadas en escenarios similares en las lagunas 513 y Palcacocha, para su posterior proceso en el entorno GIS de acuerdo con metodologías propuestas en otros estudios similares. Cada uno de los componentes del modelamiento en este proceso de cadenas fue crucial; es decir, entender el comportamiento físico, y, lo más importante, el acoplamiento de estos procesos, donde el resultado de un componente fue necesario para el siguiente.

Se realizaron varias iteraciones para lograr un acoplamiento razonable de cada uno de los modelos. Teniendo ya establecidos los parámetros finales de modelamiento se procedió al modelado, el cual se basó en la determinación y generación de mapas de peligro para la ciudad de Huaraz. De

acuerdo a estos modelamientos se determinaron tres escenarios, bajo, intermedio y alto; se obtuvo en su gran mayoría niveles de peligro bajo a intermedio, con áreas totales afectadas que varían desde $0,82 \text{ km}^2$, $0,85 \text{ km}^2$ y $0,95 \text{ km}^2$ respectivamente para cada nivel de peligro para los diferentes periodos de retorno. Las áreas afectadas se ubican en un 55 % en la margen izquierda del río Quillcay. Las principales calles afectadas son: las avenidas Manco Cápac, Antonio Raymondí, Fitzcarrald, Agustín y las calles aledañas. Hay que precisar que los resultados solo muestran el nivel de peligro físico. En próximas investigaciones se debe incluir el aspecto de la vulnerabilidad de la población e infraestructura.

Otro punto importante que hay que tener en cuenta es la verificación de campo, especialmente en sitios topográficos de difícil acceso y donde el modelo de elevación digital (DEM) puede generar errores de topografía y precisión.

Se necesitan más investigaciones sistemáticas para evaluar la influencia de los escenarios, como son la elección del tipo de modelo, y los parámetros en las alturas finales, así como otros parámetros de



velocidades máximas.

6. Referencias

- [1] S. A. Wegner, *Lo que el agua se llevó, consecuencias y lecciones del aluvión de Huaraz de 1941*. Perú: Corporación Globalmark, Ed. Huaraz, 2014.
- [2] J. Klimeš, M. Benešová, V. Vilímek, P. Bouška, and A. Cochachin-Rapre, “The reconstruction of a glacial lake outburst flood using HEC-RAS and its significance for future hazard assessments: An example from Lake 513 in the Cordillera Blanca, Peru,” *Nat. Hazards*, vol. 71, no. 3, pp. 1617–1638, 2014.
- [3] M. Christen, J. Kowalski, and P. Bartelt, “RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain,” *Cold Reg. Sci. Technol.*, vol. 63, no. 1–2, pp. 1–14, 2010.
- [4] F. Frank, B. W. McArdell, C. Huggel, and A. Vieli, “The importance of entrainment and bulking on debris flow runout modeling: Examples from the Swiss Alps,” *Nat. Hazards Earth Syst. Sci.*, vol. 15, no. 11, pp. 2569–2583, 2015.
- [5] J. Gan and Y. X. Zhang, “Numerical simulation of debris flow runout using RAMMS: A case study of Luzhuang gully in China,” *C. - Comput. Model. Eng. Sci.*, vol. 121, no. 3, pp. 981–1009, 2019.
- [6] F. Zimmermann, B. W. McArdell, C. Rickli, and C. Scheidl, “2D runout modelling of hillslope debris flows, based on well-documented events in Switzerland,” *Geosci.*, vol. 10, no. 2, pp. 1–17, 2020.
- [7] Z. Engel, J. Česák, and V. R. Escobar, “Rainfall-related debris flows in Carhuacocha Valley, Cordillera Huayhuash, Peru,” *Landslides*, vol. 8, no. 3, pp. 269–278, 2011.
- [8] H. L. Pan, Y. J. Jiang, J. Wang, and G. Q. Ou, “Rainfall threshold calculation for debris flow early warning in areas with scarcity of data,” *Nat. Hazards Earth Syst. Sci.*, vol. 18, no. 5, pp. 1395–1409, 2018.
- [9] J. Suárez, *Deslizamientos*. Santander, Colombia: Universidad Industrial de Santander UIS, 2009, vol. 1.
- [10] E. Destro, W. Amponsaha, E. I. Nikolopoulos, L. Marchi, F. Marra, D. Zoccatelli, and M. Borga, “Coupled prediction of flash flood response and debris flow occurrence: Application on an alpine extreme flood event,” *Journal of Hydrology*, vol. 558, pp. 225–237, Mar. 2018.
- [11] B. Bout, L. Lombardo, C. J. van Westen, and V. G. Jetten, “Integration of two-phase solid fluid equations in a catchment model for flashfloods, debris flows and shallow slope failures,” *Environmental Modelling & Software*, vol. 105, pp. 1–16, 2018.
- [12] E. Bladé, L. Cea, G. Corestein, E. Escolano, J. Puertas, E. Vázquez-Cendón, J. Dolz, and A. Coll, “Iber: herramienta de simulación numérica del flujo en ríos,” *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, vol. 30, no. 1, pp. 1–10, 2014.
- [13] M. A. Somos-Valenzuela, R. E. Chisolm, D. S. Rivas, C. Portocarrero, and D. C. McKinney, “Modeling a glacial lake outburst flood process chain: the case of Lake Palcacocha and Huaraz, Peru,” *Hydrology and Earth System Sciences*, vol. 20, no. 6, pp. 2519–2543, 2016.
- [14] R. Worni, C. Huggel, J. J. Clague, Y. Schaub, and M. Stoffel, “Coupling glacial lake impact, dam breach, and flood processes: A modeling perspective,” *Geomorphology*, vol. 224, pp. 161–176, Nov. 2014.
- [15] M. Álvarez, J. Puertas, E. Peña, and M. Bermúdez, “Two-dimensional dam-break flood analysis in data-scarce regions: The case study of Chipembe dam, Mozambique,” *Water (Switzerland)*, vol. 9, no. 6, p. 432, Jun. 2017.
- [16] H. Capart, D. Young, and Y. Zech, *Dam-Break Induced Debris Flow*, W. McCaffrey, B. Kneller, and J. Peakall, Eds. Wiley, 2009.
- [17] C. Rodríguez-Morata, S. Villacorta, M. Stoffel, and J. A. Ballesteros-Cánovas, “Assessing strategies to mitigate debris-flow risk in Abancay province, south-central Peruvian Andes,” *Geomorphology*, vol. 342, pp. 127–139, Oct. 2019.
- [18] M. Mergili, S. P. Pudasaini, A. Emmer, J. T. Fischer, A. Cochachin, and H. Frey, “Reconstruction of the 1941 GLOF process chain at Lake Palcacocha (Cordillera Blanca, Peru),” *Hydrology and Earth System Sciences*, vol. 24, no. 1, pp. 93–114, 2020.
- [19] J. F. Melendez de la Cruz, “Geomorfología y peligros de la subcuenca del río Auqui (Calcahuayhuanca), Huaraz, Ancash,” *Investigaciones Sociales*, vol. 12, no. 21, pp. 145–160, 2014.
- [20] J. L. Prieto Calderón, R. M.-A. López, J. Taboada-Castro, I. Montequi-Martín, and G. Sanz-Lobón, “Rotura de la presa de Vega de Tera, simulación hidráulica de la propagación de la avenida (Zamora, España),” *DYNA*, vol. 84, no. 204, pp. 45–54, Dec. 2017.
- [21] V. Vilímek, M. L. Zapata, and J. Stemberk, “Slope movements in Callejón de Huaylas, Peru,” *Acta Universitatis Carolinae, Geographica*, vol. 35, pp. 39–51, 2000.
- [22] M. Ilbay-Yupa, R. Zubieta Barragán, and W. Lavado-Casimiro, “Regionalización de la precipitación, su agresividad y concentración en la cuenca del río Guayas, Ecuador,” *Revista de Ciencias de la Vida*, vol. 30, no. 2, pp. 57–76, 2019.
- [23] P. Duque-Sarango, D. M. Patiño, and X. E. López, “Evaluación del Sistema de Modelamiento Hidrológico HEC-HMS para la Simulación Hidrológica de una Microcuenca Andina Tropical,” *Información tecnológica*, vol. 30, no. 6, pp. 351–362, 2019.
- [24] D. Schneider, C. Huggel, A. Cochachin, S. Guillén, and J. García, “Mapping hazards from glacier lake outburst floods based on modelling of process cascades at Lake

- 513, Carhuaz, Peru,” *Advances in Geosciences*, vol. 35, pp. 145–155, 2010.
- [25] Y. Deubelbeiss and C. Graf, “Two different starting conditions in numerical debris flow models—Case study at Dorfbach, Randa (Valais, Switzerland),” in *Jahrestagung der Schweizerischen Geomorphologischen Gesellschaft Reunión anual de la Sociedad Suiza de Geomorfología*, 2013, pp. 125–138.
- [26] H. Il Kim and K. Y. Han, “Linking Hydraulic Modeling with a Machine Learning Approach for Extreme Flood Prediction and Response,” *Atmosphere (Basel)*, vol. 11, no. 9, p. 987, 2020.

Basics of Methodology of Creating and Developing the Class of Natural Engineering Systems in Water Resources Management

Vladimir L. Bondarenko ^{*,a} , Elguja D. Khetsuriani ^b , Allaberdy I. Yliasov ^c ,
Elena A. Semenova ^d , Valery N. Azarov ^e 

^a*Novocherkassk Engineering And Land Reclamation Institute, Don State Agrarian University, Novocherkassk, Rostov-on-Don, Russia*

^b*M. I. Platov South Russian state Polytechnic University (NPI), Don state technical University, Rostov-on-Don, Novocherkassk, Russia*

^c*Asian branch of the WTL (United States), Moscow, Russia*

^d*School of Caucasian hospitality of the Pyatigorsk branch of the North Caucasus Federal University, Pyatigorsk, Russia*

^e*Volgograd State Technical University, Volgograd, Russia*

Abstract.- This study uses the system approach as a method of describing the relations and interaction among natural and technogenic objects and the people to elaborate the basics of the methodology of creating and developing the Natural Environment-Object of Activity-Population (NENV-OA-PPL) class of natural engineering systems (NES) for managing water resources in various branches of economy. The methodological basics elaborated for the class of natural engineering systems in question have been translated into practice in designing, building, and running water-management facilities, referred to as objects of activity, in conformity with applicable regulatory environmental requirements in effect. The basic system notions defined for the considered class of systems according to the study results are structure, component, element, and the paramount role of the whole that includes the notions of natural and technogenic constituents and environmental acceptability. The study results were time-tested in the course of designing and building the Zelenchukskaya Hydroelectric-Pumped Storage Power Plant (HPSPP) located in the basin geosystem of the Upper Kuban in the Karachay-Cherkess Republic in the North Caucasus.

Keywords: system; natural environment; object of activities; monitoring; environmental status; water quality.

Fundamentos metodológicos para la creación y desarrollo de clase de sistemas de ingeniería natural en la gestión de recursos hídricos

Resumen.- Este estudio utiliza el enfoque sistémico como método de descripción de las relaciones e interacción entre los objetos naturales y tecnogénicos y las personas para elaborar los conceptos básicos de la metodología de creación y desarrollo de la clase Entorno natural-Objeto de actividad-Población (NENV OA-PPL) de sistemas de ingeniería natural (NES) para la gestión de los recursos hídricos en diversas ramas de la economía. Los fundamentos metodológicos elaborados para la clase de sistemas de ingeniería natural en cuestión se han llevado a la práctica en el diseño, la construcción y el funcionamiento de las instalaciones de gestión del agua, denominadas objetos de actividad, de conformidad con los requisitos ambientales reglamentarios aplicables en vigor. Las nociones de sistema básico definidas para la clase considerada de acuerdo con los resultados del estudio son estructura, componente, elemento y el papel primordial del todo que incluye las nociones de constituyentes naturales y tecnogénicas y aceptabilidad ambiental. Los resultados del estudio fueron probados en el curso del diseño y construcción de la Central Hidroeléctrica de Almacenamiento por Bombeo Zelenchukskaya ubicada en el geosistema de cuenca del Alto Kubán en la República de Karachay-Cherkess en el Cáucaso Norte.

Palabras clave: sistema; entorno natural; objeto de actividades; monitoreo; estatus ambiental; calidad del agua.

Received: December 21, 2020.

Accepted: February 24, 2021.

1. Introduction

As noted by Michael E. Webber in 1990 [1], modern paradigms of social development in the Man-Nature-Society global system are characterized by ten interlinked issues enumerated

* Correspondence author:

e-mail: nimi.TbiP@yandex.ru (V. Bondarenko)

hereafter in the descending order of priority: *Energy, Water, Food, Ecology, Poverty, War and Terrorism, Diseases, Education, Democracy, and Population*. The most important of them are the ones related to *energy, water, food, and ecology* without resolving which the other six will be impossible to resolve either.

If to proceed from the system-level relation among energy, water, food, and ecology, the issue with water as a renewable natural resource is shaped at the system level by two circumstances. On the one hand, it is shaped by the natural global moisture cycle in an amount of $5777 \times 10^3 \text{ km}^3$; this cycle is continuously repeated due to an influx of solar energy of 13,500 TW [2], that reaches the terrestrial surface with an area of $149,1 \times 10^6 \text{ km}^2$ and the World Ocean with an area of $361,1 \times 10^6 \text{ km}^2$.

On the other hand, the cycle is shaped by modern technologies of utilizing water resources in multiple kinds of economic activity, where the main areas of water consumption and utilization are, respectively, the multipurpose water supply of urban households and populated localities, agricultural production in irrigated lands and farming; electricity generation at hydropower plants (HPP), pumped storage power plants (PSPP), combined heat power plants (CHPP), and nuclear power plants (NPP), industrial manufacturing.

The quantitative indices of natural hydrological processes observed from the hierarchical level of the global formation of water resources within the Earth's biosphere $W_{bserth} = 1 \times 10^{10} \text{ km}^2$ to regional basin geosystems, for example, in the North Caucasus in the South of Russia, including the geosystems of such rivers as Kuban (with water catchment area $F_{wtctch} = 57,900 \text{ km}^2$), Terek $F_{wtctch} = 43,200 \text{ km}^2$, and Lower Don $F_{wtctch} = 105,100 \text{ km}^2$, are shaped depending on global hydrological processes registered in dry, medium-, and high-water years. The economic activities practiced in the confines of the basin geosystems largely shape the considered region's environmental state (ENVST) as a factor of environmental safety (ENVS) and quality of water resources in use [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. At the system level of energy and

entropy, the environmental safety in the considered space of the basin geosystem is determined by the kind and intensity of *matter, energy, and information streams* from natural objects and economic facilities as sources determined by system-level integrated environmental monitoring [13]. As shown by the daily realities of economic activity, the utilization of water resources in actual production and engineering processes determines the relations and interaction within the natural environment of the basin geosystem in the form of its biotic and abiotic components, technogenic waterworks facilities represented by various kinds of hydraulic structures (impounding reservoirs, water intakes, etc.) called objects of activity, and the population living in the basin geosystem's territory [14, 15, 16]. The system-level relations and interaction between the natural environment and the vital necessities of the population determine a particular class of natural engineering systems called Natural Environment-Object of Activity-Population systems (Figure 1). Their influence extends to vast spaces in the ground layers of the atmosphere at heights of up to 10 km, where rain and snow are formed ($W_{(atm.)} = F_{wtctch} \cdot 10 \text{ km}^3$, upper earth crust layers down to 300 m in depth ($W_{(lts.)} = F_{wtctch} \cdot 0,3 \text{ km}^3$, where the subterranean flow of water is formed that runs into the fluvial network of the water catchment territory. This spatial interaction among the objects of activity, the natural environment, and the population determines the distinctness of the considered class of natural engineering systems. This distinctness requires conducting respective methodological studies for resolving the complicated issues of ensuring a sustainable level of environmental safety in the influence domains of the object of activity as a factor shaping the health and living standards of the population.

2. Materials and Methods

Methodologically, the creation and development of the NENV-OA-PPL class of natural engineering systems is based on system analysis, as a method of analyzing management arrangements, and system-based approach, as a way to create and use

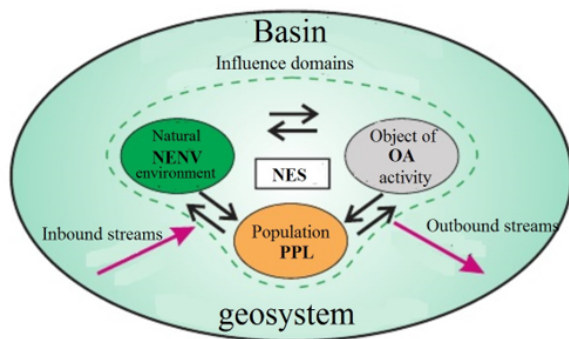


Figure 1: NENV-OA-PPL NES

as normative the description of the considered system, where water resource utilization objects are considered part of the NENV-OA-PPL natural engineering systems. Both, system analysis and system approach used in studying the NENV-OA-PPL class of natural engineering systems are based on the core notion in which a part is not perceived without perceiving the whole and is measured against the whole.

The system is a very broad notion covering combinations of interrelations and interaction, organization, etc. The first person is ever known to use it was Aristotle. The underlying ideas of the systems theory are found in Hegel's works and confine to the following: the whole is something bigger than the sum of its parts and determines their nature; parts cannot be perceived if considered beyond the whole; they are interdependent and involved in permanent interrelations [3].

The study of NENV-OA-PPL natural engineering systems involves considering the natural environment proper, the object of activity as a set of various types of hydraulic facilities, and vital necessities of the population living within the basin geosystem, where the quantitative and qualitative indicators of its water resources are shaped [4, 5, 17, 13, 8, 10, 1, 11, 18].

To understand the relations and interaction among the components of the considered system and their constituent elements, it is necessary to apply a certain methodology, where the analogy method and physical modeling play a certain role. For example, this method is applied to scaled-down models in the studying of design withdrawal of

amounts of water from a water facility. The analogy method is based on the main idea of equivalence between an actual and an abstract system. This relation of equivalence sheds light on the identical aspects in their behavior.

The basis of the methodology (BM) of creating and developing the NENV-OA-PPL class of natural engineering systems must have axiomatic prerequisites, which makes it necessary to form such by meeting the following conditions [4, 5, 19, 8, 11, 20]:

1. existence,
2. multitude of structural formations,
3. unity,
4. sufficiency

Proceeding from the modern requirements on ensuring environmental safety in the domains of influence of the object of activity, that extend to the entire territory of the basin geosystem, where quantitative and qualitative water resource indicators are formed, it is necessary to conduct a set of studies for evaluating the influence of objects of activity on quantitatively and qualitatively determined natural environments.

The procedure of evaluating environmental effects (EEE) as the technogenic component, that is related to and interacts with the natural component as part of NENV-OA-PPL NES, involves using the system-level integrated environmental monitoring process chart (Figure 2), tested and certified in the building of the Zelenchukskaya HPSPP and operation of irrigation systems in the basin of river Kuban in the Stavropol' krai. The system-level integrated environmental monitoring stipulates geochemical, atmochemical, lithochemical, hydrochemical, biochemical, and geobotanical tests based on which the integral evaluation of ES in OA influence domains is formed.

The aspects that are considered as part of the natural constituent as a component of the natural environment are the kinds of influence on the soil cover, hydrosphere of the fluvial network within the limits of the water catchment area, in the surface layers of the atmosphere, and the Earth crust's upper layers of down to 30 m in depth.

The aspects that are considered as part of the technogenic constituent as a component of

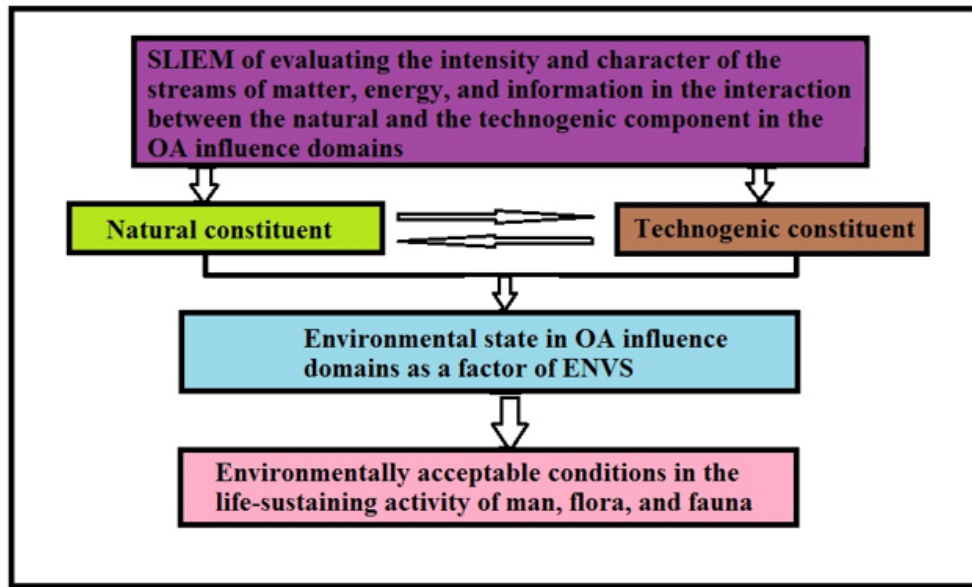


Figure 2: Process chart of conducting the system-level integrated environmental monitoring for evaluating the environmental state in the influence domains of the object of activity (technogenic constituent) interacting with the natural environment (natural constituent) as part of NENV-OA-PPL natural engineering systems

the natural environment are a set of hydraulic engineering structures intended for regulating the water runoff and using it in practice in production and engineering processes of economic activities.

3. Results

The notion of the system, where a part can not be perceived without perceiving the whole and is measured against it, is the basic notion formed in the methodology for the NENV-OA-PPL class of natural engineering systems for utilizing water resources in various process flow charts of economic activity.

In the NENV-OA-PPL class of natural engineering systems, a part is considered a separate component with its elements, for example, a natural environment with a diversity of biotic elements of fauna and such abiotic elements as atmosphere, upper earth crust levels, soil mantle, fluvial network, the changes in which under the influence of the technogenic component of OA are analyzed within the framework of the NENV-OA-PPL natural engineering system. The interaction

and relations among the components of this system are managed following the main principle of system-level integrity. This principle determines the prevalence of the whole over its parts and is the core and dominant feature in the working processes of this system within the spatial confines of the considered basin geosystem [21, 22, 23, 15, 13, 24].

The system-level conditions [25, 21, 3] formulated for validating the theoretical basics of the methodology for the NENV-OA-PPL class of natural engineering systems are exposed below.

- Condition (1) shapes the very existence of this natural engineering system as its fundamental feature; therefore, we characterize the system's existence through its forms in space or as time, or as the motion of matter, energy, and information streams, or as various combinations of these three forms.
- Condition (2) determines the multitude of natural (biotic and abiotic) and technogenic components in the form of various engineering structures, buildings, etc. This condition has to be taken into account when building NENV-OA-PPL natural engineering systems.
- Condition (3) is unity and can be understood

in two ways: on the one hand, as the relation between natural and technogenic components that favors the generation of additive, un-additive, additive-unadditive properties new to both, them taken separately and in their totality; on the other hand, they can be understood as separate components. This condition is of fundamental significance.

- Condition (4) is sufficiency understood in the sense in which the need for a sufficient amount of material (and conditions) for creating something is discussed.

The parameters and characteristics [21, 22, 23, 26, 27] determined based on the elaborated system-level integrated environmental monitoring methodology applied in building and the initial phase of running the Zelenchukskaya HPSPP and by the example of the basin geosystem of the Upper Kuban, with spatial confines $W_{bngs} = 113,300 \text{ km}^3$, water catchment area $F_{(wtctch.)} = 11 \times 10^2 \text{ km}^2$, near-terrestrial atmospheric layers $W_{(atm.)} = 11 \times 10^4 \text{ km}^3$, and upper earth crust $W_{(ercr.)} = 3,3 \times 10^3 \text{ km}^3$ are exposed below.

1. The nature and kinds of technogenic effects on the natural environment were determined against the respective classification features and characterized by the supply to and withdrawal from the ambient medium. These parameters were used to determine the boundaries of influence domains (I-V) and their division in low-active (I), active (II-IV), and very active (III-IV). The environmental safety levels in influence domains I-V of the object of activity were evaluated by determining quantitative (Π_j) and qualitative (Π_i) criteria indicators of the environmental state as a factor of environmental safety.
2. The dust burden on the water catchment territory of the Upper Kuban was measured (Figure 1).
3. Aggregate atmochemical pollution of the air basin in the Upper Kuban.
4. According to the evaluation of their state, the surface waters within the fluvial network of the Upper Kuban met utility and drinking water supply requirements.
5. The influence of the Zelenchukskaya HPSPP on the soil mantle and underlying rock does not lead to irreversible processes in the humic layer formation.
6. The protection level of the subterranean waters in the upper earth crust layers, where the subterranean water flow forms within the water catchment territory of the fluvial network of the Upper Kuban, was evaluated as well.
7. The environmental state (Figure 3) as a factor of environmental safety within the basin geosystem of the Upper Kuban [2], where the population numbers 350K people, was exposed to integral estimation. The estimation results were used to measure the severity of environmental issues across the municipal entities in the water catchment territory according to the five-grade scale, with 1, 2, 3, 4, and 5 points indicating satisfactory, high, critical, crisis, and catastrophic severity, respectively.

4. Discussion

The main condition for the sustainable performance and development of the NENV-OA-PPL class of natural engineering systems is that the considered natural engineering system is located in the natural environment surrounded, in its respect, by the basin geosystems of higher hierarchical levels to the Earth's biosphere. The analogy between the mechanisms and principles of the Earth's biospheric structures and the principles effective in local basin geosystems, within which NENV-OA-PPL natural engineering systems work, is considered as the working assumption of the methodological basics for studying and developing the class of systems in question [19]. As shown by analyzing more than ten thousand active objects of

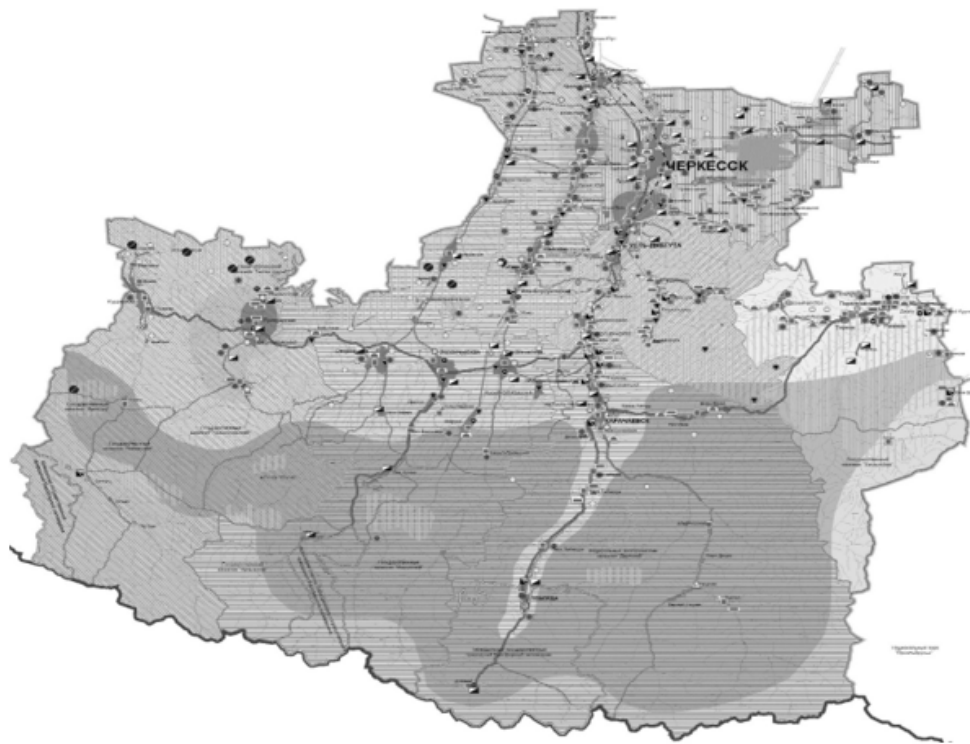


Figure 3: Integral assessment of environmental state within the basin geosystem of the Upper Kuban

activity within the confines of the basin geosystems of rivers Kuban ($W_{BGC} = 597,49 \times 10^3 \text{ km}^3$), Terek ($W_{BGT} = 4439 \times 10^3 \text{ km}^3$), and Lower Don ($W_{BGLD} = 1040,3e3 \text{ km}^3$), where a steady growth in water consumption is observed, the NENV-OA-PPL natural engineering system considered at the system level is determined as a certain pattern of regular communication and interaction among the natural environment, the object of activity, and the residing population. This system-level regularity in the relations among the natural and technogenic components and the population as part of the considered natural engineering system is formed in the phase of making project-based decisions about building objects of activity in the fluvial network for ensuring the targeted management of water resources utilization. At the spatiotemporal level, the management of water resource formation and utilization determines the system integrity principle that preconditions the prevalence of the whole over the parts of the considered system. The paramount role of integrity for NENV-OA-PPL natural engineering is core to the interaction and relations both, within the

system and between the system and the ambient medium [2, 28]. The activity of NENV-OA-PPL natural engineering systems stipulates the set of intrasystem changes in natural structural formations (fluvial network, surface layers of the atmosphere, upper earth crust layers, vegetation, etc.), where the environment must hold the leading role in the interaction and relations with objects of activity. The environment's leading role within a basin geosystem is possible on the condition that the environment's objective reality is optimally epitomized.

In NENV-OA-PPL natural engineering systems, this reality is understood as the active framework of control from outside for maintaining the development within OA influence domains and, respectively, through maintaining the paramount role of the whole played by the hydrological processes that form water runoff within the spatial confines of a basin geosystem as part of basin geosystems of higher levels in the hierarchy of the Earth's biosphere. The system interaction and relations between the spatial confines of a particular basin geosystem and the basin geosystems of

higher hierarchical levels in the global moisture cycle determine the prevalence of the whole over its components and their elements that belong to the considered NENV-OA-PPL natural engineering system. This prevalence shows in the subordination of individual parts to the system's overall performance in the environment [22, 23, 26, 27, 18].

The system-level principles distinguished according to the results of the system analysis of the interaction and relations among the natural environment, objects of activity, and population within active NENV-OA-PPL natural engineering systems work as follows:

- integrity determines the impossibility of reducing a system's properties to the sum of properties of the whole;
- structuredness determines the possibility of describing a system through networks of interaction and relations among its components, the dependence of each component with elements, properties, and relations on its place within the spatial confines of a basin geosystem and functions performed within the whole;
- hierarchy, where each component of the system is one of the components of the system of a higher hierarchical level;
- multiplicity of descriptions of the considered system means that, to study appropriately, it is necessary to build a multitude of various models each of which will describe a certain aspect of the system.

We should note that the core and dominant operational principle among all of the identified system principles in the considered class of systems is the paramount role of the whole that determines the control from outside over the interaction and relations among the system's components [3, 29, 30, 31, 32, 16, 20].

The notion of the paramount role of the whole formulated while developing the basics of the methodology of the class of systems in question is formulated below. The paramount

role of the whole in NENV-OA-PPL natural engineering systems determines the dominant influence of natural moisture cycles in basin systems of higher hierarchical levels within the Earth's biosphere on the formation of water runoff in the influence domains of an object of activity and the consistent purposeful utilization of a part of water runoff in production and engineering processes of economic activities by creating the Object-of-Activity complex that interacts with and is related to the Natural Environment and the residing Population.

According to the study results, the basic system notions for developing the methodology of creating NENV-OA-PPL natural engineering systems have been determined that reflect the essential properties of the relations among the natural environment, objects of activity, and population for all of their contradictions and in the development within the considered system.

The notion of structure in NENV-OA-PPL natural engineering systems indicates those things, that remain relatively unchanged after the transformations related to the water runoff formation within a basin geosystem, and their purposeful utilization. The structure of a NENV-OA-PPL natural engineering system is determined by its function, by the location of the natural environment and technogenic objects of activity relative to the water facility, and, timewise, by the ordered position and state of these components in space and across the seasons within the basin geosystem.

The notion of a component includes abiotic (climate, hydrosphere, atmosphere, upper earth crust layers, soil mantle) and biotic (flora and fauna) constituents, technogenic constituents in the form of objects of activity, and social constituents represented by the people living in the OA influence domains.

The notion of an element is presented as an indivisible part involved in interactions and relations within a considered component.

These basic notions were used to formulate the following spatiotemporal features of the NES in question as axiomatic assumptions [4, 6, 33, 9, 34, 35].

- The properties of the NENV-OA-PPL class of natural engineering systems show in the interaction and relations among the objects of activity, natural environment, and population within the spatial confines of the basin geosystem, where water resources are formed;
- The system's properties, in general, are not reduced to the sum of the properties of its parts, i.e., objects of activity, natural environment, population, and their elements;
- The NENV-OA-PPL class of natural engineering systems is open to the influence of the ambient natural medium within the spatial confines of the considered basin geosystem that is part of systems of higher hierarchical system levels;
- The steady rise in the capabilities of NENV-OA-PPL natural engineering systems affects the ambient natural medium and socio-economic environment within the spatial confines of the basin geosystem;
- The population has an integral demand for utilizing water resources in its life-sustaining and economic activities;
- There is a steady trend for expanded opportunities for ensuring an increase in the system's effective capacity (P) by adopting more advanced technologies of utilizing water resources with a better performance coefficient;
- There is a steady trend for a decelerated growth in lost capacity (G) in the system due to its rising P;
- A better resource efficiency and reduced energy costs in the phases of project-based decision making, building, and operation;
- The position of objects of activity within the spatial confines of the basin geosystem determines the spatial structure of the NENV-OA-PPL class of natural engineering systems;
- The temporal structure shows in pragmatically operating the objects of activity across the seasons;
- The functional structure is determined by the objects of activity in their purposeful interaction and relations with the natural environment and the people living within the spatial confines of the basin geosystem, depending on specific water consumers and users.

The interrelation of the natural processes of water runoff formation within the confines of the basin geosystem under the influence of hydrological processes in the systems of higher hierarchical levels related to the global moisture cycle determine the natural constituent of the prevalence of the whole over the natural and technogenic components in the considered NENV-OA-PPL natural engineering systems. The technogenic constituent of the prevalent role of the whole is determined by the intrasystem relations and interaction among the objects of activity, natural environment, and population, which determines the spatial boundaries of the OA influence domains within the local basin geosystem [4, 5, 17, 6, 8, 10].

The natural constituent of the paramount role of the whole is originally formed by the global moisture cycle within the Earth's biosphere and uses hierarchical system interrelations to reach local basin geosystems, in which NENV-OA-PPL natural engineering systems are functional.

The system-level relations among the processes that take place in the Earth's biosphere and form natural processes, hydrological ones included, within basin geosystems of lower hierarchical levels, determine the natural constituent of the paramount role of the whole in the water runoff formation in NENV-OA-PPL natural engineering systems. It should be noted that the hydrological processes formed within spatial confines of basin geosystems at different hierarchical levels depend on concrete natural factors of location on the Earth's surface.

The technogenic constituent of the paramount role of the whole in the purposeful operation

of objects of activity included in NENV-OA-PPL natural engineering systems is the subsequent system-level link in the sequence of practically utilizing water resources in various production and engineering processes of economic activities. As shown by field and laboratory test results, this constituent fully depends on the performance, design, and other characteristics of OA, which is important to consider in design decision making [3, 23, 4, 5, 17, 8, 10, 11, 12, 24].

The performance dependence of the objects of activity included in NENV-OA-PPL natural engineering systems on the paramount role of the whole is determined by environmental acceptability (EA) that is a peculiar copy of a natural system, in the transformation of energy forms with entropy growth on the wane [8]. The environmental acceptability of objects of activity is manifested in design fineness and performance reliability in various modes of running these objects as part of NENV-OA-PPL natural engineering systems and, at the system level, in maintaining and regulating the intensity of matter, energy, and information streams within the spatial confines of a basin geosystem, which allows maintaining the paramount role of the integrity of the technogenic component. According to the investigation results and in furtherance of the notion of the technogenic constituent of the prevalent role of integrity, the functional and constructional environmental acceptability of objects of activity as parts of NENV-OA-PPL natural engineering systems can be characterized by the following conceptual assumptions [3, 4, 5, 6, 33, 8, 10, 11].

1. The environmental acceptability of objects of activity depends on the design fineness of hydraulic facilities in use, related structures, including fish protection and passing works as well as functional and auxiliary buildings;
2. Environmental acceptability is characterized by the interaction and relations between the objects of activity and the natural environment, which shapes the environmental state under the influence of changes produced in the motion of matter, energy, and information streams;

3. Environmental acceptability depends on the resource intensity and energy efficiency of the production and engineering processes related to the building and operation of objects of activity;
4. Environmental safety is the main indicator used in assessing objects of activity for environmental acceptability proceeding from the cause-consequence logic of interaction and relations between the objects of activity and the natural environment within the spatial confines of influence domains;
5. Environmental acceptability attends self-organization as the all-purpose model of interrelated transformations in the natural environment under the influence of objects of activity included in NENV-OA-PPL natural engineering systems;
6. The environmental acceptability of objects of activity facilitates the dominant role of natural transformations in the natural environment and reduction in the pace of entropy growth rates;
7. The environmental acceptability of objects of activity included in NENV-OA-PPL natural engineering systems facilitates the adaptation and adjustment of the object-of-activity design parts to the ambient medium via structural transformations in the natural environment;
8. Environmental acceptability is interlinked with self-organization in terms of enhancement of the structural parts of objects of activity included in NENV-OA-PPL natural engineering systems;
9. The environmental acceptability of objects of activity as part of NENV-OA-PPL natural engineering systems is characterized by quantitative and qualitative design and performance indicators;
10. The environmental acceptability of objects of activity as part of NENV-OA-PPL natural engineering systems is assessed by system-level integrated environmental monitoring.

The object of activity in a NENV-OA-PPL natural engineering system is the core technogenic component that closes on itself in the influence

domains of certain spatial confines of the basic geosystem in the causal relation with environmental acceptability and forms an environmental state manifested in environmental safety as an important factor of preserving the health and living standards of people, flora, and fauna.

5. Conclusions

We have elaborated the basics of the methodology of creating and developing the NENV-OA-PPL class of natural engineering systems inalienably linked with the issues of energy, water, food, and ecology, proceeding from the unity of action among nature, the economic activities are undertaken to utilize water resources, and the current issue with water as a renewable natural resource that comes second among the four most important issues of energy, water, food, and ecology. The basics of the methodology in question have been elaborated by system analysis, system approach, and system-level integrated environmental monitoring conducted to evaluate the kind of effects produced by the objects of activity already and being built in the North Caucasus and rely on the results of studying the relations and interaction among the technogenic components of the objects of activity, the natural components of the natural environment, and the population residing in the considered confines of the basin geosystem as part of the NENV-OA-PPL natural engineering system.

Methodologically, the basics of the methodology of creating and developing the NENV-OA-PPL class of natural engineering systems rely on the system analysis and system approach used to validate the fundamental principles of the paramount role of whole, integrity, structuredness, hierarchy, and multiplicity; the basic notions of structure, component, and elements; the basic indicators [14], and the conceptual assumptions on environmental acceptability in the interaction between the object of activity and the population, which forms environmental states in the influence domains of objects of activity as a factor shaping the health and living standards of the population.

Acknowledgments

We thank the Director of the Zelenchukskaya HPSPP for granting us the opportunity to conduct the system-level integrated environmental monitoring in the course of building the HPSPP and in the initial period of its operation.

6. References

- [1] M. E. Webber, "Energy, Water, and Food Problems Must Be Solved Together," *Scientific American*, vol. 312, no. 2, 2015.
- [2] G. R. Davis, "Energy for Planet Earth," *Scientific American*, vol. 263, no. 3, pp. 7–15, Sep. 1990.
- [3] V. L. Bondarenko, A. I. Yliasov, and E. D. Khetsuriani, *Scientific and Methodological Basics of Natural Engineering Systems for Utilizing Water Resources by the Example of Basin Geosystem Territories*. Novocherkassk: Novocherkassk Engineering and Land Reclamation Institute; M. I. Platov South Russian state Polytechnic University (NPI), 2019.
- [4] R. Pereira Gomes, J. A. de Paula Silva, M. C. Carvalho Junior, W. Castro Amorin Alburquerque, P. S. Scalize, A. Rodrigues Galvão Filho, D. d. J. Pires, J. D. Gonçalves Vieira, and L. C. Carneiro, "Evaluation of the raw water quality: physicochemical and toxicological approaches," *Environmental Geochemistry and Health*, vol. 41, no. 6, pp. 2425–2442, Dec. 2019.
- [5] S. Lu, X. Wu, H. Sun, W. Li, and Y. Tang, "The multi-user evolutionary game simulation in water quality-based water source system," *Environmental Geochemistry and Health*, vol. 42, pp. 863–879, 2020.
- [6] E. D. Khetsuriani, V. L. Bondarenko, and N. A. Polianskii, "Methodological Bases of Creation and Development of a New Type of Natural and Technical Systems of Multipurpose Water Use in Urban Areas," in *Proceedings of the International science and technology conference "Earth science", IOP Conference Series: Earth and Environmental Science*, 2019.
- [7] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "Methodological Framework for the Assessment of the Zones of Influence of Technological Water Systems Multipurpose Water Supply Urban Households and the Economy,," *Biosphere compatibility: human, region, technology*, vol. 28, no. 4, pp. 25–35, 2019.
- [8] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "The results of the research on the pipelines protection from Dreissena on the water intake technological complexes of multi-purpose water supply systems for urban farms," in *Proceedings of the International Scientific Conference "Construction and Architecture: Theory and Practice of Innovative*

- Development", IOP Conference Series: Materials Science and Engineering*, 2019.
- [9] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "Lightweight constructions in technical water supply systems of thermal and nuclear power plants," in *Proceedings of the International Scientific Conference "Construction and Architecture: Theory and Practice of Innovative Development", IOP Conference Series: Materials Science and Engineering*, 2019.
- [10] E. D. Khetsuriani, D. S. Larin, and E. V. Staseva, "Objectives of increase of capacity and lifetime of municipal solid waste dump according to density index study ," in *Proceedings of the XV International Scientific-Technical Conference "Dynamics of Technical Systems", IOP Conference Series: Materials Science and Engineering*, 2019.
- [11] Z. Wu, Z. Wu, H. Li, X. Zhang, and M. Jiyang, "Developing a strategic framework for adopting water-saving measures in construction projects," *Environ Geochem Health*, vol. 42, no. 3, pp. 955–968, 2020.
- [12] M. Groß, C. Klippstein, P. Maurer, Y. Salazar, T. Schwab, K. Treffry-Goatley, and J. Voortman, *EDS®- Water Management*. Denkendorf: Festo Didactic GmbH and Co. KG., 2014.
- [13] E. D. Khetsuriani, V. L. Bondarenko, and A. I. Yliasov, *Scientific and Methodological Basics of Environmental Safety in Watch Catchment Engineering Complexes of Multipurpose Water Supply Systems*. Novochoerkassk: Novochoerkassk Engineering and Land Reclamation Institute; M. I. Platov South Russian state Polytechnic University (NPI), 2020.
- [14] E. D. Khetsuriani, "Results of experimental research on ensuring environmental safety of water supply systems of industrial and residential settlements," *Bulletin of the Volgograd state architectural and construction University. Journal of Construction and architecture [in Russian]*, vol. 54, no. 73, pp. 165–172, 2018.
- [15] E. D. Khetsuriani, V. L. Bondarenko, and T. E. Khetsuriani, "Assessing Primary Role in Ensuring Environmental Security in the Zones of Influence of Water Intake Technology Complex Multi-Purpose Systems of Water Supply of Urban Economy," *News of Higher Educational Institutions. Construction*, vol. 717, no. 9, pp. 83–90, 2018.
- [16] M. Groß, C. Klippstein, P. Maurer, Y. Salazar, T. Schwab, K. Treffry-Goatley, J. Voortman, and C. Wehlers, *Water Treatment*. Denkendorf: Festo Didactic GmbH and Co. KG., 2013.
- [17] E. D. Khetsuriani, V. L. Bondarenko, O. V. Klimenko, and E. A. Semenova, "Methodological measurement basis environmental safety in construction and operation of water management facilities," in *Proceedings of the International Conference on Modern Trends in Manufacturing Technologies and Equipment (ICMTMTE)*, 2019.
- [18] M. Groß, P. Maurer, Y. Salazar, T. Schwab, K. Treffry-Goatley, and C. Wehlers, *Water Supply*. Denkendorf: Festo Didactic GmbH and Co. KG., 2013.
- [19] E. D. Khetsuriani and V. L. Bondarenko, "The Basics of Methodology of Creation of Nature-Technical Systems in Water Management in Multipurpose Water Supply in Urban Farms and Settlements," *Bulletin of the Volgograd state architectural and construction University. Journal of Construction and architecture [in Russian]*, vol. 54, no. 73, pp. 165–172, 2020.
- [20] M. Groß, C. Klippstein, Y. Salazar, and T. Schwab, *Wastewater disposal*. Denkendorf: Festo Didactic GmbH and Co. KG., 2014.
- [21] V. L. Bondarenko, V. V. Privalenko, A. V. Kuvalkin, E. S. Polyakov, and S. G. Pryganov, *Solving Environmental Problems of Hydraulic Engineering by the Example of Basin Geosystem of the Upper Cuban*. Rostov-on-Don: South Scientific Center of the Russian Academy of Sciences, 2009.
- [22] A. M. Chernyae, *Water Resource Potential of Russia*. Yekaterinburg: Agroekologiya Publishers of the Federal Water Resources Agency, 1998.
- [23] A. M. Chernyayev, *River Basins*. Yekaterinburg: Agroekologiya Publishers of the Federal Water Resources Agency, 1999.
- [24] M. Groß, C. Klippstein, P. Maurer, Y. Salazar, T. Schwab, K. Treffry-Goatley, and J. Voortman, *Wastewater Treatment*. Denkendorf: Festo Didactic GmbH and Co. KG., 2014.
- [25] V. L. Bondarenko, V. V. Gutenev, V. V. Privalenko, and E. S. Polyakov, "Environmental Influence Estimations at the Designing of Zelenchukskaya Hydroelectric Power Plant Reservoir," *Theoretical and Applied Ecology*, no. 1, pp. 46–55, 2007.
- [26] I. S. Rumyantsev and D. V. Kozlov, "Environmental Problems of Russia's Water Reservoirs and Constructional Hydroecology," *Gidrotekhnicheskoye stroitel'stvo (Hydraulic engineering)*, no. 6, pp. 53–54, 1999.
- [27] I. S. Rumyantsev and R. Kromer, *Utilizing Engineering Biology Techniques in Hydraulic Engineering and Environment-Oriented Construction Practices*. Moscow: MSU Publishers, 2003.
- [28] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, F. S. Nozimov, and T. E. Khetsuriani, "Soft Float-Design Watch Catchment Facility As Part of Water Catchment Process Complex of System of Multipurpose Water Supply of Municipal and Economic Facilities,," *Engineering and Construction Bulletin of the Caspian Region*, vol. 33, no. 2, pp. 23–33, 2020.
- [29] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "Development of protective measures providing environmental safety in areas affected by water-intake constructions of urban households ," in *Proceedings of the International Scientific Conference "Construction and Architecture: Theory and Practice*

of Innovative Development", IOP Conference Series: Materials Science and Engineering, 2019.

- [30] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "Assessment of qualitative indicators of drinking water and their influence on human health, as ecological safety factor of population," in *Proceedings of the International Scientific Conference "Construction and Architecture: Theory and Practice of Innovative Development", IOP Conference Series: Materials Science and Engineering, 2019.*
- [31] E. D. Khetsuriani, R. V. Fedotov, and K. S. Abrosimova, "Assessment of the Impact of Integrated Pollution on the Monitoring of Biodiversity in Fisheries," in *Proceedings of the International science and technology conference "Earth science", IOP Conference Series: Earth and Environmental Science, 2019.*
- [32] E. D. Khetsuriani, V. P. Kostuykov, D. S. Larin, and T. E. Khetsuriani, "Determination of an adequate mathematical model of research results on determining fish protection efficiency," *Znanstvena misel*, vol. 9, no. 1, pp. 89–93, 2017.
- [33] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "Innovative design solutions to ensure the environmental safety in the existing water intake technological complexes of water systems for urban farms," in *Proceedings of the International Scientific Conference "Construction and Architecture: Theory and Practice of Innovative Development", IOP Conference Series: Materials Science and Engineering, 2019.*
- [34] E. D. Khetsuriani, V. L. Bondarenko, A. I. Yliasov, and E. A. Semenova, "Evaluation of environmental safety of waterworks for urban water supply systems," in *Proceedings of the International Scientific Conference "Construction and Architecture: Theory and Practice of Innovative Development", IOP Conference Series: Materials Science and Engineering, 2019.*
- [35] E. D. Khetsuriani, T. S. Kolmakova, M. A. Yakimenko, and T. E. Khetsuriani, "Ecological safety of the water environment - the key to a healthy future," *Bulletin of the Volgograd state University of architecture and construction. Series Construction and architecture*, vol. 73, no. 54, pp. 156–164, 2018.

Probabilistic model of the relationship between ENSO and the occurrence of agricultural droughts in southern Peru

Juan Cabrera * , Jesús Mejía , Edilberto Guevara-Pérez 

Programa de Doctorado en Recursos Hídricos, Escuela de Post Grado, Universidad Nacional Agraria La Molina. Lima, Perú

Abstract.- In the article, the information on rainfall in the Province of Candarave, located in southern Peru, is analyzed and contrasted with different indices of the ENSO phenomenon to define whether there is influence of this phenomenon on the dry seasons in the region and build a probabilistic model using copula functions. The EMI (El Niño Modoki Index), ONI (Oceanic Niño Index), TNI (Trans-Niño Index), ICEN (Coastal El Niño Index), and the temperature anomalies ENSO34 (from zone 34) and ENSO 1 +2 (from zone 1 + 2) were analyzed and correlated with the three-month standardized precipitation index, SPI3, usually taken as an indicator of the occurrence of agricultural droughts. The results show that there is an association between the SPI3 index and the EMI, ENSO34, TNI and ONI indices at a significance level of 5%. The goodness-of-fit analysis shows that the Gumbel-type copula is the most representative of the phenomenon being evaluated; therefore, the corresponding probabilistic model was built. The results allow inferring at a probabilistic level the possible occurrence of agricultural droughts based on the occurrence of the ENSO phenomenon.

Keywords: ENSO; probabilistic model; agricultural drought; copula functions.

Modelación probabilística de la relación entre ENSO y sequías agrícolas en el sur de Perú

Resumen.- En el artículo se analiza la información de precipitaciones en la Provincia de Candarave, ubicada en el Sur del Perú y se contrasta con diferentes índices del fenómeno ENSO para definir si existe influencia de este fenómeno sobre las temporadas secas en la región y construir un modelo probabilístico usando funciones cópula. Se analizaron los índices EMI (El Niño Modoki Index), ONI (Oceanic Niño Index), TNI (Trans-Niño Index), ICEN (Índice Costero El Niño), y las anomalías de temperatura ENSO34 (de la zona 34) y ENSO 1+2 (de la zona 1+2), y se correlacionaron con el índice de precipitación estandarizada trimensual, SPI3, tomado usualmente como indicador de la ocurrencia de sequías agrícolas. Los resultados muestran que existe dependencia del índice SPI3 y los índices EMI, ENSO34, TNI y ONI a un nivel de significancia del 5%. El análisis de la bondad de ajuste muestra que la cópula tipo Gumbel es el más representativo del fenómeno que se evalúa por lo que, se construyó el modelo probabilístico correspondiente. Los resultados permiten inferir a nivel probabilístico la posible ocurrencia de sequías agrícolas en base a la ocurrencia del fenómeno ENSO.

Palabras clave: ENSO; modelo probabilístico; sequía agrícola; funciones cópula.

Recibido: 02 de febrero, 2021.

Aceptado: 05 de abril, 2021.

1. Introducción

Durante las últimas décadas, la ocurrencia de eventos extremos pareciera haberse intensificado en Perú [1], con un incremento en las precipitaciones en el norte y reducción de estas en la zona sur del país, ocurrencias que podrían

estar asociadas al proceso de cambio climático que actualmente envuelve al planeta. Las fuertes precipitaciones en el norte están usualmente asociadas a la presencia de El Niño-Oscilación del Sur (ENSO, por sus siglas en inglés), mientras que las sequías en el sur se asocian a sistemas de circulación anticiclónica, como La Alta de Bolivia; sin embargo, no se descarta la influencia de ENSO sobre la ocurrencia de sequías, como demostró Vega [2], quien estableció que ambos fenómenos estaban relacionados y que se requerían más estudios especialmente en lo relacionado a

* Autor para correspondencia:

Correo-e: 20180857@lamolina.edu.pe (J. Cabrera)

sequías agrícolas y meteorológicas.

El problema de las sequías en Perú afecta directamente a la población rural, cuya principal actividad económica es la agricultura; ocasiona pérdidas monetarias y trastoca el normal desarrollo de actividades socioeconómicas [3, 4, 5, 6], dando lugar a oleadas masivas de migración, como las ocurridas en la primera mitad de la década de 1980-1990 [7].

Con la finalidad de ampliar el conocimiento sobre el tema en el sur de Perú, en esta investigación se propone estudiar la probabilidad de ocurrencia conjunta de las sequías y el fenómeno ENSO, desarrollando un modelo probabilístico en base a una función cópula [8, 9, 4], utilizando el procedimiento sugerido por Rayens [10] y Salvadori [4]. Con este objetivo, se evalúan algunos índices El Niño, tales como anomalías de temperatura del agua superficial del océano en las zonas 1+2 (ENSO1+2) y 34 (ENSO34), índice oceánico El Niño (ONI), índice costero El Niño (ICEN), índice El Niño Modoki (EMI), índice de evolución El Niño (TNI); los cuales se contrastan con el índice de precipitación estandarizada de orden 3 (SPI3), seleccionado como indicador de ocurrencia de sequías. La finalidad es definir el nivel de influencia entre los mencionados índices, para luego construir el modelo en cópulas que represente la probabilidad conjunta de ocurrencia y permita estimar las probabilidades de ocurrencia de las sequías para diferentes escenarios.

La técnica propuesta ha sido utilizada en otros países [11, 12, 13, 14], y en el mismo Perú [6], para establecer relaciones entre estadísticos de sequía (duración, intensidad y magnitud) sin embargo estos estudios solamente analizan las sequías y no su relación con otros fenómenos climáticos. Recientemente, Cabrera y Mejía [15] realizaron un primer intento de construir un modelo probabilístico bivariado entre estas variables, pero utilizando menos estaciones y una menor longitud de registro, encontrando poca representatividad de las funciones marginales encontradas y recomendando el uso de probabilidades empíricas. De esta manera, el presente artículo representa el segundo avance de un modelo probabilístico mucho más complejo construido para comprender

las relaciones entre variables climáticas asociadas a la ocurrencia de sequía agrícola y servirá como referencia para futuras investigaciones y como insumo de posibles planes de gestión y/o mitigación de sequías en la región de estudio.

2. Metodología

2.1. Área de estudio y datos

Se ha seleccionado para la investigación la zona de Candarave ubicada en el Departamento de Tacna, en el sur de Perú. El área está definida por las Coordenadas UTM 350000 - 370000 Este y 8082000 - 8096000 Sur, con altitudes que varían entre 2800 y 3800 m s.n.m. La principal actividad económica es la agricultura; el riego es operado por cuatro comisiones principales: Cairani, Candarave, Huanuara y Quilahuani (ver Figura 1). La región tiene un clima moderadamente frío, con temperaturas que oscilan entre los 6,8 °C (temperatura media máxima en noviembre) y 1,9 °C (temperatura media mínima en julio). La precipitación total mensual en Candarave varía entre 0 y 55 mm, siendo diciembre, enero, febrero y marzo, los meses húmedos; y, mayo, junio y julio, los meses más secos; éstos últimos definen una escasez permanente de agua.

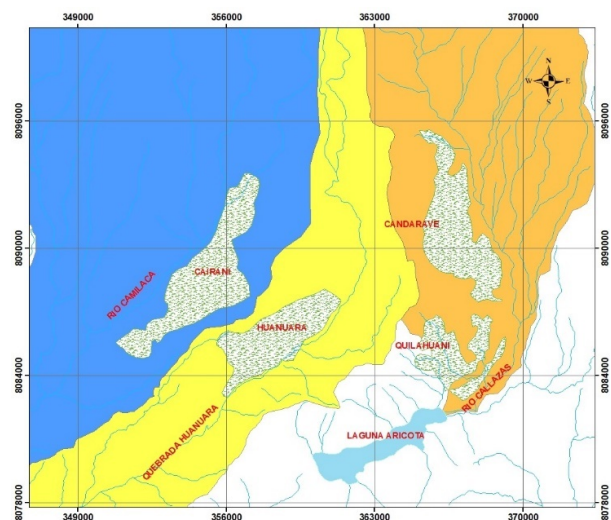


Figura 1: Zona de estudio. Los triángulos muestran la ubicación de las estaciones climáticas utilizadas

Para la investigación se utilizaron las precipitaciones mensuales del período 1964-2020

registradas en tres estaciones climáticas del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) ubicadas dentro del área de estudio. En la Tabla 1 se presentan las características geográficas de ubicación de dichas estaciones: latitud, longitud y altitud.

Tabla 1: Características de ubicación geográfica de las Estaciones climáticas utilizadas

Estación	Coordenadas geográficas – Zona 19 Datum WGS 84		
	Longitud (m)	Latitud (m)	Altitud (m s.n.m.)
ARICOTA	70°14' O	17°19' S	2843
CANDARAVE	70°15' O	17°16' S	3415
CAIRANI	70°22' O	17°17' S	3205

La información correspondiente a los índices El Niño se descargó de la base de datos del Centro Climático APEC (APCC, por su acrónimo en inglés) y del Instituto Geofísico del Perú (IGP).

2.2. Identificación de sequías

Las sequías se definen como desviaciones climáticas negativas respecto a condiciones consideradas normales de una región [3, 16]. Esta condición de humedad insuficiente puede ser causada por un déficit en la precipitación o un retraso del inicio a lo largo de un cierto período de tiempo.

Dependiendo de la duración temporal, las sequías pueden afectar en mayor o menor medida la disponibilidad hídrica y las condiciones y actividades asociadas del entorno. De esta forma se habla de una sequía meteorológica, si el déficit de precipitaciones está por debajo de lo normal; sequía agrícola, si el déficit de precipitaciones afecta el normal crecimiento de las plantas; sequía hidrológica, si el déficit afecta el almacenamiento subterráneo y el ciclo hidrológico; y, sequía socioeconómica, si el déficit afecta la oferta y demanda de un cierto producto. El presente estudio se enfoca en la ocurrencia de sequías agrícolas y, por tanto, en lo sucesivo se utilizará el término “sequía” para referirse a este tipo en especial.

Como se ha mencionado, el análisis convencional de ocurrencia de sequías se realiza en base al estudio de cuantiles o en base a índices,

denominados “índices de sequía”. En este trabajo se eligió el Índice de Precipitación Estandarizada (SPI, por su acrónimo en inglés) [17] como indicador de la ocurrencia de sequías. La selección del índice SPI se debió al bajo requerimiento de información necesario para el cálculo en comparación con otros como, por ejemplo, el Índice de Severidad de Sequías de Palmer (PDSI, por su acrónimo en inglés) [18]; pero también por la eficiencia que ha mostrado en la caracterización y ocurrencia de sequías [19, 20, 3, 4, 6, 16].

El SPI se calcula “agregando precipitaciones mensuales” en escalas de 1, 3, 6, 12 y 24 meses, de acuerdo con el efecto que se quiere capturar. Para sequías meteorológicas, cuyos efectos se manifiestan a corto plazo, se recomienda la escala de 1 mes; para sequías agrícolas, escala 1 o 3 meses; para sequías hidrológicas, escalas de 6, 12 o 24 meses. Las series agregadas se ajustan a una distribución de probabilidad y luego de transformar a una distribución normal estándar, quedando definido el índice SPI como se muestra en la ecuación (1):

$$SPI = \frac{y - \mu}{\sigma}, \quad (1)$$

donde y es la precipitación agregada mensual; μ es la media de la variable precipitación agregada mensual; y σ es la desviación estándar de la variable precipitación agregada mensual.

El modelo (1) arroja para el índice SPI valores positivos y negativos. Los valores positivos representan precipitaciones por encima de un valor medio; es decir, representan la ocurrencia de un período húmedo; los valores negativos representan precipitaciones por debajo de la media; es decir, un período seco. En base a esta caracterización, McKee [17] define cuatro categorías de sequía, las cuales se presentan en la Tabla 2.

De acuerdo con lo indicado, en el presente estudio se usa el SPI a escala de 3 meses (SPI3), como indicador de ocurrencia de sequía agrícola.

2.3. Identificación de las zonas El Niño

El fenómeno El Niño Oscilación del Sur (ENSO, por sus siglas en inglés) es un ciclo de desviaciones periódicas de la temperatura superficial del agua

Tabla 2: Categorías de sequías en base al índice SPI [17]

SPI	Categoría
0,00 a -0,99	Sequía leve
-1,00 a -1,49	Sequía moderada
-1,50 a -1,99	Sequía severa
-2,00 o menos	Sequía extrema

en el Océano Pacífico ecuatorial, y en la atmósfera, denominadas anomalías. Estas desviaciones representan incrementos o reducciones de temperatura que pueden afectar los patrones climáticos en todo el mundo, ocasionando precipitaciones intensas en algunas regiones y períodos muy secos en otras, datos que están disponibles en <https://psl.noaa.gov/data/climateindices/list/>).

El estudio de este fenómeno ha permitido definir regiones en el Pacífico Ecuatorial donde la ocurrencia de estas desviaciones está directamente asociada con los cambios climáticos en diferentes partes del planeta. En el caso de Perú, las anomalías que se presentan en la región 1+2 influyen fuertemente en la ocurrencia de precipitaciones extremas en el norte del país. La Figura 2 muestra las regiones “El Niño”.

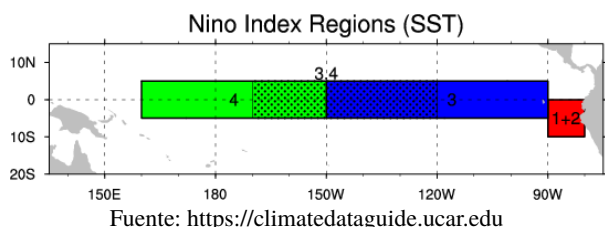


Figura 2: Regiones “El Niño”

Un primer indicador de la presencia del ENSO son las anomalías de temperatura del agua superficial del mar en cada una de las regiones “El Niño”; sin embargo, teniendo en cuenta la complejidad del fenómeno, se han desarrollado diferentes “índices” para definir su ocurrencia. Para el presente estudio se toman en cuenta las anomalías de temperatura superficial del mar en las zonas 1+2 (ENSO1+2) y 34 (ENSO34), el índice oceánico El Niño (ONI), el índice costero El Niño (ICEN), el índice El Niño Modoki (EMI), y el índice de evolución El Niño (TNI).

2.4. Medidas de asociación

La asociación entre el índice SPI3, considerado como indicador de sequía agrícola, y los índices ENSO indicados, se analizó utilizando enfoques cualitativos y cuantitativos. La dependencia cualitativa se evaluó mediante un análisis gráfico, y fue utilizado también para eliminar índices ENSO que muestran dependencia entre sí. La dependencia cuantitativa se evaluó mediante el coeficiente de correlación lineal r de Pearson, el coeficiente de correlación de rango de Spearman y el de Kendall.

2.5. Función cópula

Una cópula es una función multivariada que une o acopla funciones unidimensionales de distribución marginal correspondientes al grupo de variables que se busca analizar [10]. El concepto fue introducido por Sklar [9, 4], quien propuso que toda función de distribución F , n -dimensional, puede descomponerse en dos partes, una función de distribución marginal F_i y una cópula C , que representa la dependencia de la distribución.

Sklar [10, 21] estableció la relación entre una función de distribución acumulada conjunta F_X y una cópula mediante el llamado Teorema de Sklar. En este teorema se propone que, para una función de distribución acumulada conjunta F_X con distribuciones marginales F_{X1} y F_{X2} , existe una cópula bidimensional C tal que $\forall x \in R^2$ (ecuación (2)):

$$F_X(x_1, x_2) = C(F_{X1}(x_1), F_{X2}(x_2)). \quad (2)$$

El teorema puede generalizarse para cópulas n -dimensionales como se muestra en la ecuación (3):

$$F_X(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \quad (3)$$

Asimismo, Sklar demostró que, si las distribuciones marginales son continuas, la cópula es única.

Para el presente estudio se seleccionaron dos cópulas elípticas (normal y t), y tres cópulas arquimedianas (Clayton, Frank y Gumbel), por ser las más usuales en Hidrología y Climatología [8, 21]. En la Tabla 3 se presenta un resumen de las definiciones de cada una de las cópulas seleccionadas para el caso trivariado.

Tabla 3: Definición de cópulas trivariadas utilizadas para el estudio [4]

Copula	Cópula trivariada $C_{\theta}(u, v)$	Parámetros
Normal	$C_N(u; \Sigma) = \int_{-\infty}^{\phi^{-1}(u_1)} \dots \int_{-\infty}^{\phi^{-1}(u_3)} \frac{1}{\sqrt{(2\pi)^3 \Sigma }} \exp\left(-\frac{1}{2} X' \Sigma^{-1} X\right) dx$	Σ es la matriz de correlación
t(Student)	$C_t(u; \Sigma; \nu) = \int_{-\infty}^{t^{-1}(u_1)} \dots \int_{-\infty}^{t^{-1}(u_3)} \frac{\Gamma\left(\frac{\nu+3}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \sqrt{(\pi\nu)^3 \Sigma }} \left(1 + \frac{x' \Sigma^{-1} x}{\nu}\right)^{-\frac{\nu+3}{2}} dx$	Σ es la matriz de correlación; ν es el número de grados de libertad
Clayton(*)	$\left(\left(U_1^{-\theta_2} + U_2^{-\theta_2} - 1\right)^{\theta_1/\theta_2} + U_3^{-\theta_1} - 1\right)^{-1/\theta_1}$	$\theta_2 > \theta_1$
Frank(*)	$-\theta_1^{-1} \ln\left(C_1^{-1}\left(1 - \left(1 - C_2^{-1}\left(1 - e^{-\theta_2 U_1}\right)\left(1 - e^{-\theta_2 U_2}\right)\left(1 - e^{-\theta_1 U_3}\right)\right)^{\theta_1/\theta_2}\right)\right)$	$C_1 = 1 - e^{-\theta_1}$, $C_2 = 1 - e^{-\theta_2}$; $\theta_2 > \theta_1$
Gumbel(*)	$\exp\left(-\left[\left(-\ln U_1\right)^{\theta_2} + \left(-\ln U_2\right)^{\theta_2}\right]^{\theta_1/\theta_2} + \left(-\ln U_3\right)^{\theta_1}\right)^{1/\theta_1}$	$\theta_2 > \theta_1$

(*) Si la cópula es simétrica, $\theta_2 = \theta_1 = \theta$

Para estimar los parámetros de cópulas, se utilizó el enfoque de máxima pseudo-verosimilitud (MPL) por ser el más recomendado para análisis multivariados [21]. El método consiste en estimar pseudo-observaciones y, en base a ellas, estimar, a su vez, la función de pseudo verosimilitud a maximizar.

Para determinar el mejor ajuste a los datos y en consecuencia seleccionar la cópula, se utiliza la prueba paramétrica de Cramér-von Mises (S_n) definida mediante la ecuación (4):

$$S_n = \int_{[0,1]^d} \mathbb{C}_n(u)^2 dC_n(u), \quad (4)$$

donde: d , es el número de variables en análisis; C_n es el valor calculado utilizando la cópula empírica; y, \mathbb{C}_n es la diferencia entre el valor calculado con la cópula empírica y el calculado con la familia de cópulas en evaluación como se observa en la ecuación (5):

$$\mathbb{C}_n = \sqrt{n} (C_n - C_{\theta_n}). \quad (5)$$

Un valor alto de estadístico S_n indica el rechazo de la hipótesis nula H_0 , que propone que “la cópula empírica C pertenece a la familia de cópulas C_{θ} ”. Para el análisis indicado, se utilizó el paquete “Copula” en lenguaje R disponible en <https://CRAN.R-project.org/package=copula>.

3. Resultados y discusión

3.1. Relación entre variables

En una primera etapa, se procedió a la estimación del índice SPI3 para cada una de las estaciones incluidas en el estudio. En la Figura 3 se presentan los resultados obtenidos para el índice SPI3 calculado para cada una de las estaciones: a) Aricota; b) Cairani; c) Candarave (índice SPI3 en las ordenadas y meses en las abscisas. Las series de los índices así calculados se utilizan como variable de entrada en la construcción del modelo cópula. En la Tabla 4 se presentan las características promedio de las sequías evaluadas según el índice SPI3, extraídas de las de las gráficas dadas en la Figura 3. Los resultados muestran que el número de eventos registrados se incrementa con la altitud de la zona y, consecuentemente, el período entre eventos se reduce. Asimismo, se observa que la magnitud y la intensidad aumentan con la altitud pero el máximo SPI negativo registrado (pico) se reduce (Tabla 4). Este comportamiento indica que, a mayor altitud, mayor número de sequías, más severas pero de menor intensidad.

Con la serie generada de índices SPI3 se procedió a realizar un análisis gráfico para evaluar la dependencia con los índices ENSO. Los resultados se presentan en la Figura 4, en la cual se observa una fuerte dependencia entre los índices ENSO1+2 e ICEN, ENSO34 y ONI, TNI y EMI (ver Figura 4 para la estación Aricota); por lo

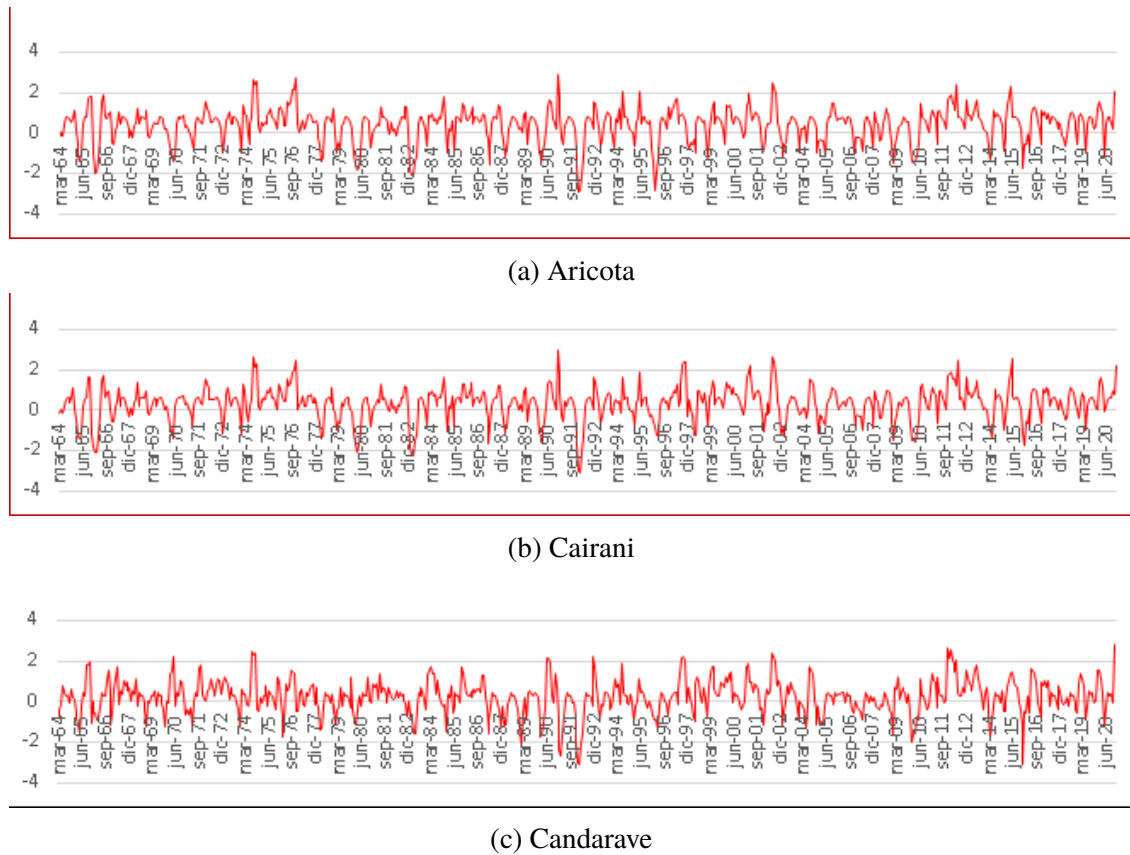


Figura 3: Resultados del índice SPI3 calculado para las estaciones usadas. Índice SPI3 vs. meses.

Tabla 4: Características promedio de las sequías evaluadas con el SPI3

Características	Aricota	Cairani	Candarave
Numero de eventos	25	24	31
Duración promedio (meses)	3,36	3,33	3,10
Magnitud	-3,74	-3,95	-3,11
Intensidad	-1,09	-1,13	-0,99
Pico	-1,48	-1,50	-1,53
Período entre eventos (meses)	28	28	22

que se descartó el uso de los índices ICEN, ONI y TNI. Para determinar la bondad de ajuste de las correlaciones se calcularon los coeficientes de correlación entre los índices restantes y el índice SPI3, según se muestra en la Tabla 5; los valores calculados muestran que en las estaciones Aricota y Cairani existe una correlación muy baja pero significativa (con $p - valor < 0,05$) entre el índice SPI3 y los índices ENSO34 y EMI; por tanto, se descartó el índice ENSO1+2.

En la estación Candarave, se encontró una correlación muy baja pero significativa solamente entre los índices SPI3 y EMI, esto implica que la influencia de las anomalías en la temperatura del mar de la zona 3-4 se pierde a una altitud de 3400 m s.n.m. aproximadamente; sin embargo, teniendo en cuenta que el $p - valor$ es 0,06 (cercano a la significancia de 0,05), se tomará en cuenta para la construcción del modelo.

3.2. Aplicación de cópulas

Debido a que las variables que se analizan no se ajustan a distribuciones de probabilidad conocidas, se decidió utilizar la probabilidad empírica y se calcularon las pseudo-observaciones para cada variable. Las familias de cópulas elípticas y Arquímedeanas, descritas en la Tabla 3, se ajustaron utilizando el método de máxima pseudo-verosimilitud (MPL) y la performance de las diferentes familias de cópulas se evaluó mediante la prueba de Cramér-von Mises (S_n). En la Tabla 6 se presenta el resumen de los estadísticos y

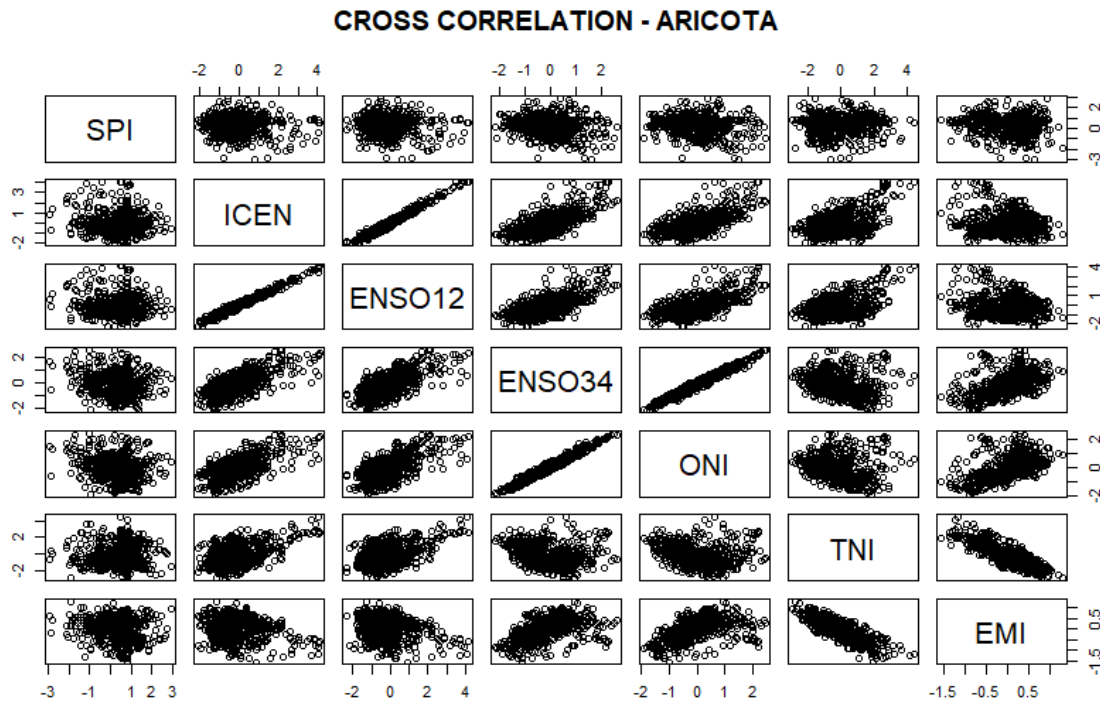


Figura 4: Análisis de correlaciones entre variables en estudio. Estación Aricota. Cada índice se encuentra graficado en el eje vertical de la fila correspondiente y en el eje horizontal de la columna correspondiente, como señalan los círculos rojos para el caso del SPI

parámetros estimados para cada familia de cópulas. En todas las estaciones, la cópula t (Student) mostró un menor valor del estadístico S_n y, por tanto, un mejor ajuste a los datos; asimismo, los valores p muestran significancia al 5%. Por otro lado, se observa que las cópulas elípticas, normal y t (Student), son las que arrojan menor valor en el estadístico S_n , lo que indica que este tipo de cópulas tiene mejor performance para la evaluación de este tipo de variables climáticas.

Finalmente, se procedió a generar una serie aleatoria de 10.000 datos, cantidad mínima recomendada para mantener estables los valores del estadístico de bondad de ajuste de la cópula [22]. Los resultados de la simulación se muestran en la Figura 5. Estas series sintéticas tienen una longitud mucho mayor que la original, representando un espacio muestral más grande y, por tanto, con mayor probabilidad de encontrar las distribuciones marginales de cada variable.

La generación conjunta de estas series busca preservar la dependencia entre variables e incrementar el número de datos para facilitar el cálculo de las

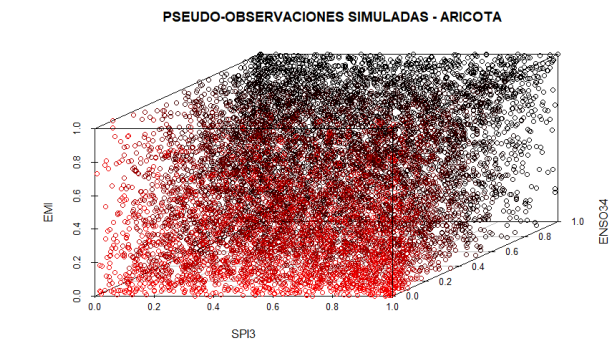


Figura 5: Simulación de 10000 pseudo observaciones para la estación Aricota.

funciones marginales.

El objetivo de generar estas series sintéticas en base a la función cópula es preservar la relación encontrada entre variables, la cual podría perderse si se generan series sintéticas para cada variable de forma independiente, así como definir un espacio muestral mayor que permita estimar valores de probabilidad empírica con mejor aproximación en para los casos en que no existan distribuciones

Tabla 5: Correlaciones calculadas para las estaciones en estudio. Se observa que solamente los índices EMI y ENSO34 tienen significancia al 5 % en todas las estaciones

Estación	Índice ENSO	Tipo de correlación	Coficiente de correlación	Valor p	Significante al 5 %
ARICOTA	EMI	Kendall	-0,0931	0,0003	Si
		Spearman	-0,134	0,0005	Si
		Pearson	-0,1231	0,0013	Si
	ENSO1+2	Kendall	-0,0291	0,2635	No
		Spearman	-0,0433	0,2592	No
		Pearson	-0,0961	0,0121	Si
	ENSO34	Kendall	-0,0861	0,0009	Si
		Spearman	-0,1287	0,0008	Si
		Pearson	-0,178	0,0000	Si
CAIRANI	EMI	Kendall	-0,1237	0,000	Si
		Spearman	-0,1752	0,000	Si
		Pearson	-0,1622	0,000	Si
	ENSO1+2	Kendall	-0,0178	0,493	No
		Spearman	-0,0284	0,4591	No
		Pearson	-0,0554	0,1487	No
	ENSO34	Kendall	-0,0993	0,000	Si
		Spearman	-0,1449	0,000	Si
		Pearson	-0,164	0,000	Si
CANDARAVE	EMI	Kendall	-0,0615	0,017	Si
		Spearman	-0,0907	0,0179	Si
		Pearson	-0,121	0,0015	Si
	ENSO1+2	Kendall	-0,0283	0,2716	No
		Spearman	-0,0424	0,2685	No
		Pearson	-0,0521	0,1745	No
	ENSO34	Kendall	-0,0469	0,0687	No
		Spearman	-0,0708	0,0648	No
		Pearson	-0,0851	0,0263	Si

Tabla 6: Prueba de bondad de ajuste y parámetros de la cópula multivariada. La cópula con mejor ajuste se indica en negrita

Estación	Cópula	Sn	Valor p	Parámetros
ARICOTA	Gumbel	0,65747	0,0004995	1,0596
	Clayton	0,60015	0,0004995	0,141
	Frank	0,60287	0,0004995	0,66602
	Normal	0,14599	0,002498	-0,14221; -0,11902; 0,53578
	t Student(*)	0,12484	0,0004995	-0,1312; -0,142; 0,5658; 8,8248
CAIRANI	Gumbel	0,5676	0,0004995	1,0533
	Clayton	0,5374	0,0004995	0,1168
	Frank	0,54041	0,0004995	0,56528
	Normal	0,11195	0,01049	-0,14750; -0,16195; 0,53580
	t Student(*)	0,093311	0,01548	-0,1476; -0,1846; 0,5689; 7,0776
CANDARAVE	Gumbel	0,62131	0,0004995	1,0771
	Clayton	0,5254	0,0004995	0,17864
	Frank	0,53792	0,0004995	0,79468
	Normal	0,16219	0,0004995	-0,067464; -0,110520; 0,535740
	t Student(*)	0,13395	0,0004995	-0,05478; -0,11335; 0,56548; 7,55966

(*) Los tres primeros parámetros de la cópula t Student representan la matriz de correlaciones; el cuarto parámetro es el número de grados de libertad. El estadístico Sn no está definido para grados de libertad no enteros, por lo que se estimó para el entero más cercano.

marginales que se adecúen a los datos.

En una etapa posterior, estas distribuciones marginales (o las correspondientes probabilidades

empíricas), se utilizarán para la estimación de los períodos de retorno y probabilidad de ocurrencia de diferentes escenarios conjuntos, información

necesaria para la elaboración de planes de mitigación y riesgo de sequías.

4. Conclusiones

En este estudio, se analizó la relación entre la ocurrencia del fenómeno ENSO, definida en base a los índices ENSO1+2, ENSI34, ICEN; TNI; EMI y ONI, y la ocurrencia de sequías agrícolas, definida en base al índice SPI3. Para el efecto, se compararon diferentes modelos probabilísticos basados en cópulas para construir la frecuencia conjunta entre variables. En primer lugar, se correlacionaron entre sí los diferentes índices mencionados para descartar aquellos que no tenían relevancia o aquellos que presentaban una fuerte dependencia respecto a otros, reduciendo el número de índice de interés a tres: ENSO34, EMI y SPI3. El siguiente paso fue evaluar cinco familias de cópulas para probar cuál de ellas presenta mejor ajuste a los datos, encontrándose que la cópula trivariada t de Student tiene mejor performance en todas las estaciones. Los resultados muestran que la cópula de Student es una función robusta y adecuada para el análisis del caso presentado, permitiendo a su vez evaluar la probabilidad de ocurrencia y períodos de retorno de sequías agrícolas en conjunto con la ocurrencia del ENSO, información que es relevante para desarrollar planes de gestión y mitigación de las sequías en la región. Asimismo, se encontró que la influencia de las anomalías de la temperatura del mar en la región ENSO34 se pierde alrededor de la cota 3400 m s.n.m., hecho que contrasta con el incremento en el número de sequías y la mayor severidad registrada por encima de dicha altitud, por lo que se sugiere evaluar la construcción de un modelo probabilístico que incluya la variable “temperatura del aire” para conseguir mayor representatividad en regiones de mayor altitud.

5. Referencias

- [1] Intergovernmental Panel on Climate Change, “Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects.” Cambridge University, Cambridge, United Kingdom and New York, NY, USA, Reporte Técnico, 2014.
- [2] F. Vega, “Variabilidad espacio-temporal de las sequías en el Perú y el peligro asociado al Fenómeno del Niño.” Tesis para optar el grado de Magister Scientiae en recursos hídricos, Universidad Nacional Agraria La Molina, Lima, Perú, 2018.
- [3] C. Knutson, “Methods and Tools for Drought Analysis and Management,” *Eos*, vol. 89, no. 22, 2018.
- [4] G. Salvadori, C. De Michele, N. Kottegoda, and R. Rosso, *Extremes in Nature: An approach using Copulas*. Springer Netherlands, 2007.
- [5] L. Santos, I. Cordery, and I. Iacovides, *Coping with Water Scarcity Addressing the Challenges*. Springer, Dordrecht, 2009.
- [6] SENAMHI, “Análisis de riesgo de sequias en el Sur del Perú.” Ministerio del Ambiente del Perú, Reporte Técnico, 2016.
- [7] INEI, *Perú: Migraciones Internas*. Lima. INEI, 2009.
- [8] A. AghaKouchak, “Entropy-copula in hydrology and climatology,” *Journal of Hydrometeorology*, vol. 15, no. 6, pp. 2176–2189, 2014.
- [9] L. Rüschendorf, *Mathematical Risk Analysis Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer Berlin Heidelberg, 2013, ch. Copulas, Sklar’s Theorem, and Distributional Transform.
- [10] B. Rayens and R. Nelsen, “An Introduction to Copulas by Roger B. Nelsen,” *Technometrics*, vol. 42, no. 3, p. 317, 2000.
- [11] M. Azam, S. Maeng, H. Kim, and A. Murtazaev, “Copula-Based Stochastic Simulation for Regional Drought Risk Assessment in South Korea,” *Water*, vol. 10, no. 4, p. 359, 2018.
- [12] L. Hangshing and P. Dabral, “Multivariate Frequency Analysis of Meteorological Drought Using Copula.” *Water Resources Management*, vol. 32, pp. 1741–1758, 2018.
- [13] F. Serinaldi, B. Bonaccorso, A. Cancelliere, and S. Grimaldi, “Probabilistic characterization of drought properties through copulas,” *Physics and Chemistry of the Earth Parts A/B/C*, vol. 34, no. 10, pp. 596–605, 2009.
- [14] K. Xu, D. Yang, X. Xu, and H. Lei, “Copula based drought frequency analysis considering the spatio-temporal variability in Southwest China,” *Journal of Hydrology*, vol. 527, pp. 630–640, 2015.
- [15] J. Cabrera and J. Mejía, “Relationship between drought occurrence and ENSO in Southern Peru: a copulas analysis,” in *38th IAHR World Congress-“Water: Connecting the World”*, Panama, 2019.
- [16] D. A. Wilhite and M. Glantz, “Understanding the drought phenomenon: The role of definitions,” *Water International*, vol. 10, no. 3, pp. 111–120, 1985.
- [17] T. Mckee, N. Doesken, and J. Kleist, “The relationship of drought frequency and duration to time scales.” in *Eighth Conference on Applied Climatology*, Anaheim, California, 1993, pp. 17–22.
- [18] L. Nkemdirim, *Encyclopedia of Atmospheric Sciences*,

- 2nd ed. Academic Press, 2015, ch. HYDROLOGY, FLOODS AND DROUGHTS | Palmer Drought Severity Index, pp. 224–231.
- [19] A. Cancelliere, G. Di Mauro, B. Bonaccorso, and G. Rossi, “Drought forecasting using the Standardized Precipitation Index,” *Water Resources Management*, vol. 2, no. 5, pp. 801–819, 2007.
- [20] A. Cancelliere and J. Salas, “Drought length properties for periodic-stochastic hydrologic data,” *Water Resources Research*, vol. 40, no. 2, 2004.
- [21] M. Sadegh, E. Ragno, and A. AghaKouchak, “Multivariate Copula Analysis Toolbox (MvCAT): Describing dependence and underlying uncertainty using a Bayesian framework,” *Water Resources Research*, vol. 56, no. 6, pp. 5166–5183, 2017.
- [22] C. Genest, B. Rémillard, and D. Beaudoin, “Goodness-of-fit tests for copulas: a review and a power study,” *Insurance: Mathematics and Economics*, vol. 44, no. 2, pp. 199–213, 2009.

Genesis and geochemistry of the soils of urban landscapes of the Black Sea coast of Russia

Lalita V. Zakharikhina * , Lyudmila S. Malyukova 

Russian Research Institute of Floriculture and Subtropical Crops, Sochi, Russia

Abstract.- In soils of Sochi (the Black Sea coast of Russia), in comparison with natural climatic soil types (burozems, zheltozems (acrisols)), the acid-alkali properties are transformed; in abrozems (technosols), an additional humus content is observed. The pH-water index shifted from 5,8 to 7,5 (average values for the genetic horizons of soils), the degree of saturation with alkali increased 1,3 times in the upper horizon (from 72,9 to 97,7 %) and 1,5 times (from 64,8 to 97,3 %) in the structural-metamorphic BM horizon (argic). Abrazems (technosols) are characterized by a decrease in the humus content relative to the background soils by about 2 times. Assessment of the content in soils of different landscapes of Sochi of a wide range of chemical elements (61) that are excessive and deficient concerning the background has shown the following. Soil contamination is mainly due to its enrichment with elements from mid-soil horizons and underlying parent rocks that can enter the upper soil horizons during excavation works and backfilling of construction sites and roads with local soil. The exception is Ca which enters the soil as a result of urban technogenesis. Pollutants, the contents of which must be controlled when monitoring urban soils, are: Ca, Cu, Ni, Mg, Mn, Cd, K and, additionally, in subordinate urban landscapes: Cs, Ga, Be, Rb, V, Fe, Li, Al, as well as 9 rare earth elements (REE): Dy, Tb, Sm, Ho, Eu, Gd, Sc, Y, Er. The total pollution index for the city's soils is not large; it is a value characteristic of the permissible level of pollution ($Z_c > 16$) on slopes, a moderately hazardous pollution category ($Z_c = 16 - 32$) on a flat surface adjacent to the slope and a dangerous pollution category ($Z_c < 32$) in the coastal zone on the territory of sanatoriums.

Keywords: soil of subtropics of Russia; transformation in urban conditions; geochemical properties of soils.

Génesis y Geoquímica de Suelos de Paisajes Urbanísticos en la Costa Rusa del Mar Negro

Resumen.- Comparados con los suelos naturales zonales (tierras parda y amarilla), los suelos de la ciudad de Sochi situada en la costa Rusa del Mar Negro tienen propiedades ácidas alcalinas transformadas; se ve un contenido adicional de humus en los suelos del tipo abrozemo. El pH del agua cambia de 5,8 a 7,5 (valores medios para los horizontes genéticos de suelos); la saturación con álcalis en los horizontes superior y metamórfico estructural BM (árgico) aumentó por casi treinta (de 72,9 a 97,7 %) y cincuenta (de 64,8 a 97,3 %) por ciento. Los abrazemos tienen un contenido del humus dos veces más bajo en comparación con los suelos de fondo. La evaluación del contenido en suelos de diferentes paisajes de Sochi de una amplia gama de elementos químicos (61) que son excesivos y deficientes en relación con el fondo ha demostrado lo siguiente: La contaminación del suelo se debe principalmente a su enriquecimiento con elementos de los horizontes medios del suelo y rocas madre subyacentes que pueden ingresar a los horizontes superiores durante los trabajos de excavación y el relleno de sitios de construcción y carreteras con suelo local. La excepción es el Ca que entra el suelo por medio del tecnogénesis urbano. En el transcurso del monitoreo de suelos urbanos hay que controlar el contenido de Ca, Cu, Ni, Mg, Mn, Cd, K y también el contenido de Cs, Ga, Be, Rb, V, Fe, Li, Al y de nueve elementos de tierras raras (ETR) (Dy, Tb, Sm, Ho, Eu, Gd, Sc, Y, Er) en paisajes urbanos subordinados. El índice total de contaminación de los suelos en la ciudad no es alto y tiene un valor admisible ($Z_c > 16$) en las pendientes, encaja en una categoría de peligro ligero ($Z_c = 16 - 32$) en la superficie plana adyacente a la pendiente y en una categoría de peligro alto ($Z_c < 32$) en la zona litoral del territorio de los sanatorios.

Palabras clave: suelos subtropicales de Rusia; transformación en entorno urbano; propiedades geoquímicas de suelos.

Received: January 04, 2021.

Accepted: April 01, 2021.

* Correspondence author:
e-mail: zlv63@yandex.ru (L. V. Zakharikhina)

1. Introduction

The increase of urban areas causes a growing interest of international soil science in the genesis of urban soils. The change in the global ecological

potential of soils is due to an increase in the technogenic load on the topsoil. Areas with natural and arable land with actively functioning soil are being reduced. Prediction of changes in the ecological functions of the urban topsoil in order to find new ways to preserve these functions is the most important problem of modern soil science. It is possible to solve it only with a comprehensive study of the transformation of all the properties of urban soils that are formed in different natural and technogenic conditions.

In recent decades, the nomenclature and classification of urban soils, reflecting their transformation processes with increasing urbanization, have been widely discussed [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. The processes of reducing the content of humus, total concentrations of N, and the microbial activity of urban soils have been studied [11]. The geochemical features of the soils of the urban landscapes, the spectra of chemical elements whose contents in soils significantly exceed the background indicators have also been in the focus of research. It has been established that the content of heavy metals positively correlates with clay and silt particles [12], is connected with the texture of the soil, the content of organic matter in it [13] and the time of technogenic load on the soil [14, 15]. For various environmental conditions, differing in the degree and nature of anthropogenic impact, a similar composition of pollutants has been established: Cd, Cr, Ni, Zn, Cu, Pb, As, Fe, V [16, 17, 18, 19, 20, 21, 22, 23]. Concentrations of these elements in soils of different types of land use in cities vary within 2–4 orders of magnitude [24, 25, 26]. The degree of variation in the concentration of elements decreases in the following row: Zn > Cr > Ni > Pb > Cu > As > Cd > Hg (Li, 2018). Quite often, the integrated potential environmental risks (PER) of contaminated soils vary from low [27], elevated [26, 28], significant and very high levels [13, 29, 25]. In a number of cases, geochemical anomalies take place in urban conditions, caused not by technogenesis, but by the natural factors of the composition of regional parent rocks [30, 31] or by location near active volcanoes [32]. The alkalization of soils takes place everywhere due to the intake of metal

salts occurring as a result of urban technogenesis and having, as is known, an alkaline reaction [33, 34, 35].

With extensive data on the transformation of soils in urban areas, an objective assessment of their changes with an increasing technogenic load in specific regional conditions is important.

Soils of Sochi, located on the Black Sea coast of Russia in the subtropical climate zone, have developed on rocks that are very rich in chemical elements (black tar shales, dark blue mudstones, siltstones). In Sochi during the Olympic construction in 2014, additional elements from the fragmented local rocks used for backfilling of roads and construction sites were involved in the migration processes. A study of soils in the zone of such an active impact can help to identify new theoretical aspects of the geochemistry of technogenesis. In addition, international scientific publications lack information on the spread and behavior of chemical elements in the urban soils of the subtropical zone of Russia.

Studying the soil genesis of Sochi, identifying geochemical indicators of their transformation processes in typical urban landscapes is the main research objective.

2. Material and Methods

2.1. Objects

Soils of Sochi were studied in the water divide – slope – coastline system of two ecological and geochemical profiles (Figure 1). The latter are located in the Khostinsky district of the city between the lower reaches of the Sochi and Bzugu rivers, along the banks of the Gnilushka stream flowing between them. All of these streams flow into the Black Sea. The first (conditionally background) profile is located on the territory of the Sochi Arboretum Park, the second is located in the zone of active urban technogenesis adjacent to the park from the southeast. The anthropogenic factor, both of the studied region and the whole territory of the city, is mainly due to the typical urban infrastructure, conveyor lines, thermal power plants, household waste, etc. For Sochi, which has historically developed as a resort city [36],

industrial facilities are not typical. The terrain is low-mountain of erosion-denudation type, near the coastal zone characterized by a wide spread of the processes of hill-creeping, planar erosion and slow movement of the weathering crust. The upper points of the profiles are located at an altitude of approximately 120 m, the lower points - at a height of 10 m above sea level.

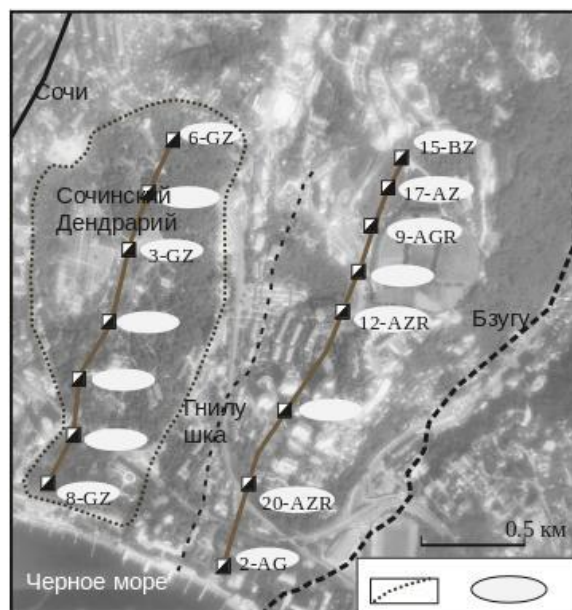


Figure 1: Schematic map of soil sampling in Sochi. 1 - border of the Sochi Arboretum; 2 - sampling points, numbers of sampling points. Latin letters used for soil indices: GZ – yellow soil. BZ – burozem (brown soil). AZR – reclaimed technosol (abrazem). AGR– agro-abrazem. AZ– technosol (abrazem). AG– agrozem

One of the most challenging tasks in characterizing the geochemical properties of soils in urban landscapes is the selection of similar sites not affected by anthropogenic impact for assessing the natural geochemical background (background soil-geochemical parameters), a comparison with which reveals the technogenic geochemical anomalies that are not related to the natural composition of rocks and soils.

By the conditions and type of soil formation, the soils of the Arboretum park are closest to the studied urban soils (transformed zheltozems and burozems (Acrisols)). However, although to a

much lesser extent, this area is also experiencing anthropogenic urban pressure. Therefore, at a considerable distance from the city on the territory of the Caucasian State Reserve upstream the Achipse river (the right tributary of the Mzymta river, 40 km northeast of the lower point of the technogenic profile), close in genesis natural soils (lithosolic burozems (acrisols)) were studied, which, like urban soils, are characterized by zonal soil formation.

The parent rocks selected for the comparison are similar. On the territory of the Arboretum park and in the urban area of active technogenesis, the parent rocks are represented by Oligocene clays with interlayers of siltstones and sandstones; upstream the Achipse river, the parent rocks are mainly terrigenous (clastic) rocks represented by Jurassic mudstones, with interbeds of siltstones and sandstones.

2.2. Methods

Soil diagnostics was carried out in accordance with the Russian Soil Classification, 2008 [37], at the first mention of the name of the soil, its name is given in accordance with the World Reference Base for Soil Resources, WRB. The following indicators were evaluated in the soils: humus - according to Tyurin in the modification of Central Research Institute of Agrochemical Services for Agriculture (GOST 26213-91) with the selection of plant residues during sample preparation by further oxidation of the organic substance with a solution of potassium dichromate in sulfuric acid and the subsequent determination of trivalent chromium, equivalent to the content of organic matter, using a photoelectric colorimeter; water and salt pH (1 mEq/L KCl) extracts (soil-solution ratio = 1:2,5) potentiometrically (GOST 26423-85, GOST 26483-85, respectively); hydrolytic acidity according to the Kappen method in the modification of Central Research Institute of Agrochemical Services for Agriculture titrimetrically (extract of 1 mEq/L CH₃COONa, in the soil: solution ratio of 1:2,5, GOST 26212-91); the content of exchange forms of calcium and magnesium trilonometrically with the extraction of 1 mEq/L NaCl (GOST 26487-85).

The gross contents of 61 chemical elements in soils were determined by quantitative methods - mass spectrometric and atomic emission analyzes with inductively coupled plasma (ICP-MS and ICP-AES) according to the certified method of Scientific Council on Analytical Research Methods No.499-AES/MS “Determination of the elemental composition of rocks, soils, and bottom sediments by inductively coupled plasma atomic emission and inductively coupled plasma mass spectral methods”. Equipment used: inductively coupled plasma mass spectrometer Elan-6100 (Perkin Elmer, USA); inductively coupled plasma atomic emission spectrometer Optima-4300 DV (PerkinElmer, USA).

The procedure for decomposition of samples in an open system using HF, HNO₃, HCl and HClO₄ was applied. The decomposition stage of each sample was controlled using stable isotopes. The correctness of the analysis was confirmed by the use of standard samples. The chemical yield of all analyzed elements using this method of dissolving the samples, as a rule, is 90-100 %.

The tolerance in the determination of the content of chemical elements using these methods and the application of an external standard is ≤ 6 %. Method detection limits for soils are hundredths of micrograms per gram ($\mu\text{g/g}$) for trace elements and hundredths of % for macrocomponents. The decomposition method and subsequent analysis of the obtained solution by ICP-MS + ICP-AES is described in detail in [38].

The content of mobile forms of rare-earth elements (REE) in soils was identified by ICP-MS + ICP-AES methods according to the certified method of Scientific Council on Analytical Research Methods No.500-MS “Determination of the elemental composition of nitric acid and acetate-ammonium extracts from soils by mass spectrometry with inductively coupled plasma”. When extracting mobile REE forms from soils, an ammonium acetate extract (pH = 4,8) was used. The ratio of solid to liquid = 1:10.

2.3. Statistical processing

The average geochemical backgrounds of the elements for the region in general ($C\phi$) in the soils

of the Arboretum and the Caucasus Nature Reserve are calculated as the arithmetic mean values of the element contents at the 7th and 18th observation points, respectively. The calculation method took into account the small range of the gross contents of chemical elements ($C\phi$).

The calculation of the clarkes of the elements concentration (K_k) in soils was made using the following ratio: $K_k = C\phi/K$ for soils, where K is the general content of elements in continental soils (soil clarkes) [39]. The whole range of values $1 > K_K > 1$ is considered in the present study; when interpreting the results, the traditional for geochemistry terminology is used: $K_K > 1$ – excessive elements, $K_K < 1$ – scarce chemical elements.

The degree of similarity of the elemental composition of soils of two background territories was determined by calculating the rank correlation coefficients for concentration clarkes (K_k).

To determine the degree of soil pollution, the total pollution index $Z_c = \sum K_{ci} - (n - 1)$ was calculated, where K_{ci} is the concentration coefficient of the i th element in the soil, equal to the ratio of its actual content at the observation point (C_i) to the local background ($C\phi_i$) [40, 41]. The content of trace elements in terms of MPC (maximum permissible concentrations) (hygienic norms GB 2.1.7.2041-06) and APC (approximate permissible concentrations) (hygienic norms GB 2.1.7. 2511 – 09) was established.

To determine the sources of soil pollution in the subordinate urban landscapes, rank correlation coefficients were calculated using the K_{ci} indices of the average structural and metamorphic horizon BM of the soil of the autonomous landscape (clay horizon (argic)) and the upper horizons of the soils of anthropogenic profile.

The approximate assessment scale of the danger of soil pollution by the Z_c total pollution index has the following categories: acceptable - less than 16, moderately dangerous - 16-32, dangerous - 32-128, and extremely dangerous - more than 128 (Saet, 1990). Permissible gross concentrations (MPC, APC) of the elements in the soil under consideration are: Sb-4,5 (MPC) mg/kg, Mn-1500 (MPC) mg/kg,

V-150 mg/kg (MPC), Pb 32 (MPC) and 130 (APC) mg/kg, As 2,0 (MPC) and 10,0 (APC) mg/kg, Cd-2,0 (APC) mg/kg, Cu-132 mg/kg (APC), Ni-80 mg/kg (APC), Zn-220 mg/kg (APC), (hygienic norms GB 2.1.7.2041-06, GB 2.1.7. 2511-09).

3. Results and discussion

3.1. General characteristics of urban soils

To understand the peculiarities of the formation of the geochemical composition of urban soils, it is important to have the information on the morphological structure of their profiles and physico-chemical properties caused by technogenesis. The features of changes in the soil structure of typical urban landscapes of the studied technogenic profile are clearly shown in the schematic diagram (Figure 2).

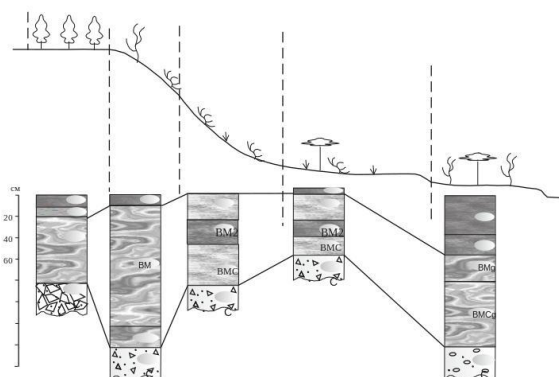


Figure 2: A generalized scheme of changes in the structure of soil profiles of characteristic urban landscapes of Sochi in the system of water divide - slope - coastline. Horizon indices are given in accordance with the Russian Soil Classification, 2008 (Shishov, 2008): AY - gray humus (fulvic); BM - structural-metamorphic (argic); BMg - structurally-metamorphic gleysolic (argic with gleysolic features); PB - agro-abrazem (hortic); W - poorly developed humus (fulvic); PU - agro dark humus (hortic with signs of chernic)

Typical burozems (Acrisols (Clayic)) with a natural profile structure are developed on an autonomous water-dividing surface in the area

of minor technogenic impact (the vicinity of the weather station) in a virtually non-transformed hornbeam-beech forest.

For the upper parts of the slopes of Sochi, local plantations of subtropical cultures (used or abandoned gardens) are quite common. Within the studied profile, the soil of such a landscape on the territory of an abandoned hazelnut garden has been described. The upper soil horizon is formed here as a result of agrogenic transformation of the middle structural-metamorphic BM horizon (clayey (argic) horizon). The middle and lower parts of the profile have not been transformed. The soil can be defined as structural-metamorphic agro-abrazem (Fulvic Anthrosols (Aric, Clayic)).

Down along the terrain, on slopes with a steepness of about 20-2500, most common are soils altered to the state of structural-metamorphic abrazems (Technosols (Clayic)). As a result of excavation work carried out for terrain leveling, humus horizons have been completely lost (cut off). In view of the absence of well-developed grass cover (only the cover of the trimmed lawn) and shrubbery in such areas, the formation of humus horizons is difficult. On the surface of abrazems, lie mineral horizons, composed of displaced and mixed mineral matter, similar in morphological and mesomorphological characteristics to the material of structural-metamorphic BM horizons of the natural zheltzems and burozems.

In local areas of the long-standing transformation of natural soils, thin organic-mineral horizons have been formed on the surface of abrazems. Here, soils can be defined as structurally-metamorphic reclaimed abrazems (FulvicTechnosols (Clayic)).

The subordinate landscapes of coastal terraces are actually occupied by sanatoriums along the entire city coast. The history of the development of this coastal area has been associated with the construction of a sanatorium infrastructure and the creation of arboretums of exotic plants [36]. To cultivate the soils, highly humified material of the organomineral soil horizons, delivered mainly from the chernozem area of Russia (chernozem horizons (chernic)), was applied to the surface of the local soils. The thickness of the newly created

organomineral horizons of the cultivated soils of the sanatorium area of the Sochi agglomeration varies on average from 15 to 20 cm, sometimes reaching 30-35 cm. The soils have the parameters of structural-metamorphic dark agrozems (Chernic Teric Anthrosols (Aric, Clayic)).

The change in the physical and chemical properties of the soils within the technogenic profile was evaluated in comparison with the soils of the Arboretum park (Table 1). A significant transformation of their acid-alkali properties is noted, and for abrozems, additionally, the transformation of the humus parameters is observed. The pH-water index shifted from 5,8 to 7,5 (average values for the genetic horizons of soils), the degree of saturation with alkali increased 1,3 times in the upper horizon (from 72,9 to 97,7 %) and 1,5 times (from 64,8 to 97,3 %) in the structural-metamorphic BM horizon. Abrozems (technosols) are characterized by a decrease in the humus content relative to the background soils by about 2 times. The transformation of the acid-alkali properties of soils is caused by the ingress with surface runoff and drainage water of metal salts occurring in the water due to technogenesis and, as is known, having an alkaline reaction [33, 5, 34].

3.2. Features of the geochemistry of zonal soils outside the area of active technogenesis

The assessment of the features of the elemental composition of zonal soils (Table 2), which do not experience the active technogenic impact of the urban agglomeration, was carried out by calculating the clarkes of concentrations of chemical elements for the soils of the Arboretum park and soils of the Caucasus State Nature Reserve. A comparative analysis of the regional characteristics of the elemental composition of zonal soils with the general spread of elements in the continental soils (clarkes of soils) was performed [39]. For background samples of the abovementioned soils, geochemical formulas were constructed in which the clarkes of the elements concentration (K_k) were ranked in decreasing order (Table 3). In the formulas, above the line are the elements with regional background concentrations which exceed their values in continental soils or

are close to them (hereinafter, excess elements with $K_k \geq 1$). Below the line are the elements which are deficient ($K_k < 1$).

In general, the zonal soils developed in this case on clays and mudstones with interlayers of siltstones and sandstones are characterized by a rich elemental composition in comparison with the average indices for the continental soils. Out of the total number of the studied chemical elements, the values of K_k exceeding or close to 1 are typical both for the soils of the Arboretum and for the soils of the Caucasus Nature Reserve, more than half (52 % of the elements for the soils of the Arboretum park, 68 % - for the nature reserve). The number of elements, whose contents significantly exceed clarkes values of $K_k \geq 1.5$, is 10 for the soils of the Arboretum park and 36 for the soils of the reserve.

A richer composition of the soils of the reserve in comparison with the soils of the Arboretum is due to the genesis of rocks upstream the Achipse river. As indicated above, this place is characterized by the development of terrigenous (clastic) mudstones. Soils are formed here on considerably fragmented (due to the high steepness of the slopes reaching 35-400) mudstones that occur, as shown by the field morphological description of soils, close to the surface. Close contact with crushed rock rich in chemical elements is the cause of a relatively high content of elements in the soils of this area in comparison with soils of the Arboretum park.

On the whole, the rich elemental composition of soils in both territories is determined by the features of the geochemical composition of sedimentary clay rocks on most of the Sochi coast. According to the data on the clarkes of rocks that compose the upper part of the continental crust [42], the content of a large spectrum of chemical elements is increased in clays and shales relative to other sedimentary rocks. Elements whose clarkes are approximately twice the clarkes of other rocks are the following: Li, K, Rb, Cs, Al, Sn, P, As, Cu, Ni, Co, Ag, Zn, Sc, Cd, Th, Pb, Ta, V, Nb, Cr, W and rare-earth elements related (by analogy with [43]) to the groups of light and medium REE (La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy).

A comparison of the geochemical properties of

Table 1: Physical and chemical properties of soils

Horizon	Depth (cm)	pH		Humus (%)	Ca ⁺²	Mg ⁺²	Hidrolytic acidity	Soil saturain with alkali (%)
		water	KCl					
The upper part of the Arboretum park, typical yellow soil (Acrisols (Clayic)), section 6								
AY	0-10	5,6	5,4	7,23	22,3	5,9	6,1	82,1
AYB	10-20	5,8	4,8	5,17	16,1	3,9	5,9	77,2
BM	20-70	6,5	5,8	4,61	19,9	5,3	2,0	92,7
Middle part of the park, gleysolic yellow soil (Stagnic Acrisols (Clayic)), section 3								
AY	0-10	5,6	4,9	7,2	18,3	2,9	5,9	78,2
BM	10-27	6,1	5,2	5,3	16,5	2,7	3,4	84,8
BMg1	27-60	6,2	4,7	4,2	14,6	4,1	3,6	83,7
BMg2	60-70	6,1	4,0	4,0	5,9	3,9	3,9	83,1
Lower part of the park, gleysolic yellow soil (Stagnic Acrisols (Clayic)), section 8								
AY	0-10	5,5	4,7	2,9	5,2	1,8	5,9	54,2
BM	10-30	5,7	4,1	1,9	2,9	1	5,5	41,6
BMg1	30-50	5,8	4,0	1,9	4,0	0,9	9,7	33,5
BMg2	50-70	5,6	3,7	1,8	6,1	3,1	17,6	34,3
Technogenic profile, water divide surface of the Bzugu River and Gnilushka stream, section 15, typical burozem (Acricisols (Clayic))								
AY	0-0,5	7,0	6,9	3,7	20	4,4	0,7	97,9
AYBM	0,5-12	7,2	6,9	2,5	21,4	3,9	0,6	97,6
BM	30-90	7,9	7,3	2,3	24,8	4	0,6	98,1
BMC	90-115	8,1	7,2	2,0	25,1	4,1	0,5	98,9
Upper part of the slope, typical agro-technosol. (Fulvic Anthrosols (Aric, Clayic)), section 9								
PB	0-2	6,7	6,6	5,3	39,8	5,3	1,9	96,1
BM	2-92	7,5	7,3	2,4	27,7	2,6	0,6	98,0
BMg	92-115	7,9	7,3	1,9	24,6	3,2	0,6	98,1
Middle part of the slope, structural-metamorphic technosol (Technosols (Clayic)), section 12								
BM1	0-20	7,6	6,9	2,4	26,2	2,1	0,6	98,7
BM2	20-45	8,2	7,4	2,3	24,9	2,5	0,4	98,0
Coastal terrace, "Avantgarde" sanatorium, dark structural-metamorphic agrozem (Chernic Anthrosols (Aric, Clayic)), section 2								
PU1	0-35	7,07	6,75	6,67	38,96	2,20	0,74	98,23
PU2	35-45	7,70	6,99	4,31	20,81	3,81	0,55	97,81
BMg	45-70	7,63	6,17	4,24	12,90	4,26	0,92	94,91
BMCg	70-120	7,35	6,00	4,06	20,27	4,62	1,20	95,40

the soils of the two studied areas reveals their similarity in the composition of both excess and deficient elements. The coefficient of K_k rank correlation for the soils of the Arboretum park and the soils of the Caucasus Nature Reserve is 0,799. Taking into account the amount of data used, $n = 53$, $r_{5\% \text{ crit}} = 0,271$.

This makes it possible to refer both areas to the same geochemical category, according to the analysis of which the group of excess elements ($K_k \geq 1,5$) should include: Zn, Mn, Cs, Sn, As, Ta, W, Li, Bi, Cu, Ni, Pb, Co, Th, Sc, Ag, Rb, Tl, while the group of excess elements ($K_k \leq 0,5$) should include: Mo, Hf, Zr, Sr, Ca. As can be seen from the above list of excess chemical elements, most are included in the spectrum of elements contained in elevated concentrations in clay and shales, characteristic of the territories

under consideration and for the region as a whole. It is only necessary to note the absence of excess elements in soils of rare-earth elements, as well as the absence of Mn and Bi.

The high content in the soils of the Arboretum park of Zn ($K_k = 5,9$) should also be noted. For the soils of the reserve with relatively elevated contents of most elements, the K_k level of Zn is also high, but it is 2,3 times lower than in the Arboretum. This feature is due to the local geochemical features of the geological foundation of the Arboretum park territory. In accordance with the scheme of mineralogical zoning upwards the terrain of this region, upstream the Sochi and Bzugu rivers, there are several mineralogenic zones and ore areas with different compositions of excess elements. Two geochemical anomalies are characterized by the presence of excess Zn in the

Table 2: The content of chemical elements in the soil of Sochi and in the continental soils (Clarke-K). Na, Mg, Al, K, Ca, Ti, Mn, Fe in %. Other elements in mg/kg.

Element	K	Arboretum background	Reserve background	Eluvial landscape		Transeluvial landscape				Superaquatic landscape				
				Soil Index (point number)								AZ (19)	AZR (20)	AG (2)
				BZ (15), AY horizon	BZ (15), BM horizon	AZ (17)	AGZ (9)	AZ (18)	AZR (12)					
Li	30	40,77	83,60	43,10	57,00	53,60	38,10	51,90	39,90	39,90	46,20	35,20		
Be	1,5	1,22	2,70	1,45	1,79	1,70	1,22	1,72	1,42	1,27	1,53	1,19		
Na	0,63	0,75	0,93	0,66	0,84	0,75	0,50	0,68	0,65	0,53	0,61	0,73		
Mg	0,63	0,37	0,84	0,53	0,67	0,61	0,50	0,68	0,49	0,36	0,78	0,65		
Al	7,13	5,05	8,84	5,40	7,35	6,61	4,85	6,56	4,90	4,97	6,14	5,08		
K	1,36	0,92	1,67	1,19	1,51	1,34	1,05	1,36	1,09	0,81	1,38	1,00		
Ca	1,37	0,42	0,14	1,07	0,34	0,75	1,21	2,93	2,24	0,91	5,36	3,02		
Sc	8	9,90	16,90	11,20	14,40	13,10	10,40	14,10	10,50	10,40	13,60	10,40		
Ti	0,37	0,37	0,40	0,29	0,34	0,32	0,24	0,31	0,25	0,44	0,35	0,34		
V	90	80,57	150,90	90,70	122,90	108,9	76,30	115,8	91,60	92,70	108,8	90,90		
Cr	60	60,40	78,70	55,90	73,30	66,30	50,00	74,90	48,00	60,70	70,90	70,90		
Mn	0,05	0,08	0,29	0,12	0,06	0,06	0,05	0,09	0,07	0,15	0,06	0,12		
Fe	3,8	2,51	4,66	2,69	3,76	3,38	2,48	3,83	3,23	3,29	3,32	3,64		
Co	9	9,56	23,40	11,00	13,20	11,20	7,56	17,20	11,60	19,30	11,40	11,20		
Ni	20	24,13	51,80	37,10	42,50	41,70	28,80	42,50	38,00	24,00	41,50	34,80		
Cu	23	26,10	63,40	33,60	40,00	57,00	43,30	42,30	38,00	26,60	35,30	104,3		
Zn	60	352,5	155,70	99,70	109,30	98,90	126,3	107,1	95,80	73,50	98,90	977,8		
Ga	20	12,08	24,20	14,50	18,90	17,10	12,70	17,60	13,40	12,20	17,00	12,50		
As	6	9,28	23,50	8,08	8,81	8,83	7,46	10,80	11,50	9,73	9,28	11,60		
Rb	70	78,90	133,40	94,20	119,60	109,3	83,10	108,3	86,70	73,90	108,5	77,40		
Sr	220	82,37	75,00	84,50	77,80	95,30	83,40	127,8	118,8	83,80	177,1	175,9		
Y	25	15,73	14,10	17,30	20,00	20,80	16,90	20,60	19,20	14,90	21,60	15,20		
Zr	300	136,8	84,50	123,90	129,60	135,7	101,8	110,9	104,0	140,2	141,9	95,70		
Nb	11	14,37	15,00	11,30	12,20	12,00	9,14	11,10	9,70	16,00	12,90	12,00		
Mo	2	1,03	0,94	0,96	1,08	0,93	0,70	1,16	1,85	1,21	1,22	1,64		
Ag	0,1	0,12	0,19	0,05	0,05	0,05	0,05	0,05	0,10	0,05	0,05	2,65		
Cd	0,16	0,21	0,16	0,44	0,25	0,22	0,29	0,25	0,69	0,15	0,23	0,93		
Sn	1,1	2,62	3,59	2,60	2,69	2,66	2,27	3,45	2,23	2,42	2,66	9,55		
Sb	0,9	0,94	0,99	0,91	0,92	0,96	0,72	1,06	1,19	0,96	0,73	1,61		
Cs	3	4,57	12,80	5,70	7,67	7,07	5,18	7,10	5,21	4,79	7,03	4,03		
Ba	500	372,2	457,00	306,30	303,90	306,9	242,8	334,8	283,1	278,4	310,6	467,0		
La	40	26,47	22,00	24,50	28,20	31,20	22,80	30,30	25,00	26,10	28,30	25,60		
Ce	50	61,03	57,80	58,80	67,60	68,90	52,10	72,60	60,90	70,80	63,40	56,90		
Pr	7	5,97	5,54	5,96	6,54	7,07	5,31	6,94	6,25	6,25	6,44	5,69		
Nd	33	25,47	25,70	26,60	30,20	31,30	22,90	31,30	27,90	26,60	28,90	23,70		
Sm	5	4,68	5,89	5,01	5,89	6,28	4,64	6,19	5,48	4,96	5,65	4,50		
Eu	1	0,96	1,18	1,07	1,15	1,27	1,00	1,26	1,15	0,95	1,18	1,00		
Gd	4	4,02	4,70	4,59	4,90	5,34	4,20	5,29	4,98	4,15	5,04	3,98		
Tb	0,7	0,56	0,63	0,65	0,69	0,76	0,60	0,77	0,71	0,56	0,74	0,55		
Dy	5	3,38	3,44	3,96	4,13	4,60	3,56	4,58	4,10	3,53	4,46	3,20		
Ho	0,6	0,60	0,54	0,63	0,69	0,81	0,63	0,73	0,73	0,60	0,77	0,54		
Er	2	1,88	1,60	2,03	2,10	2,37	1,89	2,16	2,07	1,86	2,37	1,55		
Tm	0,4	0,26	0,21	0,28	0,28	0,32	0,26	0,30	0,28	0,28	0,33	0,21		
Yb	3	1,96	1,55	2,08	2,07	2,40	2,05	2,10	2,00	2,02	2,42	1,47		
Lu	0,35	0,26	0,21	0,28	0,27	0,33	0,26	0,29	0,28	0,29	0,33	0,22		
Hf	7	3,51	2,18	3,00	3,04	3,37	2,68	2,62	2,50	3,53	3,37	2,16		
Ta	1	2,12	2,30	1,24	1,51	1,72	1,35	1,51	1,24	1,74	1,57	1,15		
W	1	2,13	2,11	1,69	1,79	1,81	1,73	1,71	1,33	2,36	2,65	1,86		
Tl	0,3	0,38	0,49	0,39	0,47	0,42	0,36	0,41	0,35	0,44	0,45	0,37		
Pb	20	29,87	43,70	28,30	23,50	32,40	22,40	24,90	22,00	34,40	17,60	149,5		
Bi	0,2	0,32	0,50	0,30	0,34	0,29	0,26	0,30	0,24	0,31	0,30	1,15		
Th	6,5	10,90	11,90	9,97	11,30	10,80	9,22	10,70	9,31	11,60	10,90	7,59		
U	1,5	2,10	1,44	1,84	2,05	2,06	1,47	1,72	1,82	2,59	1,98	1,61		
ΣREE	185	163	162	165	189	197	150	200	172	188	186	155		
Zc				7,45	12,14	12,94	5,31	19,56	13,06	6,66	24,02	46,59		

composition. Under such conditions, the increased territories is quite likely. content of this element in the soils of the adjacent The relatively high concentrations of chemical

Table 3: Clarkes of concentrations of chemical elements for zonal natural soils in comparison with the total element contents for continental soils

Area	Geochemical formula
Arboretum park	Zn(5.9)-Sn(2.4)-W. Ta(2.1)-Th(1.7)-Bi, As(1.6)-Cs, Mn, Pb(1.5)-U, Li(1.4)-Cd, Nb,Tl(1.3)-Sc, Ce, Ni, Ag, Na(1.2)-Cu, Rb(1.1)-Co, Sb, Ti, Cr, Gd, Ho, Eu(1.0) Er, Sm, V, Pr(0.9)-Be, Tb, Nd, Lu(0.8)-Ba, Al, K, Dy, La, Fe, Tm, Yb (0.7)-Y, Ga,; Mg(0.6)-Mo, Hf, Zr(0.5)-Sr(0.4)-Ca(0.3)
Caucasian Reserve	Mn(5.7)-Cs(4.3)-As(3.9)-Sn(3.3)-Li, Cu(2.8)-Co, Zn, Ni(2.6)-Bi(2.5)-Ta(2.3)-Pb(2.2)-Sc, W(2.1)-Rb, Ag(1.9)-Th, Be(1.8)-V(1.7)-Tl(1.6)-Na(1.5)-Nb(1.4)-Mg, Cr(1.3)-Al,K, Fe, Ga, Eu, Sm, Gd, Ce(1.2)-Sb, Ti (1.1)- Cd(1.0)-U(1.0) Ba, Ho, Tb (0.9)-Er, Pr, Nd (0.8)-Dy(0.7)-Lu, Y, La(0.6)-Tm, Yb, Mo(0.5)-Sr, Hf, Zr(0.3)-Ca(0.1)
Regional averaged background	Zn(4.2)-Mn(3.6)-Cs(2.9)-Sn(2.8)-As(2.7)-Ta(2.2)-W,Li,Bi(2.1)-Cu(2.0)-Ni(1.9)-Pb, Co(1.8)-Th,Sc(1.7)-Ag(1.6)-Rb,Tl(1.5)-Nb,Na,Be,V(1.3)-Ce,U,Cr,Cd(1.2)-Gd,Sb,Eu,Sm,(1.1)-Ti,Al,Mg,K,Ho (1.0) Fe,Ga,Er,Tb(0.9)-Ba,Pr,Nd(0.8)-Dy,Lu(0.7)-La,Y,Tm,Yb(0.6)-Mo(0.5)-Hf,Zr,Sr(0.4)-Ca(0.2)

elements in the soils of the two studied areas are confirmed by the assessment of their contents relative to the established MPC and APC (Table 4). The content of some elements in actually natural soils are increased relative to normalized values. Maximum excess values were registered for As (4,6 - 11,8 MPC and 2,3 APC) and for Zn (1,6 MPC). In the soils of the Caucasian reserve, additionally, the excess value for Pb (1,4 MAC) was registered.

3.3. Geochemical features of urban soils

The geochemical features of urban soils studied within the technogenic profile were estimated by calculating the concentration coefficients of chemical elements. The average content of elements in the soils of the Arboretum park was taken as a geochemical background. These soils, as shown above, do not significantly differ in the qualitative composition of the elements from the soils of the Caucasus Nature Reserve, which are not affected by anthropogenic impact.

The total pollution index (Z_c) of soils is expected to increase from an autonomous landscape to a subordinate one (Table 2). In the eluvial landscape in the upper humus horizon of soils, the value of Z_c is 7,5; on a slope, the indicator increases on average 1,5 times; on a flat surface adjacent to the slope, the value of Z_c increases in relation to the autonomous position by 3,2 times; and in the coastal zone (the territory of the sanatorium “Avantgarde”) this value increases by 6,2 times.

The lowest Z_c is typical for agro-abrazems developed on the territory of a hazelnut garden (point 9), where the impact of technogenesis has

been weak. A small spectrum of excess elements and their low K_c values are similar to those for the soil of an autonomous landscape.

The features of chemical pollution of urban soils are clearly reflected in geochemical formulas (Table 5), which show elements with K_c over 1,3, ranked in decreasing order of this indicator.

The maximum concentration coefficients for chemical elements located in the first place of geochemical formulas grow sequentially from 2,6 on the eluvial surface to 2,2–6,7 on the slope and to 12,1–17,1 in the subordinate coastal landscape.

The dominant position in most of the above formulas is occupied by Ca, which once again confirms the conclusion about a significant alkalization of soils as a result of urban technogenesis. Pollutants with elevated K_c in all elemental landscapes should include: Cu, Ni, Mg, Mn, Cd and K. In addition, elevated K_c in subordinate urban landscapes contain: Cs, Ga, Be, Rb, V, Fe, Li, Al, as well as 9 REE: Dy, Tb, Sm, Ho, Eu, Gd, Sc, Y, Er. The content of REEs is due to the geochemical features of the rocks of the territory, which differ, as noted above, by their elevated clarkes. On the one hand, most of the listed chemical elements are typical of almost any technogenic geochemical anomaly [40]. Their source is emissions of vehicles, local heating systems, household waste and runoff. On the other hand, they all determine the geochemical specialization of the clayey rocks of the territory [42] and can enter the soil during excavation works and the backfilling of construction sites and roads with local soil.

Table 4: Gross microelements content in MPC (p) and APC (o)

Territorial levels		V	Mn	Ni	Cu	Zn	As		Cd	Sb	Pb	
		p		o			p	o	p		p	o
Regional background	Arboretum	0,54	0,50	0,30	0,40	1,60	4,64	0,93	0,21	0,21	0,93	0,23
	Reserve	1,01	1,91	0,65	0,96	0,71	11,75	2,35	0,16	0,22	1,37	0,34
Urban soils. Soil indices (point numbers)	BZ (15), A horizon	0,60	0,83	0,46	0,51	0,45	4,04	0,81	0,44	0,20	0,88	0,22
	BZ (15), BM horizon	0,82	0,40	0,53	0,61	0,50	4,41	0,88	0,25	0,20	0,73	0,18
	AZ (17)	0,73	0,40	0,52	0,86	0,45	4,42	0,88	0,22	0,21	1,01	0,25
	AGZ (9)	0,51	0,33	0,36	0,66	0,57	3,73	0,75	0,29	0,16	0,70	0,17
	AZ (18)	0,77	0,57	0,53	0,64	0,49	5,40	1,08	0,25	0,24	0,78	0,19
	AZR (12)	0,61	0,50	0,48	0,58	0,44	5,75	1,15	0,69	0,26	0,69	0,17
	AZ (19)	0,62	0,98	0,30	0,40	0,33	4,87	0,97	0,15	0,21	1,08	0,26
	AZR (20)	0,73	0,39	0,52	0,53	0,45	4,64	0,93	0,23	0,16	0,55	0,14
AG (2)	0,61	0,78	0,44	1,58	4,44	5,80	1,16	0,93	0,36	4,67	1,15	

Table 5: Concentration factors of chemical elements for urban soils exceeding the value of 1,3. Values of Kc are given in parentheses

Elemental landscapes	Soil indices (point numbers)*	Chemical elements and their concentration factors
Eluvial	BZ (15), AY horizon	Ca(2,6)-Cd (2,1)-Mn (1,6)-Mg (1,4)-K (1,3)-Cu (1,3)
	BZ (15), BM horizon	Mg (1,8)-Ni (1,8)-Cs (1,7)-K (1,6)-Ga (1,6)-Cu (1,5)-V (1,5)-Rb (1,5)-Fe, Be, Al (1,5)-Sc (1,4)-Co (1,4)-Y (1,3)-Sm (1,3)
Transeluvial	AZ (17)	Cu (2,2)-Ca (1,8)-Ni (1,7)-Mg(1,6)-Cs, K (1,5)-Ga, Be, Rb, Dy, Tb (1,4), V, Fe, Sm, Ho, Eu, Gd, Sc, Y, Li, Al, Er (1,3) 9REE
	AGZ (9)	Ca (2,9)-Cu (1,7)-Cd (1,4)-Mg (1,4)
	AZ (18)	Ca (6,7)-Mg, Co, Ni, Cu (1,8)-Cs, Sr, Fe, K, Ga (1,5)-V, Sc, Be, Tb, Rb, Dy (1,4), Sm, Eu, Sn, Gd, Y, Al, Al, Li (1,3) 7REE
	AZR (12)	Ca (5,3)-Cd (3,3)-Mo (1,8)-Ni (1,6)-Cu (1,5)-Sr (1,4)-Mg, Fe, Sb, Tb (1,3) 1REE
Super-aquatic	AZ (19)	Ca(12,1)-Mg(2,2)-Sr (2,1)-Ni (1,8)-Cs, K (1,6)-Cu, Ga (1,5)-Fe, Sc, V, Rb, Tl, Y, Li, Be (1,4)-Al, Tb, Dy, Ho, Gd, Co, Sm, Eu (1,3) 8REE
	AZR (20)	Ca(12,8)-Sr, Mg (2,1)-Ni (1,7)-Cs, K (1,5)-Ga, Rb, Sc, Y, Cu, V (1,4), Fe, Tb, Dy, Ho, Er, Be (1,3) 5REE
	AG (2)	Ag(17,1)-Ca (7,2)-Pb (5,0)-Cd (4,4)-Cu (4,0)-Sn, Bi (3,6)-Zn (2,8)-Sr (2,1)-Mg (1,7)-Sb (1,7)-Mo, Mn (1,6)-Fe, Ni (1,4)-Ba, As (1,3)

This is confirmed by the similarity of the range of excess elements of the close-to-rock BM horizon of the soil of eluvial position (point 15, BM horizon) with the composition of chemical elements that accumulate below the relief in structural metamorphic abrazems.

The calculation of the rank correlation coefficients (PKK) by the Kc indicator between the BM horizon and the upper horizons of the soils of the technogenic profile showed the highest values of the PKK (0,78 on average, at $r_{5\% \text{ crit}} = 0,271$) for structural-metamorphic abrazems (Table 6). As these soils evolve into abrazems that are regraded, with thin primitive humus horizons developed on the surface, their geochemical connection with the BM horizon weakens.

The PKK index decreases from 0,78 to 0,52 on the slope and from 0,77 to 0,68 in the subordinate landscape.

In addition to a decrease in the PKK index in regraded abrazems relative to structural and metamorphic abrazems, the following parameters also decrease: the number of excess REEs with Kc above 1,3 (from 9 to 1) (Table 5) and the total gross content of the REE group under discussion (by about 13 % on the slope and, insignificantly, by 1,4 % in the subordinate landscape) (Table 2). It is likely that the decrease in REE contents in abrazems regraded at the initial stage of organic matter accumulation is due to their inclusion in the composition of mobile organomineral complexes. The latter are removed from the

Table 6: Coefficients of Kc rank correlation of the upper horizons of soils with Kc transition horizon BM of the soil of the autonomous landscape

Soils in accordance with the Classification of Soils of Russia, 2008	Soils in accordance with WRB	Soil indices (point numbers)	BZ (15), BM horizon
Typical burozem	Acrisols (Clayic)	BZ (15)	0,685
Structural metamorphic abrazem	Technosols (Clayic)	AZ (17)	0,803
Structural metamorphic agro-abrazem	Fulvic Anthrosols (Aric, Clayic)	AGZ (9)	0,586
Structural metamorphic abrazem	Technosols (Clayic)	AZ (18)	0,756
Structural metamorphic abrazem, reclaimed	Fulvic Technosols (Clayic)	AZR (12)	0,521
Structural metamorphic abrazem	Technosols (Clayic)	AZ (19)	0,774
Structural metamorphic abrazem, reclaimed	Fulvic Technosols (Clayic)	AZR (20)	0,685
Dark agrozem	Chernic Terric Anthrosols (Aric, Clayic)	AG (2)	- 0,026

soil and accumulate at geochemical barriers, as evidenced by a significant (4,3-fold) increase in the percentage of REE mobile forms (relative to gross contents) from the autonomous landscape to the subordinate (Table 7).

Agro-abrazems and agrozem typical for the Sochi agglomeration stand out from the described scheme of geochemical transformation of urban soils. Agro-abrazems were also formed in the material of the BM horizons but experienced an additional factor during the cultivation of subtropical fruit crops.

Dark agrozem, developed in a subordinate landscape on a coastal lowland in the territory of the “Avantgarde” sanatorium, has a negative correlation of geochemical connection with the BM horizon, since its upper humus horizon is composed of genetically alien material for the region, imported, as noted above, from the chernozem strip of Russia (chernic horizons).

The soil of the territory of the “Avantgarde” sanatorium has a different qualitative composition of excess elements and relatively high values of their Kc. Here, the following elements are added to the range of pollutants: Ag, Pb, Zn, Sn, Sb, and As. This is one of the rare areas within the landscape profile where the location of Ca is typical for soils not at the beginning of the geochemical series, but after silver. The analysis of the geological conditions of the studied area makes it possible to regard the presence of Ag, As, and Sb at the beginning of a series of priority elements as a possible natural geochemical anomaly. As noted above, in the upper reaches of the Sochi and Bzugu

rivers mineralogical zones and ore regions are quite widespread. Among them are two zones of gold sulfide mineralization, for which Ag and As and, in some cases, Sb are always the main complementary elements [40].

Analysis of the qualitative composition of excess elements in the soils of different elementary landscapes of the urban area makes it possible to identify two main factors in the formation of geochemical features of the topsoil of the city.

The first factor is natural-technogenic, due to the high content of chemical elements in the rocks of the subtropics of the Black Sea coast, which is reflected in the composition of natural soils and their urban derivatives. For urban soils, an additional source of elements is the excavated deep BM horizons and fragmented local rocks used for the backfilling of roads and construction sites.

The second factor is technogenic, which manifests itself mainly in the coastal sanatorium zone of the city. The relatively high content of elements in the soils of sanatoriums is due to:

1. the subordinate position in the terrain of this territory,
2. the high content of chernozem humus in the soils, which has a high absorption capacity and
3. a long period of absence of excavation work related to the construction of new facilities.

The sanatorium zone of Sochi was built quite a long time ago (30–40 years of the last century) [36, 44]. After that, the work accompanied by

Table 7: Percentage of the content of mobile forms of REE relative to their gross concentrations

Elements	Soil indices (point numbers)							
	Bz (15)	Az (17)	Agz (9)	Az (18)	Azr (12)	Az (19)	Azr (20)	Ag (2)
Sc	0,03	0,02	0,02	0,17	0,06	0,55	0,35	0,02
Y	0,45	0,41	0,19	1,31	0,99	1,36	1,76	0,30
La	0,16	0,16	0,11	0,63	0,52	0,75	1,06	0,16
Ce	0,14	0,10	0,07	0,50	0,41	0,60	0,91	0,11
Pr	0,20	0,21	0,13	0,75	0,56	0,71	1,10	0,17
Nd	0,23	0,26	0,15	0,77	0,61	0,71	1,04	0,20
Sm	0,32	0,35	0,19	0,87	0,73	0,76	1,13	0,24
Eu	0,46	0,46	0,26	1,19	0,87	1,07	1,53	0,33
Gd	0,41	0,43	0,24	1,10	0,90	0,97	1,41	0,30
Tb	0,38	0,39	0,20	1,03	0,80	0,92	1,27	0,29
Dy	0,30	0,30	0,15	0,87	0,66	0,89	1,14	0,22
Ho	0,35	0,28	0,15	0,97	0,66	0,88	1,18	0,24
Er	0,22	0,19	0,10	0,74	0,48	0,73	0,97	0,19
Tm	0,16	0,15	0,08	0,60	0,43	0,69	0,88	0,17
Yb	0,12	0,10	0,05	0,48	0,32	0,50	0,70	0,12
Lu	0,12	0,11	0,06	0,48	0,33	0,58	0,73	0,13
Average value (AV)	0,25	0,25	0,14	0,78	0,58	0,79	1,07	0,20
AVi/AV(15)		1,00	0,56	3,12	2,32	3,16	4,28	0,80

the replacement, displacement, and mixing of the upper soil horizons was rare and limited in terms of area size. At the same time, in the urban area, outside the coastal sanatorium zone, in recent years, the upper soil horizons have been quite actively displaced and renovated during construction work in preparation for the Olympic games.

On-site studies of the ecological and geochemical state of soils in sanatorium territories have revealed separate areas (mainly in the floodplains of small streams flowing into the sea) that refer to an extremely dangerous level of soil contamination (Z_c above 128) [45].

4. Conclusion

In general, the natural zonal soils of the Black Sea Sochi area developed on clay shales and mudstone contain a wider range of chemical elements in comparison to their content in continental soils. The group of excess elements includes: Zn, Mn, Cs, Sn, As, Ta, W, Li, Bi, Cu, Ni, Pb, Co, Th, Sc, Ag, Rb, Tl. Most of them are included in the spectrum of elements contained in elevated concentrations (in comparison with other elements) in clays and shales, characteristic of the territories under consideration and for the region as a whole.

In the process of urban anthropogenesis, zonal soils have been transformed in urban conditions to the state of agro-abrozems, abrozems, degraded abrozems and agrozems. They differ in the structure of the soil profile, morphology, and humus content, but, at the same time, they are characterized by similar acid-alkali properties.

The highest concentration coefficients for urban soils are typical for Ca, which is due to their alkalization as a result of the penetration of metal salts associated with technogenesis, which are known to have an alkaline reaction, into soils with surface runoff and drainage water. Other pollutants with elevated K_c in all elemental landscapes include: Cu, Ni, Mg, Mn, Cd, and K and, additionally, in subordinate urban landscapes: Cs, Ga, Be, Rb, V, Fe, Li, Al, as well as 9 rare earth elements (REE): Dy, Tb, Sm, Ho, Eu, Gd, Sc, Y, Er.

Relatively low urban soil pollution has been established, for which Z_c on the slopes in most cases is within the permissible level of pollution ($Z_c > 16$), and in the abrozems of the subordinate landscapes can be referred to a moderately hazardous pollution category ($Z_c = 16 - 32$). The soils of the sanatorium coastal zone are most polluted (Z_c of over 32 referring to the dangerous category of pollution), requiring







detailed monitoring to identify areas with a high level of pollution requiring reclamation.

5. References

- [1] B. F. Aparin and E. Y. Sukhacheva, "Principles of soil mapping of a megalopolis with St.Petersburg as an example," *Eurasian Soil Science*, vol. 47, no. 7, pp. 650–661, 2014.
- [2] W. Burghardt, "Soil in urban and industrial environments," *Journal of Plant Nutrition and Soil Science*, vol. 157, no. 3, pp. 205–214, 1994.
- [3] University of Essen, *First International Conference on soils of urban, industrial, traffic and mining areas*, W. Burghardt and C. Dornauf, Eds. Universitat-GH Essen, 2000.
- [4] A. Lehmann and K. Stahr, "Nature and Significance of Anthropogenic Urban Soils," *Journal of Soils and Sediments*, vol. 7, no. 4, pp. 242–260, 2007.
- [5] M. A. Naeth, H. A. Archibald, C. L. Nemirsky, L. A. Leskiw, J. A. Brierley, M. D. Bock, A. J. Vanden Bygaart, and D. S. Chanasyk, "Proposed classification for human modified soils in Canada: Anthroposolicorder," *Canadian Journal of Soil Science*, vol. 92, no. 1, pp. 7–18, 2012.
- [6] T. V. Prokof'eva, M. I. Gerasimova, O. S. Bezuglova, S. N. Gorbov, K. A. Bakhmatova, N. N. Matinyan, A. A. Gol'eva, E. A. Zharikova, E. N. Nakvasina, and N. E. Sivtseva, "Inclusion of soils and soil-like bodies of urban territories into the russian soil classification system," *Eurasian Soil Science*, vol. 47, no. 10, pp. 959–967, 2014.
- [7] D. G. Rossiter, "Classification of Urban and Industrial Soils in the World Reference Base for Soil Resources," *Journal of Soils Sediments*, vol. 7, no. 2, pp. 96–100, 2007.
- [8] M. N. Stroganova, M. I. Gerasimova, and T. V. Prokofieva, "Approaches to grouping the technogenic soils," *Eurasian Soil Science*, vol. 38, no. 1, pp. 66–71, 2005.
- [9] M. N. Stroganova and A. V. Rappoport, "Specific features of anthropogenic soils in botanical gardens of metropolises in the southern taiga subzone," *Eurasian Soil Science*, vol. 38, no. 9, pp. 966–972, 2005.
- [10] E. Sukhachev and B. Aparin, "Principles of soil mapping of urban areas," in *Abstract book of 9th International Soil Science Congress on "The soul of soil and Civilization"*, Turkey, Antalya, Side, 2014.
- [11] Z.-Q. Wei, S.-H. Wu, S.-L. Zhou, J.-T. Li, and Q.-G. Zhao, "Soil Organic Carbon Transformation and Related Properties in Urban Soil Under Impervious Surfaces," *Pedosphere*, vol. 24, no. 1, pp. 56–64, 2014.
- [12] M. A. Bashir, A. Rehim, J. Liu, M. Imran, and S. Naveed, "Soil survey techniques determine nutrient status in soil profile and metal retention by calcium carbonate," *CATENA*, vol. 173, no. 141–149, 2019.
- [13] K.-L. Hu, F.-R. Zhang, H. Li, F. Huang, and B.-G. Li, "Spatial Patterns of Soil Heavy Metals in Urban-Rural Transition Zone of Beijing," *Pedosphere*, vol. 16, no. 6, pp. 690–698, 2006.
- [14] K. Bu, D. Freile, V. James, and C. J. Richards, "Virinder Sidhu and Nurdan S. Duzgoren-Aydin. Geochemical Characteristics of Soils on Ellis Island. New York-New Jersey. Sixty Years after the Abandonment of the Hospital Complex," *Geoscience*, vol. 8, no. 1, p. 13, 2018.
- [15] E.-Q. Hou, H.-M. Xiang, J.-L. Li, J. Li, and D.-Z. Wen, "Soil Acidification and Heavy Metals in Urban Parks as Affected by Reconstruction Intensity in a Humid Subtropical Environment," *Pedosphere*, vol. 25, no. 1, pp. 82–92, 2015.
- [16] F. Fordyce, S. Nice, T. Lister, B. O. Dochartaigh, R. Cooper, M. Allen, M. Ingham, C. Gowing, B. Vickers, and A. Scheib, "Urban soil geochemistry of Glasgow," British Geological Survey, OpenReport OR/08/002, 2012.
- [17] S. Islam, K. Ahmed, and M. H. Al-Mamun, "Distribution of trace elements in different soils and risk assessment: a case study for the urbanized area in Bangladesh," *Journal of Geochemical Exploration*, vol. 158, pp. 212–222, 2015.
- [18] M. S. H. Khorshid and S. Thiele-Bruhn, "Contamination status and assessment of urban and non-urban soils in the region of Sulaimani City," *Environmental Earth Sciences*, vol. 75, p. 1171, 2016.
- [19] X. S. Luo, S. Yu, Y. G. Zhu, and X. D. Li, "Trace metal contamination in urban soils of China," *Sciences of the Total Environment*, vol. 421–422, pp. 17–30, 2012.
- [20] M. Vinha, M. Cabral, P. A. Dinis, and L. Mandavela, "Geochemistry of Urban Soil in the Fast-Growing Kuito City (Angola)," in *Conference of the Arabian Journal of Geosciences: Petrogenesis and Exploration of the Earth's Interior*, 2018.
- [21] P. Tume, E. González, F. Reyes, J. P. Fuentes, and G. Medina, "Sources analysis and health risk assessment of trace elements in urban soils of Hualpen," *CATENA*, vol. 175, pp. 304–316, 2019.
- [22] X.-S. Wang and Y. Qin, "Some characteristics of the distribution of heavy metals in urban topsoil of Xuzhou, China," *Environmental Geochemistry and Health*, vol. 29, no. 1, pp. 11–19, 2007.
- [23] K. Zhao, L. Zhang, J. Dong, J. Wu, and W. Fu, "Risk assessment spatial patterns and source apportionment of soil heavy metals in a typical Chinese hickory plantation region of southeastern China," *Geoderma*, vol. 360, no. 15, p. 114011, 2020.
- [24] N. Adimalla, "Heavy metals contamination in urban surface soils of Medak province, India, and its risk assessment and spatial distribution," *Environmental Geochemistry and Health*, vol. 42, no. 21, pp. 59–75, 2020.
- [25] M. S. Islam, M. K. Ahmed, M. H. Al-Mamun, and

- S. M. A. Islam, "Sources and Ecological Risks of Heavy Metals in Soils Under Different Land Uses in Bangladesh," *Pedosphere*, vol. 19, no. 5, pp. 665–675, 2019.
- [26] P. Yang, J. Patrick, D. M. Yang, and H. Li, "Spatial variability of heavy metal ecological risk in urban soils from Linfen," *CATENA*, vol. 190, p. 104554, 2020.
- [27] M. Rachwal, K. Kardel, T. Magiera, and O. Bens, "Application of magnetic susceptibility in assessment of heavy metal contamination of Saxonian soil (Germany) caused by industrial dust deposition," *Geoderma*, vol. 295, no. 1, pp. 10–21, 2017.
- [28] H. Zhuo, H. Liu, S. Fu, H. Song, and L. Ren, "Source analysis and risk assessment of heavy metals in development zones: a case study in Rizhao, China," *Environmental Geochemistry and Health*, no. 42, pp. 135–146, 2020.
- [29] M. S. Islam, M. K. Ahmed, M. H. Al-Mamun, and D. W. Eaton, "Human and ecological risks of metals in soils under different land-use types in an urban environment of Bangladesh," *Pedosphere*, vol. 30, no. 2, pp. 201–213, 2020.
- [30] S. Fernández, T. Cotos-Yáñez, J. Roca-Pardiñas, and C. Ordóñez, "Geographically Weighted Principal Components Analysis to assess diffuse pollution sources of soil heavy metal: Application to rough mountain areas in Northwest Spain," *Geoderma*, vol. 311, pp. 120–129, 2018.
- [31] M. Gaberšek and M. Gosar, "Geochemistry of urban soil in the industrial town of Maribor, Slovenia," *Journal of Geochemical Exploration*, vol. 187, pp. 141–154, 2018.
- [32] A. Petrik, M. Thiombane, S. Albanese, A. Lima, and B. de Vivo, "Source patterns of Zn, Pb, Cr and Ni potentially toxic elements (PTEs) through a compositional discrimination analysis: A case study on the Campanian topsoil data," *Geoderma*, vol. 331, pp. 87–99, 2018.
- [33] P. J. Craul, *Urban soils: Applications and practices*. New York: John Wiley and Sons, 1999.
- [34] M. N. Stroganova, "Urban Soils: genesis, taxonomy and environmental significance," Dissertation for the degree of Doctor of Biological Sciences, Higher Attestation Commission of the Russian Federation, Moscow, 1998.
- [35] L. Yang, Y. Li, K. Peng, and S. Wu, "Nutrients and heavy metals in urban soils under different green space types in Anji," *CATENA*, vol. 115, pp. 39–46, 2014.
- [36] A. L. Korkeshko, "The history of park construction on the territory of Sochi (1866-1969) The history of park construction on the territory of Sochi (1866-1969)," in *Reports of the Sochi Department of the Geographical Society of the USSR*. Leningrad: Glavpoligrafprom, 1971, vol. II, [in Russian].
- [37] L. L. Shishov, V. D. Tonkonogov, I. I. Lebedeva, and M. I. Gerasimova, *Classification of Russian soils*. Moscow: Institute of soil studies named after V.V. Dokuchaev of the Russian Academy of Agricultural Sciences, 2008.
- [38] V. K. Karandashev, A. N. Turanov, T. A. Orlova, A. E. Lezhnev, S. V. Nosenko, N. I. Zolotareva, and I. R. Moskvitina, "Use of the inductively coupled plasma mass spectrometry for element analysis of environmental objects," *Inorganic Materials*, vol. 44, no. 14, pp. 1491–1500, 2008.
- [39] A. A. Yaroshevsky, *Problems of modern geochemistry*. Novosibirsk: Novosibirsk State University, 2004.
- [40] A. I. Achkasov, I. L. Basharkevich, T. L. Onishchenko, L. N. Pavlova, B. A. Revich, Y. E. Saet, S. S. Sargsyan, R. S. Smirnova, N. Y. Trefilova, and E. P. Yanin, *Geochemistry of the environment*. Moscow: Nedra Publishers, 1990, [in Russian].
- [41] A. P. Solovov, *Geochemical methods of prospecting for mineral deposits*. Moscow: Nedra Publishers, 1985.
- [42] N. A. Grigor'ev, "Average concentrations of chemical elements in rocks of the upper continental crust," *Geochemistry International*, vol. 41, no. 7, pp. 711–718, 2003.
- [43] D. Aubert, P. Stille, A. Probst, F. Gauthier-Lafaye, L. Pourcelot, and M. Del Nero, "Characterization and migration of atmospheric REE in soils and surface waters," *Geochimica et Cosmochimica Acta*, vol. 66, no. 19, pp. 3339–3350, 2002.
- [44] L. I. Prasolov, I. N. Antipov-Karataev, and V. N. Filippova, *Soils of the Sochi station*. Leningrad: Publishing house of the All-Union Scientific Research Institute of Fertilizers and Agricultural Soil Studies named after D.N. Pryanishnikov, Leningrad branch, 1934.
- [45] L. V. Zakharikhina and A. V. Burtovoy, "Anthropogenic Evolution of the zheltozems (acrisols) of the Sanatorium Zone of Sochi," *Soil Science*, vol. 53, no. 6, pp. 820–828, 2020.

Technogenic geochemical abnormalities in the soils and bottom sediments of Voronezh

Natalia V. Kaverina ^a , Semen A. Kurolap ^a , Pavel M. Vinogradov ^a ,
Tatyana V. Dubovitskaya ^b , Natalya N. Nazarenko ^c , Roman A. Kondurov ^a 

^aDepartment of Geoecology and Environmental Monitoring, Voronezh State University, Voronezh, Russia.

^bDepartment of Physics, Voronezh State University, Voronezh, Russia.

^cDepartment of Biology and Plant Protection, Voronezh State Agrarian University named after Emperor Peter the Great, Voronezh, Russia.

Abstract.- Environmental pollution in large industrialized cities has become an urgent global problem. On the territory of the Russian Federation, a tense ecological situation is developing in almost all cities with a population of over one million and in large administrative centers. The creation of a quality urban environment is a national priority, a complex and financially costly task, the solution of which cannot be postponed any longer. The main sources of problems in the territories occupied by cities lie in the concentration of production capacities, the steady growth of the population, and as a result, the amount of transport. Due to production, industrial cities concentrate substances of natural and man-made origin forming local and even regional geochemical anomalies in various landscapes. Among all the interrelated components, a special place belongs to the depositing (accumulating) media—soils and bottom sediments. Based on the results of geochemical studies of 2018–2020, the indicators of the quality of the soils of Voronezh were determined, the levels of the content of pollutants of I-III hazard classes in the soils of urban and peripheral areas of Voronezh (benzo(a)pyrene and heavy metals) were established. The authors provide an assessment of the accumulation of pollutants in the bottom sediments of the Voronezh water reservoir, which is located on the territory of the urban district.

Keywords: urban environment; degradation; load; functional areas; pollutants; concentration and accumulation factors; total pollution index.

Anomalías Tecnogénicas Geoquímicas en Suelos y Sedimentos de Fondo en Voronezh

Resumen.- La contaminación del medio ambiente de ciudades grandes industrializadas es un problema global de actualidad. En Rusia se ve una situación difícil medioambiental casi en todas las ciudades con más de un millón de habitantes y en centros administrativos mayores. La creación de un entorno urbano de alta calidad es una tarea nacional compleja prioritaria que requiere costes materiales intensivos y la solución más urgente posible. Los orígenes de los problemas de territorios urbanizados se hallan en la concentración de capacidades de producción, un crecimiento constante de población y, con ello, la cantidad de vehículos de transporte en uso. Las sustancias de origen natural y tecnogénico concentradas por ciudades industrializadas a través de capacidades productivas forman anomalías geoquímicas locales y hasta regionales en varios paisajes. Una posición distinta entre todos los componentes interrelacionados pertenece a suelos depositantes (acumuladores) y sedimentos del fondo. Los resultados de las pruebas geoquímicas conducidas en 2018-2020 se usan para medir los índices de calidad de los suelos de la localidad. También se determinan los contenidos en el suelo de contaminantes de los clases de peligro I a III (benzopirenos y metales pesados). Se evalúa la acumulación de sustancias contaminantes en los depósitos de fondo del la represa de Voronezh que se halla por completo en el territorio del distrito urbano.

Palabras clave: medio urbano; degradación; carga; zonas funcionales; contaminantes; coeficientes de concentración y acumulación; indicador acumulativo de contaminación.

Received: January 25, 2021.

Accepted: April 01, 2021.

* Correspondence author:
e-mail:knataliy@mail.ru (N. Kaverina)

1. Introduction

The problems of the landscape-geochemical organization of the territory of the city of Voronezh lie in its complex historical past. During the Great Patriotic War, fierce battles took place in the city. From July 1942 to January 1943, the right bank of the Voronezh river was occupied and the buildings and infrastructure were destroyed in the course of intense hostilities. By the time of liberation, no more than five percent of the buildings remained intact in this part of the city.

The echo of the war sounds in Voronezh until now. Post-war cinder block buildings make up a significant part of the city's dilapidated housing stock. The urban landscape of the right bank is actually formed on the ruins and remains of buildings. As a result of the functional transformation of industrial areas, the foundations and parts of residential buildings, plants, and factories that existed in the pre-war period are exposed.

Landscaping of the territory of Voronezh and its suburbs with rapidly growing balsam poplars (*Populus balsamifera*) is also a result of the post-war period. Replacing green spaces with new sustainable and low maintenance plants is a long-term process that requires carefully considered decisions and investments [1]. The growth of the urban area has led to the occupying of large areas of adjacent forestry. Thus, the following recreational areas appeared in the city: Severny Lesopark, Figurnaya Roscha, Malaya Roshcha, Treugolnaya Roshcha, Pridonskaya Roscha, and Peschany Log.

In 1967, the "rehabilitation of the Voronezh River" was carried out within the city limits and the Voronezh water reservoir was created. The filling of the reservoir coincided with several dry years. The floodplains of the river by this time were very swampy, and the channel needed to be cleaned not only from the polluting runoff but also from the construction waste left after the restoration of the city.

The result of the improvement of the floodplain of the Voronezh River was the creation of a shallow multipurpose water reservoir. Its aquatic and terrestrial ecosystems turned out to be unstable,

which contributed to the further development of imbalance in the "reservoir-environment" system.

The operation of the reservoir as a transport artery between the cities of Voronezh and Lipetsk has been completely stopped. The fishery is represented only by sportfishing.

Despite the reduction in water consumption by industry and agriculture, the reservoir has become shallow and overgrows with duckweed regularly in the warm season.

A vivid example of the degradation of the reservoir is the Maslovsky Zaton area - the place where the river flows into the Tavrovka river reservoir. The river, surrounded by buildings and asphalt, eventually turned into a temporary watercourse. Lack of runoff led to stagnation and intensive development of blue-green algae in the warm season.

The only right decision was the ecological rehabilitation of the territory with the reclamation of land and the formation of the river bed. The work was completed in November 2019, but it was not enough to completely solve the problems of the reservoir. Residents of the city still complain about the putrid smell from the reservoir.

At present, Voronezh, in terms of the urban environment quality index, ranks only twelfth among the fifteen cities with a population of over one million in Russia, significantly inferior to the leader (Moscow) in terms of the quality of landscaping and the development of the road network. This assessment was given by the Ministry of Construction and Housing and Utilities (Ministry of Construction of Russia) according to the Decree of the President of the Russian Federation of May 7, 2018, No. 204 "On national goals and strategic objectives of the development of the Russian Federation for the period up to 2024" and the national project "Housing and urban environment".

Work on environmental impact assessment in Voronezh is carried out regularly for new construction projects and reconstruction. All relationships between natural and man-made systems are regulated by regulatory documents. A mandatory stage of engineering and environmental surveys is to obtain reliable data on the ecological

and geochemical state of the environment. The condition for the reliability of the materials obtained is information on the sources of pollution, factual data on the regional background values of the most common harmful substances.

Information about stationary sources and the mass of emissions of pollutants is accumulated by the Federal Service for Supervision of Natural Resource Usage (Rosprirodnadzor) and the Federal Service for Supervision of Consumer Rights Protection and Human Welfare (Rospotrebnadzor).

Background concentrations of substances in the air of the city of Voronezh are registered by the Federal Service of the Russian Federation for Hydrometeorology and Environmental Monitoring (Roshydromet).

The Federal Agency for Water Resources collects data for measures for the protection of water bodies and the use of water resources to ensure potable and utility water supply.

The purpose of the study was to control the pollution of soils in urban areas and bottom sediments by discharges, and emissions from various enterprises, transport and the general anthropogenic load.

2. Materials and methods

Voronezh is a complex research object in the modern landscape and functional structure of which anthropogenic (technogenic) components prevail, combined with natural systems. The territory of the city is a set of functional areas formed over a long historical period [2, 3, 4, 5, 6].

Various types of soils of urbanized landscapes have been formed on the territory of the city. The natural soil cover of the left-bank part of the city is represented by soddy forest, sandy, and sandy loam soils. Chernozems prevail on the right bank of the city. The territory of the city is characterized by a significant change in the soil cover and the formation of urban soils [7, 8, 9, 10].

The terrain of the city consists of contrasting elements. The right-bank part is located on a hilly plateau, and the left-bank part is located in a low flat area, which gradually turns into a river terrace.

The choice of objects of observation (functional areas and territories) is determined by the economic-geographical and physical-geographical characteristics of the survey area [5], materials of previous surveys [11, 12], and data from the public cadastral map.

Particular attention was paid to the following sources of pollution:

- industrial and transport enterprises;
- power supply enterprises;
- enterprises for storage and transportation of oil products;
- residential areas;
- production and domestic wastes generation areas;
- recreation areas;
- road transport network.

Soil sampling was carried out per GOST 17.4.3.01-2017 according to a previously planned and adjusted in situ scheme using a uniform grid during route trips. To neutralize the local features of the distribution of pollutants, mixed samples were taken from 5 point samples from test plots of 25 m². The sampling depth of point and mixed samples was 0–20 cm.

Bottom sediment samples were taken per GOST 17.1.5.01-80. The standard applies to bottom sediments of water bodies flowing into the sea and establishes general requirements for sampling bottom sediments of water bodies for analysis of contamination by chemical indicators.

The assessment of the state of bottom sediments of the water reservoir was carried out per regulatory document RD 52.24.609-2013 “Organization and carrying out of observations of the content of pollutants in bottom sediments of water bodies”.

Observations of the quality of the soil of the city were carried out by guidelines MU 2.1.7.730-99 “Hygienic assessment of the quality of soil in populated areas”. The list of the studied indicators includes chemical substances of various hazard classes from the list of indicators of Sanitary

Regulations and Norms SanPiN 2.1.7.1287-03 “Sanitary and epidemiological requirements for soil quality” (Table 1).

The fertility of the soil was assessed by an indirect indicator - the content of ammonium nitrogen. For plants, exchangeable ammonium is available, as well as ammonium of water-soluble salts, contained in small quantities. These two forms are extracted from the soil with a 2 % solution of potassium chloride. Since the ammonium of water-soluble salts is only a small part of nitrogen, GOST 26489-85 is considered a method for determining exchangeable ammonium.

A titrimetric method was used to determine chlorides. Their content was used to determine the degree of salinization of urban soils.

The effect of sulfur compounds in emissions from fossil fuel combustion is indirectly characterized by the content of mobile sulfur, determined by the turbidimetric method.

The extraction of 3,4-benzo(a)pyrene was carried out by liquid extraction, n-hexane is used as the extractant. The extraction was analyzed by high performance liquid chromatography (HPLC) with spectrofluorimetric detection using Fluorat-02-2M device. The determination of benzo(a)pyrene content was carried out in the branch of the Center for laboratory analysis and technical measurements (TsLATI) in the Voronezh region and the testing laboratory center of the Federal State Budgetary Institution “Center for Hygiene and Epidemiology in the Voronezh Region”.

Following the requirements of environmental surveys, the assessment of the acid-base properties of soils is carried out in any work on the study of the state of the environment. To assess soil acidity, two groups of indicators are used: pH of aqueous and salt extracts. Based on the requirements of GOST 17.5.4.01-84, the actual soil acidity (pH of the aqueous extract) was measured.

At present, exchangeable soil acidity is understood as the total number of acidic components that are retained on the surface of colloidal particles mainly by electrostatic forces [13]. Exchangeable acidity is a more harmful form of soil acidity for plants [4]; it was determined in pH units of salt extract (1M KCl) according to GOST 26489-85.

At the same time, the soil samples were examined for the content of technogenic pollution indicators - heavy metals (HM). The preparation of soil extracts for the determination of the gross metal content was carried out according to generally accepted methods, followed by instrumental analysis on the “SPEKTR-4” atomic absorption spectrometer [4].

Acid-soluble forms of heavy metals were determined on the “Spektr-5-4” atomic absorption spectrometer with flame atomization in the accredited laboratory of the branch of the Center for laboratory analysis and technical measurements (TsLATI) in the Voronezh region of the TsLATI in the Central Federal District, as well as by stripping voltammetry on the TA-4 analyzer in the ecological analytical laboratory of the Faculty of Geography, Geo-ecology, and Tourism of Voronezh State University. The error varied from 15 to 30 % in all measurement ranges.

The degree of soil pollution was estimated by the accumulation coefficient (K_o) relative to MPC (APC). For bottom sediments, the concentration factor of the chemical substance (K_c) was used, calculated as the excess of the concentrations relative to the background concentrations [14].

The identification of anomalies was carried out according to the calculations of the total indicators of pollution (Z_c).

The interpretation of the data obtained and their mathematical processing was carried out using computer technology. The selection of contours of soils of various degrees of pollution and mapping was performed using the MapInfoPro program at the Faculty of Geography, Geo-ecology, and Tourism of Voronezh State University.

3. Results and discussions

One of the main indicators of the state of soils is their acidity. The values of the indicator directly depend on the quality and quantity of anti-icing agents used, the presence of construction waste, the composition of dust, and atmospheric precipitation.

In recent years, numerous research results have been obtained on the problems of changes in the

Table 1: Main indicators for assessing the state of soils in the city of Voronezh depending on their functional purpose

No.	Indicator	Research method
1	Ammonium nitrogen, (mg/kg)	GOST 26489-85 Soils. Determination of exchangeable ammonium by the method of Central Research Institute of Agrochemical Services for Agriculture (TsINAO)
2	Chlorides, (mg/kg)	GOST 26425-85 Soils. Methods for the determination of chloride ions in aqueous extract
3	Hydrogen index of aqueous extract, (pH)	GOST 26423-85 Soils. Methods for determination of electrical conductivity, pH and solid residue of aqueous extract.
4	Hydrogen index of salt extract, (pH)	GOST 26483-85 Soils. Preparation of a salt extract and determination of its pH by the method of Central Research Institute of Agrochemical Services for Agriculture (TsINAO)
5	Sulfur compounds, (mg/kg)	GOST 26490-85 Soils. Determination of mobile sulfur by the method of Central Research Institute of Agrochemical Services for Agriculture (TsINAO)
6	Heavy metals, (mg/kg)	Regulatory document RD 52.18.191-89 Methodological guidelines. Methods for measuring the mass fraction of acid-soluble forms of metals (copper, lead, zinc, nickel, cadmium) in soil samples by atomic absorption analysis. Environmental regulations PND F 16.1:2:2.2:3.48-06 Methods for measuring the mass concentration of zinc, cadmium, lead, copper, manganese, arsenic and mercury in soils, greenhouse grounds, spropels, silts, bottom sediments, solid waste by stripping voltametry using TA-type analyzers.
7	Benz(a)pyrene, (mg/kg)	Environmental regulations PND F 16.1:2:2.2:3.39-03 Quantitative chemical analysis of soils. Methods for measuring the mass fraction of benzo(a)pyrene in samples of soils, grounds, solid waste, and bottom sediments by the method of high-performance liquid chromatography using "Lyumakhrom" liquid chromatograph

acidity of the soil cover of Voronezh [15, 4]. It was found that changes in the actual acidity are associated not only with the type of soil but also with the levels of technogenic load on the territory. The high level of anthropogenic impact on the soil of the city has led to significant geochemical transformation and a change in the nature of the functioning of the soil.

The excess of alkalinity in the soils of the city, on average, reached three pH units in comparison with the background territories for the actual acidity level and four pH units for exchangeable acidity (Figures 1, and 2). By the reaction, the soil pH can be classified as weakly acidic, neutral, and slightly alkaline. The values vary over a wide range, which significantly differs across functional areas.

For the developed industrial areas, regardless of location, slightly alkaline or neutral pH values are characteristic. The acidity value is influenced by powerful technogenic factors (Table 2). As a result, the values of the indicator in these areas are the highest ($\text{pH}_{\text{aq}} > 7$).

The obtained results of the values of the hydrogen index coincide with the data of other authors on the increased alkalinity of the soils of Voronezh and the dependence on the degree of atmospheric pollution [15, 6, 9]. The shift in the reaction is due to the impact of technogenic

emissions on the soil, which contains a large number of alkaline components.

The study revealed, for the city of Voronezh, statistically significant correlation coefficients of the pH values and the atmospheric pollution indicator [15]. With approaching the administrative boundaries of the city and weakening of the traffic load, the pH values of the aqueous and salt extracts become neutral and slightly acidic in the forests of recreational functional areas.

The pH value of the salt extract of soils is in direct correlation with the content of mobile sulfur. For functional areas, the following values were calculated: $r = +0,59$ in the recreational area, $r = +0,53$ in the residential and $r = +0,4$ in the industrial area.

In the recreational functional areas, the average degree of linear dependence of the pH of the salt extract on the accumulation of mobile sulfur and exchangeable ammonium in the soil was established (+0,59 and +0,69). When calculating the correlation coefficient between the measured values of the pH of the aqueous extract and mobile sulfur in the soils of industrial and recreational functional areas, a slight degree of linear dependence was revealed (+0,47 and +0,39, respectively). Thus, the concentrations of nitrogen and sulfur compounds identified in the soil within

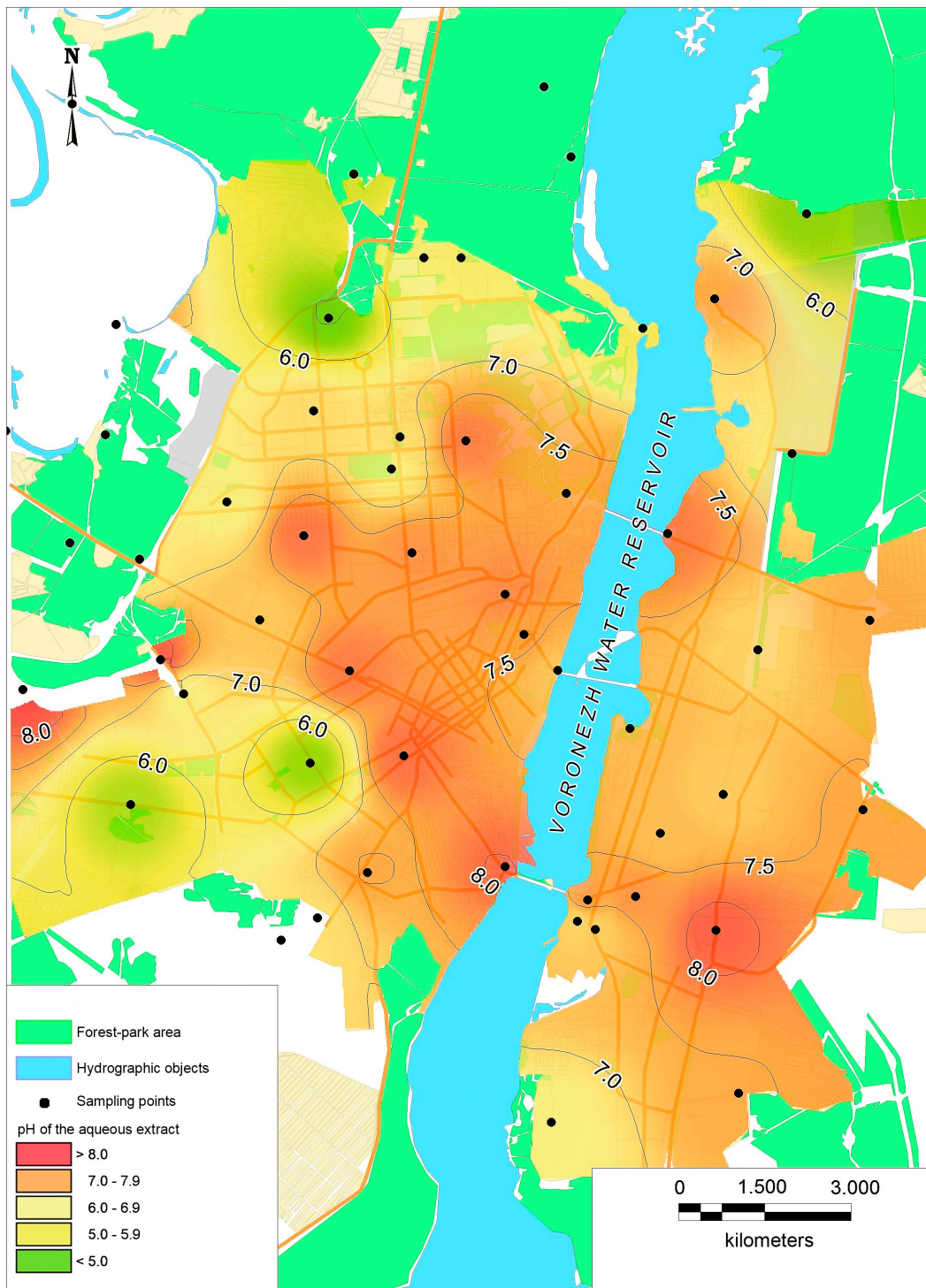


Figure 1: Distribution of the pH values of the aqueous extract in the surface soils of Voronezh

the territory of the city affect the pH value and occur due to transport, industrial, and “atmospheric” sources.

The soil cover of the city of transport functional areas has been formed under special conditions. In the selected areas, neutral indicators of the average

pH values of the aqueous and salt extracts have been established. This greatly contributes to the active accumulation of readily soluble salts, strong fragmentation of the soil surface layer, and the development of degradation processes.

Water-soluble chlorides play a significant role

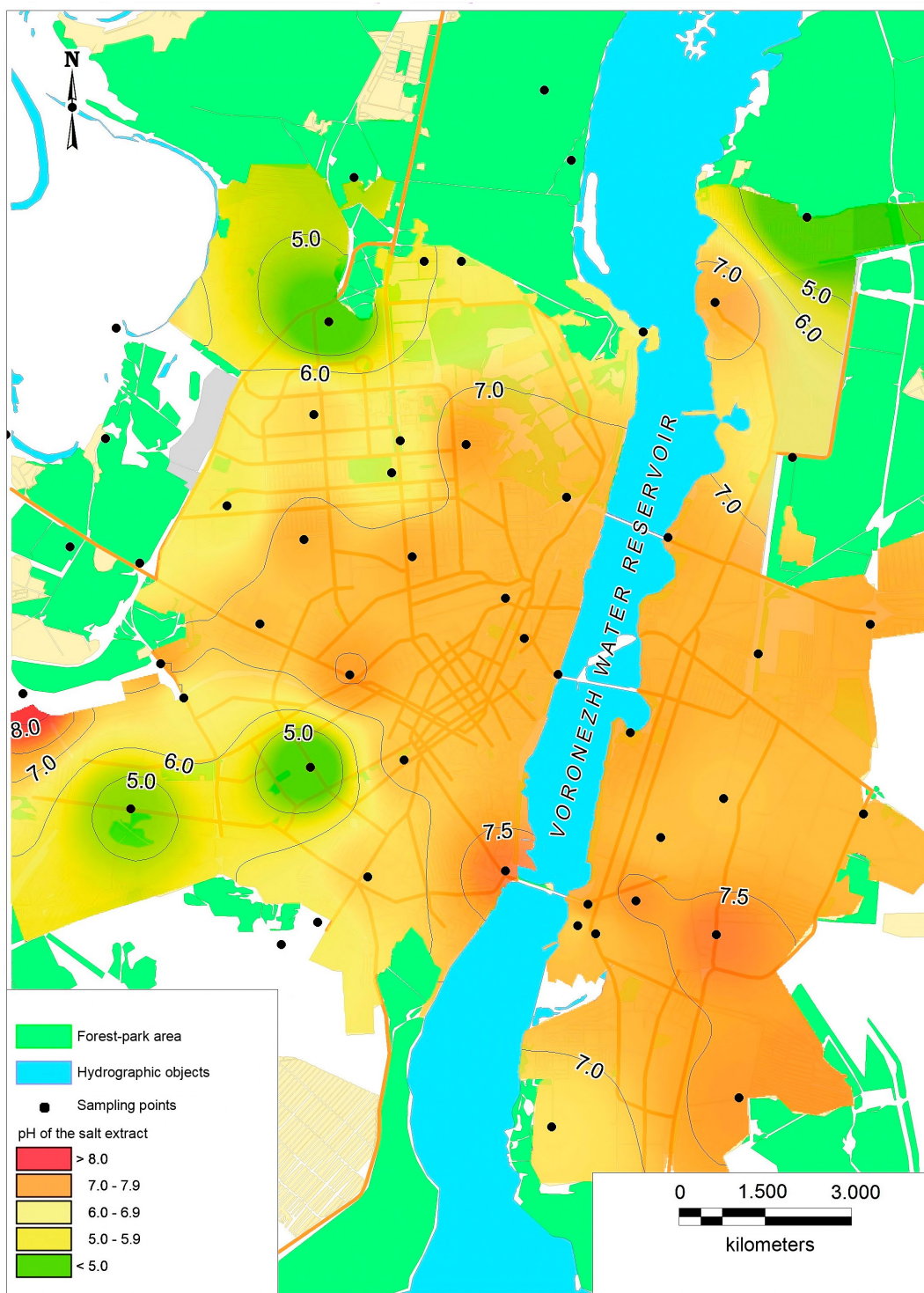


Figure 2: Distribution of the pH values of the salt extract in the surface soils of Voronezh

in technogenic urban flows. As the most mobile compounds, migrating further than other salts, and can significantly increase the impact zone (Figure 3).

In the soils adjacent to traffic intersections and

highways, a slight direct correlation was found between the pH_{salt} indicators and the concentration of oil products ($r = +0,41$). Studies carried out earlier for the city territory [14, 16, 17, 5, 9] have revealed stable correlations in the “source

Table 2: Physical and chemical properties of the surface soil layer (0–20 cm) in various functional areas of the city of Voronezh

Areas	Indicator	$pH_{aq.}$	pH_{salt}	Chlorides (mg/kg)	Mobile sulfur (mg/kg)	Exchangeable ammonium (mg/kg)
Background areas	Average	6,74	6,55	75,97	15,56	6,21
Industrial functional areas	Average	7,49	7,29	92,75	28,35	9,78
	Min	6,65	6,68	67,45	2,00	2,50
	Max	8,15	7,66	118,04	53,60	23,66
	Max-Min	1,50	0,98	50,59	51,60	21,16
Transport functional areas	Average	7,39	7,12	95,87	31,82	9,59
	Min	6,12	6,24	67,45	14,81	2,50
	Max	8,33	8,14	181,32	54,47	27,86
	Max-Min	2,21	1,90	113,87	39,66	25,36
Recreational functional areas	Average	6,01	5,73	73,8	33,83	9,13
	Min	5,2	4,21	67,45	24,51	2,49
	Max	7,21	7,17	84,31	43,16	15,28
	Max-Min	2,01	2,96	16,86	18,65	12,8
Residential functional areas	Average	6,5	6,02	92,73	31,93	9,88
	Min	5,0	3,86	50,59	12,5	2,5
	Max	7,86	7,49	118,04	52,2	26,9
	Max-Min	2,86	3,63	67,45	39,7	24,4

of pollution - soil” system. Strong positive correlations between concentrations and car traffic intensity have been reliably established.

Compared to the background areas, urban soils are rich in plant nutrients. Extremely high or low content of nutrients in soils determines the resistance of plants.

The signs of critical parameters of soil indexes are the suppression of plant associations for each key object [18].

Urban soils are not used for agricultural production, and the main removal of nutrients occurs during the annual harvesting of fallen leaves. The loss of nutrients leads to the destruction of mineral and organic components and worsens soil fertility.

The need for cleaning activities in the city is determined by the functional purpose of landscaping objects. Thus, in the changeable climate of the Voronezh region, lawns not cleared of autumn leaves are less enduring in the winter season and quickly degrade. In city parks, boulevards, and central streets, foliage harvesting is important not only for aesthetic but also for environmental reasons: all urban green spaces during the growing season receive a complex set of pollutants from the air, including compounds of heavy metals, benzo(a)pyrene, sulfur, etc.

The territory of Voronezh city is distinguished by a high content of mobile sulfur. High contents were found in various functional areas of the city, which are characterized by natural soil formation. Historical and landscape parks, green spaces of departmental territories, forests, and forest parks have preserved areas with natural soil formation and foliage harvesting is not advisable here. The soils of these territories accumulate the nutrient element (mobile sulfur) better than others.

The highest concentrations of mobile sulfur (over 16 mg/kg) are found in the forest-park and park area, as well as in the area of the Peschany Log memorial. Among the industrial functional areas, the territories adjacent to the combined heat and power production plant feature high concentrations of mobile sulfur. In the recent historical past, heat and power supply enterprises used coal. As a “legacy” from this period, the city has got “ash dumps” on the banks of the Voronezh water reservoir.

Urban soils, sandy loam, and meadow soils of the left bank of the city are insufficiently supplied with mobile sulfur (the content is less than 6 mg/kg). These soils make less than 14 % of all studied soil samples.

The content of exchangeable ammonium is a relative indicator characterizing the fertility of the

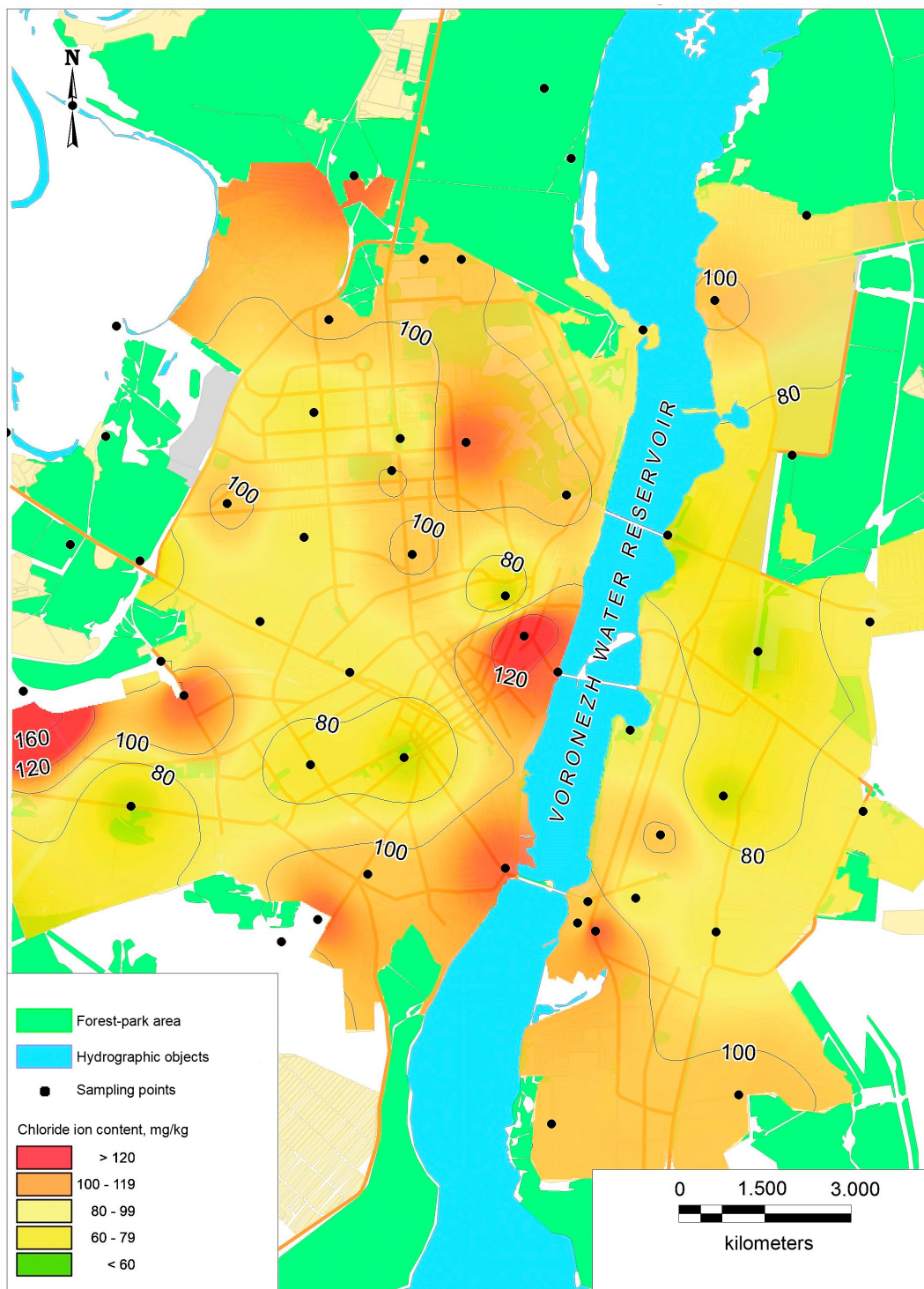


Figure 3: Distribution of chloride ion in the surface layers of soils in Voronezh

soil surface layer (Table 2). According to the level of exchangeable ammonium content, the functional areas of the city form a series: residential area (1,92 Kc) > industrial area (1,90 Kc) > transport area (1,86 Kc) > recreational area (1,77 Kc).

Under the influence of alkalization, urban soils

are transformed, the soil absorbing elements are saturated, the migration capacity of many heavy metals decreases and leads to their accumulation (Table 3).

One of the main processes affecting migration in soils is the fixation with the humic matter. It occurs

Table 3: The content of heavy metals and benzo(a)pyrene (mg/kg) in the surface soil layer (0-20 cm) of the city of Voronezh

	Indicator	Lead	Cadmium	Copper	Zinc	Benz(a)pyrene
Maximum permissible concentrations [19]	MPC (maximum permissible concentration) / APC (approximate permissible concentration)	32	0,5	33	55	0,02
Industrial functional areas	Average	69,21	1,27	23,83	14,37	0,050
	min	3,04	0,12	2,96	0,38	< 0,005
	max	453,00	8,02	93,02	123,80	0,170
	max - min	449,96	7,90	90,06	123,42	0,17
Transport functional areas	Average	25,9	0,20	21,24	98,75	0,028
	min	0,79	0,01	4,13	0,61	< 0,005
	max	134,7	1,29	94,7	499,98	0,360
	max - min	133,91	1,28	90,57	499,37	0,355
Recreational functional areas	Average	7,7	0,06	6,39	11,01	< 0,005
	min	0,59	0,02	0,42	0,26	< 0,005
	max	18,73	0,15	14,02	36,86	< 0,005
	max - min	18,14	0,13	0,19	0,2	0,00
Residential functional areas	Average	14,52	0,18	11,83	6,15	< 0,005
	min	0,00	0,00	3,32	1,98	< 0,005
	max	40,50	0,44	26,70	10,42	< 0,005
	max - min	40,50	0,44	23,38	8,44	0,00

due to the formation of salts with organic acids by metals, the adsorption of ions on the surface of organic colloidal systems, or the creation of complexes with humic acids.

The studied parameters are distributed in different ways in urban soils. Heavy metals are widespread in industrial and transport functional areas of the city. According to the present research, the soil adjacent to the railway tracks leading to the production areas is under maximum stress. Thus, on the left bank of the city in the area of the railway bridge (Otrozhka station) and the production base of the locomotive depot, lead concentrations exceed the MPC by 4,5 times, and in the area of the Pridacha railway station on the border with the aircraft building plant, the MPC is exceeded 14,2 times.

The industrial areas of the city have a high content of lead and cadmium. The average concentrations in soils exceed the MPC of lead and cadmium by more than 2 times. Within the industrial territories, the excess of MPC of lead was found in 42 % of the samples taken, cadmium in 25 %, and copper in 16,7 %.

In general, the soils of the Voronezh region are characterized by a deficiency of microelements - copper and zinc. Revealing their high

concentrations makes it possible to speak about their technogenic sources in the industrial areas of the city, such as the chemical and mechanical processes of metal processing and the processes of burning organic fuel.

Transport functional areas throughout the city concentrate heavy metals. On sections of streets with the maximum traffic intensity, pollutants of the 1st and 2nd hazard classes are accumulated. The values of the total pollution indicator for transport areas vary in the range from permissible to moderately hazardous values.

In smooth and sloped sections of transport functional areas, pollutants are accumulated in different ways. The actual accumulation of heavy metals in the upper and middle parts of the slopes of 3 to 5 degrees was established for the territory of the city. On the soils of the slightly sloping areas of the Voronezh River embankments, pollutants are practically not accumulated ($Z_c < 16$). This is due to leaching and the light texture of the soil.

High concentrations of benzo(a)pyrene are characteristic of anthropogenically transformed soils of industrial areas, as shown Figure 4.

For the areas of the left bank, the content of benzo(a)pyrene varies in the range from 0,05 to 0,08 mg/kg, for the areas of the right bank from

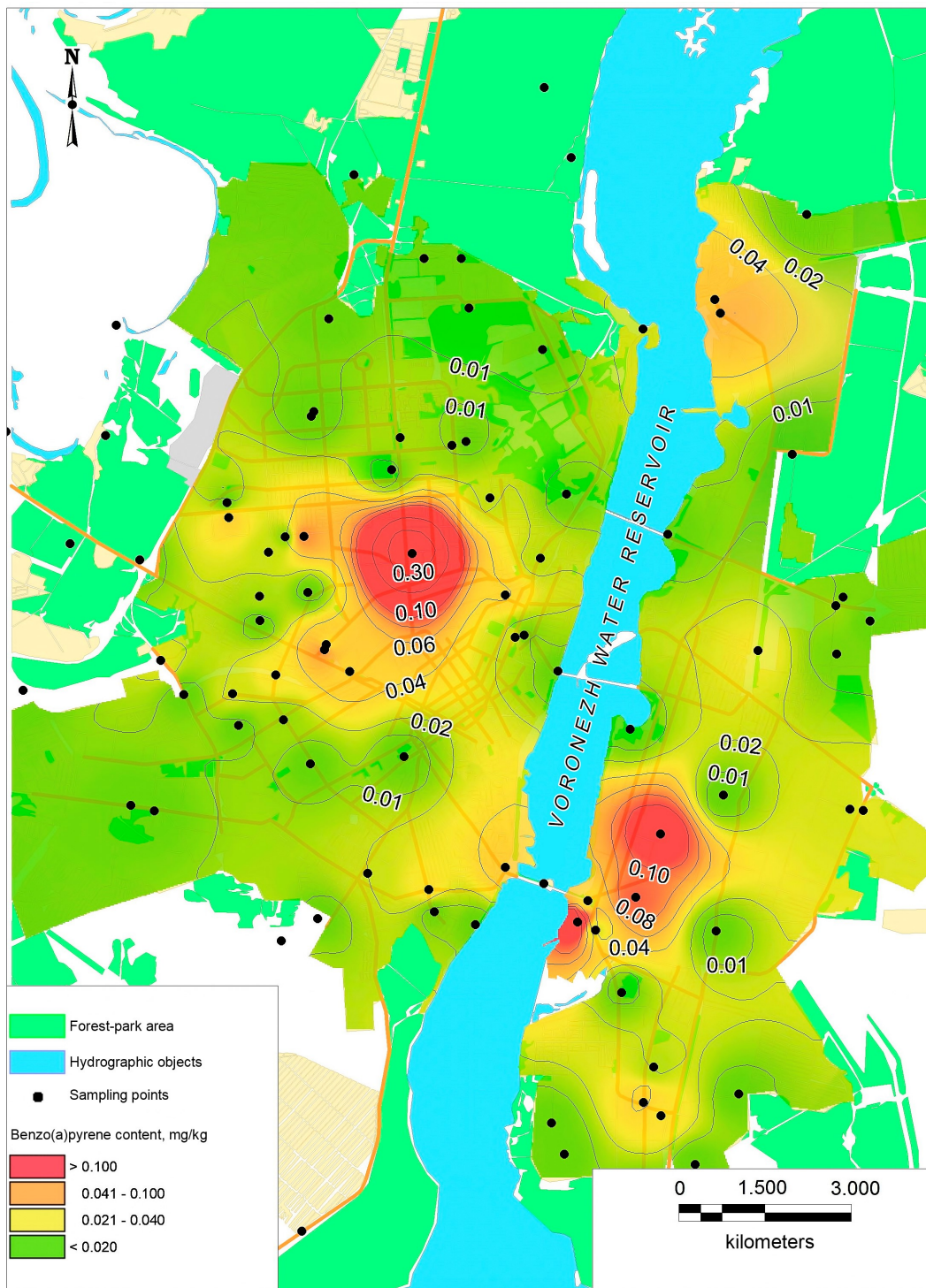


Figure 4: Accumulation of benzo(a)pyrene in the surface layers of soils in Voronezh

0,05 to 0,17 mg/kg. The soils of recreational and residential areas in the city do not contain PAHs.

On the territory of the city, the excess of MPC for benzo(a)pyrene was found in every second sample of soil from industrial areas and in every third sample from transport functional areas.

At the intersection of sanitary protection zones of enterprises with road and rail transport interchanges, the maximum concentration of benzo(a)pyrene is reached. Specification of the boundaries of "polluted" areas and identification of facilities of negative impact at the moment is

not possible. These enterprises are classified as highly hazardous industrial facilities with limited access to the territory.

When comparing the average concentrations of benzo(a)pyrene in the soils of the functional areas of the city with the MPC (0,02 mg/kg), the coefficient of environmental hazard (K_o) was calculated. The excess of the standard in the soils of the industrial areas of the city is quite high and averages 5 MPC for the left-bank part of the city and 2,5 MPC for the right-bank part. The maximum values ($K_o = 8,5$) are noted in the urban soils of the industrial area of the left bank near the Combined heat and power production plant 1.

The concentration coefficient (K_c) shows a strong degree of soil pollution in the industrial areas of the city. Particularly dangerous is the territory of the transport area in the central part of the city at the intersection of a busy road junction of motor transport (Zastava square), railway transport (Plekhanovskaya station), and the production area of the refrigeration plant ($K_o = 18$). By calculating the total indicator of soil pollution (Z_c), these territories were classified in this study as "hazardous" in terms of the degree of pollution ($32 < Z_c < 128$).

Most of the city is categorized as an area with an acceptable pollution level ($Z_c < 16$). These territories can be used without restrictions for housing construction, developing areas of cultural leisure, and business buildings.

Moderately hazardous soil pollution ($16 < Z_c < 32$) with heavy metals and benzo(a)pyrene was revealed in areas adjacent to the steep slopes of the right bank of the Voronezh water reservoir. Sloped surfaces, rainsheet, and linear erosion facilitate the migration of pollutants from the soil into the water reservoir. A significant amount of man-made compounds and suspended substances from the entire urban area enter the water of the reservoir with the runoff from the storm sewer network and treatment facilities of the left bank, as shown in Table 4.

Above the water area of the reservoir, special climatic conditions are created with changes in the speed and direction of the wind during the day. These movements are permanent and affect the

transport of substances along the entire bank line.

Analysis of the granulometric composition of the bottom sediments of the Voronezh water reservoir revealed the predominance of fractions of physical sand. In total, coarse, medium, and fine sand make up more than 60 % of the sample mass. Clay fractions, including silts, make on average about 30 % of the mass in all studied samples.

The chemical composition of the bottom sediments of the water reservoir is characterized by great diversity and is a consequence of long-term anthropogenic impact. Comparison of different parts of the reservoir was carried out in the present study according to the concentration coefficient (K_c) of each pollutant in comparison with the background indicator, as shown in Table 4.

As the background concentration of the pollutant, the bottom sediments were used, the anthropogenic impact on which is minimal. These are the sediments at the site in the north of the urban agglomeration, above the M-4 highway at the influx of the Voronezh River in the reservoir.

Long-term observations of the bottom sediments of the Voronezh water reservoir make it possible to identify the most "polluted" parts of the reservoir and to determine the periods during which the reservoir experienced the greatest technogenic loads.

The highest concentration coefficients of heavy metals (K_c) are typical for 2008 - 2010. The K_c coefficient values increase from north to south, in the direction of water flow. The greatest pollution of bottom sediments of the reservoir on the territory of the city is typical for the area below the outlet from the treatment facilities of the left bank.

The levels of pollution of bottom sediments with heavy metals increase in the following order: lead < cadmium < copper < zinc. Thus, the main heavy metals polluting the water body are zinc and copper.

4. Conclusions

Based on the results of the study, the following conclusions can be made:

In the surface layers of Voronezh soils, the pH value changes. Alkalization of the soil occurs,

Table 4: Accumulation coefficient (Ko) of heavy metals in the bottom sediments of the Voronezh water reservoir

No.	Sampling spot	Year	Lead	Cadmium	Zinc	Copper
1	Water reservoir 500 m downstream the railway bridge	2008	1,7	3,3	6,5	15,7
		2010	1,9	1,0	7,5	8,2
		2018	1,4	1,0	8,0	3,8
		2019	1,9	1,0	5,4	3,0
		2020	1,2	1,0	4,2	3,0
2	Water reservoir 500 m downstream the North bridge	2008	5,6	1,7	5,7	5,7
		2010	5,5	3,7	38,6	43,2
		2018	1,1	1,0	3,5	3,7
		2019	1,6	1,0	2,6	2,5
		2020	1,1	1,0	1,8	2,3
3	Water reservoir 500 m downstream the Chernavsky bridge	2008	9,4	0,9	13,9	5,7
		2010	1,1	1,0	6,4	6,2
		2018	1,8	3,0	4,9	2,0
		2019	1,3	1,0	2,7	2,3
		2020	1,2	1,0	1,9	2,4
4	Water reservoir between the dam of the Chernavsky bridge and the Pridachenskaya dam	2008	2,2	1,7	11,7	1,0
		2010	1,8	1,0	2,4	5,4
		2018	2,3	1,0	4,6	4,2
		2019	0,4	1,0	2,0	1,8
		2020	1,2	1,2	1,9	1,7
5	Water reservoir in the Petrovskaya embankment area	2008	8,9	0,1	17,4	8,6
		2010	8,1	3,0	37,7	34,1
		2018	1,7	1,0	4,2	2,7
		2019	1,4	1,0	3,8	2,5
		2020	1,2	0,9	2,5	2,6
6	Water reservoir 500 m downstream the bridge of Voronezh State District Power Plant	2008	16,7	1,8	44,8	25,7
		2010	12,9	7,6	27,1	81,8
		2018	1,4	3,0	8,4	4,5
		2019	2,2	5,2	6,5	2,8
		2020	2,0	3,8	5,9	3,1
7	Water reservoir 500 m downstream the outlet of "Levoberezhnye wastewater treatment facilities"	2008	3,3	19,2	28,7	60,4
		2010	33,1	180,1	235,5	81,8
		2018	7,2	16,0	22,4	46,7
		2019	10,1	22,3	16,1	24,4
		2020	12,4	20,5	15,9	22,5
8	Water reservoir 500 m upstream the waterworks	2008	1,1	0,1	2,2	0,9
		2010	4,5	2,2	25,0	43,2
		2018	1,7	1,0	12,7	18,0
		2019	2,0	1,0	8,4	9,0
		2020	1,2	1,2	7,5	6,9

which in turn affects the solubility of heavy metals and their concentration in the soil.

High levels of soil pollution are typical for the centers of geochemical abnormalities in industrial and transport functional areas.

The soil cover in Voronezh contains excessive concentrations of metals, which makes it possible to identify areas that are polluted by man-made sources.

Certain geochemical abnormalities have a diverse pattern, which often does not reflect a real assessment of the state of the environment. The distribution of elements is largely determined

by landscape and geochemical conditions. Metal content abnormalities are most evident on the windward side of pollutant emission sources. They are intensely accumulated in temporary drying up streams and in the troughs of untreated rainwater and meltwater.

The process of accumulation of zinc and copper in the bottom sediments of the reservoir is facilitated by the processes of degradation of the urban soil surface layer, which is deficient in these biogenic elements.





Landscape-geochemical conditions and functional organization of urban areas are important

factors that determine the level of anthropogenic impact on the soil and bottom sediments.

5. References

- [1] A. N. Nasonov, V. V. Kulnev, and M. V. Grafkina, "Modeling the dynamics and forecast of the environmental conditions of the air environment of residential areas," *Ecology and development of society*, vol. 2, no. 29, pp. 56–63, 2019, [in Russian].
- [2] L. M. Akimov, P. M. Vinogradov, and E. L. Akimov, "Comprehensive assessment of the ecological situation taking into account the state of the atmosphere and the functional planning structure of the city," *Bulletin of Voronezh State University. Ser. Geography. Geoecology*, no. 4, pp. 57–67, 2014.
- [3] G. A. Antsiferova, V. V. Kulnev, S. L. Shevyrev, E. V. Bepalova, N. I. Rusova, and A. E. Skosar, "Artificial water bodies of the Voronezh river basin and algobiotechnology in water quality management," *Ecology and Industry of Russia*, vol. 22, no. 8, pp. 50–54, 2018.
- [4] M. A. Klevtsova, V. V. Sivachenko, and Y. N. Davydova, "Evaluation of the ecological state of the soil cover under green plantations on the right bank of the city of Voronezh," *Bulletin of Voronezh State University. Ser. Geography. Geoecology*, no. 1, pp. 150–156, 2013.
- [5] Russian Geographical Society, Voronezh State University, *Medical and ecological atlas of the Voronezh region*, S. A. Kurolap, T. I. Prozhorina, M. A. Klevtsova, P. M. Vinogradov, N. V. Kaverina, S. Eprintsev, L. Sereda, A. Skosar, I. Popova, O. V. Klepikov, Y. I. Stepkin, N. P. Mamchik, I. V. Colnet, E. M. Studenikina, and Y. S. Kalashnikov, Eds. Voronezh: Center for Hygiene and Epidemiology in the Voronezh Region, 2019.
- [6] L. A. Mezhdova, V. A. Sedykh, Z. M. Sagova, E. M. Repina, and V. V. Kulnev, "Geoecological analysis of air pollution in the Middle Podvoronezh area," *Astrakhan bulletin of ecological education*, vol. 52, no. 4, pp. 85–92, 2019.
- [7] N. N. Nazarenko, N. V. Kaverina, K. E. Stekolnikov, and I. D. Svistova, "Polycyclic aromatic hydrocarbons in urban soils of Voronezh," *Vestnik VSU. Ser. Chemistry. Biology. Pharmacy*, no. 1, pp. 92–97, 2017.
- [8] N. N. Nazarenko and I. D. Svistova, "The content of benzo(a)pyrene in the soils of urbanized territories (on the example of the city of Voronezh) [in Russian]," *Successes of modern natural science*, no. 1, pp. 142–146, 2016.
- [9] N. S. Kasimov, V. R. Bitukova, S. M. Malkhazova, N. E. Kosheleva, E. M. Nikiforova, N. V. Shartova, D. V. Vlasov, S. A. Timonin, and V. N. Krainov, *Regions and cities of Russia: an integral assessment of the ecological state*, N. S. Kasimov, Ed. Moscow: IP Filimonov M.V., 2014, [in Russian].
- [10] A. I. Fedorova, *Tree plantations of city streets, their stability and bioindication role. Forest eco-systems of the green zone of Voronezh: collection of articles. scientific. Art.*, Voronezh: Voronezh, 1999.
- [11] R. A. Kondaurov, "Forecasting of soil pollution with heavy metals," *Science and technology in the road industry*, vol. 1, no. 64, pp. 33–36, 2013.
- [12] R. A. Kondaurov, "Results of verification of long-term environmental monitoring and forecasting of pollution fields in the area of influence of highways," *Scientific Bulletin of the Voronezh State Architectural and Construction University. Construction and architecture*, vol. 4, no. 32, pp. 85–91, 2013.
- [13] T. A. Sokolova, I. I. Tolpeshta, and S. Y. Trofimov, *Soil acidity. Acid-base buffering of soils. Aluminum compounds in the solid phase of the soil and in the soil solution: textbook. manual for universities*. Tula: "Grif and K", 2012, [in Russian].
- [14] S. A. Kurolap and O. V. Klepikova, Eds., *Integral ecological assessment of the state of the urban environment*. Russian Geographical Society - Russian Foundation for Basic Research, Voronezh State University of Engineering Technologies, Center for Hygiene and Epidemiology in the Voronezh Region, 2015.
- [15] P. M. Vinogradov, S. A. Kurolap, and O. V. Klepikov, "Geoinformation support of medical and ecological monitoring of the urban environment (on the example of the city of Voronezh)," *Bulletin of Voronezh State University. Ser. Geography. Geoecology*, no. 4, pp. 39–47, 2014.
- [16] O. Klepikov, A. Samoilov, I. Ushakov, V. Popov, and S. Kurolap, "Comprehensive assessment of the state of the environment of an industrial city," *Hygiene and sanitation*, vol. 97, no. 8, pp. 686–692, 2018.
- [17] O. Klepikov, S. Kurolap, and P. Vinogradov, "Integral ecological and hygienic assessment of the territory of an industrial center," *Sanitary Doctor*, no. 6, pp. 20–26, 2016.
- [18] L. O. Sereda, L. A. Yablonskikh, and S. A. Kurolap, "Monitoring of the ecological and geochemical state of the soil cover of the city of Voronezh," *Bulletin of the Volgograd State University. Ser. Natural Sciences*, no. 2, pp. 66–72, 2015.
- [19] Chief State Sanitary Physician of the Russian Federation, "GN 2.1.7.2041-06. Maximum permissible concentration (MPC) of chemicals in soil: Hygienic standards.," Federal Center for Hygiene and Epidemiology of Rospotrebnadzor, Moscow, Tech. Rep., 2006.

Developing multilevel statistical models of determining bottom-hole flowing pressure in commercial oil well operations

Inna N. Ponomareva * , Vladislav Ign. Galkin , Alexander V. Rastegaev ,
Sergey VI. Galkin 

Department of Oil and Gas Technologies, Perm National Research Polytechnical University, Russia

Abstract.- One of the major tasks of monitoring oil production well operations is to determine bottom-hole flowing pressure (BHFP). The overwhelming majority of wells in the Perm krai are serviced using borehole pumps, which makes it difficult to take direct bottom-hole flowing pressure measurements. The BHFP in these wells is very often determined by recalculating the parameters measured at the well mouth (annulus pressure, dynamic fluid level depth). The recalculation is done by procedures based on analytically determining the characteristics of the gas-liquid mixture in the wellbore, which is very inconsistent to perform due to the mixture's complex behavior. This article proposes an essentially different approach to BHFP measurements, that relies on the mathematical processing of the findings of more than 4000 parallel mouth and deep investigations of the oil production wells of a large oil-production region. As a result, multivariate mathematical models are elaborated that allow reliably determining the BHFP of oil-production wells in operation.

Keywords: oil production well; bottom-hole flowing pressure; BHFP determination technique; multivariate statistical model; regression analysis; multilevel modeling.

Desarrollo de modelos estadísticos multinivel para determinar la presión de flujo de fondo en operación de pozos petroleros

Resumen.- Una de las principales tareas del monitoreo de las operaciones de los pozos de producción es determinar la presión de flujo del fondo del pozo. La inmensa mayoría de los pozos en Perm krai reciben servicio mediante bombas de pozo, lo que dificulta la toma de mediciones directas de la presión de flujo del fondo del pozo. La presión de flujo de fondo en estos pozos se determina muy a menudo recalculando los parámetros medidos en la boca del pozo (presión anular, profundidad del nivel de fluido dinámico). El recálculo se realiza mediante procedimientos basados en la determinación analítica de las características de la mezcla gas-líquido en el pozo, que es muy inconsistente de realizar debido al complejo comportamiento de la mezcla. Este artículo propone un enfoque esencialmente diferente para las mediciones de la presión de flujo del fondo del pozo, que se basa en el procesamiento matemático de los hallazgos de más de 4000 investigaciones paralelas de boca y profundidad de los pozos de producción de petróleo de una gran región de producción de petróleo. Como resultado, se elaboran modelos matemáticos multivariados que permiten determinar de manera confiable la presión de flujo de fondo de los pozos productores de petróleo en operación.

Palabras clave: pozo de producción; presión de flujo del fondo del pozo; técnicas de determinación de BHFP; modelo estadístico multivariable; modelado multinivel.

Received: December 20, 2020.

Accepted: February 24, 2021.

1. Introduction

Well operation monitoring is an integral part of oil and gas production optimization [1, 2, 3].

One of the topical tasks of this monitoring is to determine the BHFP [4, 5, 6]. In addition, the BHFP is the parameter linking the work of the elements of the reservoir-well system. The BHFP level can be used to control and manage the operation of downhole equipment [7]. As a matter of practice, if a well is operated using deep-pumping equipment, the BHFP is determined by calculation [8]. If the suction manifold is

* Correspondence author:
e-mail: ponomarevaIN@pstu.ru (I. Ponomareva)

equipped with a special instrument (sensor), the pressure measured with its help and referred to as the suction pressure is recalculated to the BHFP. As a rule, these calculations are not impaired by any particular difficulties and are made fairly stably. If the deep-pumping equipment configuration has no room for installing a measuring instrument at the suction manifold, the BHFP is determined by recalculating the parameters that are measured at the well mouth and include dynamic fluid level depth and annulus pressure [9]. This recalculation is mathematically based on the hydrostatic equation. In this case, the quantity to determine is the wellbore fluid density [10]. This fluid is a gas-liquid mixture the parameters of which are very difficult to describe by analytical equations [11, 12, 13]. In this article, all of the BHFP determination procedures based on recalculating the parameters measured at the well mouth and recalculated considering the gas-liquid mixture density are called density-based techniques.

In the Perm krai oil is produced at more than 5.000 wells only 30 % of which are equipped with downhole instruments. The BHFP at other wells is determined by recalculating the mouth parameters by various density-based procedures [14, 15, 16]. The accuracy of these techniques has been evaluated by several specialized investigations [17, 18, 19], which have revealed their low consistency and significant errors in the BHFP measurements, especially at deposits with reservoir oil highly saturated with gas.

The current investigations are aimed at choosing optimal BHFP control techniques [20, 21]. When the analytical description of physical processes is inapplicable for whatever reason, it is relevant to apply statistical techniques based on processing collected facts mathematically [22]. Statistical techniques are often easier to apply and more accurate than the attempt to elucidate analytical regularities in the behavior of complex physical systems. This approach does not require any significant reductions and suppositions, applies for any distribution laws, to systems of any complexity and multiplicity of states, and is restricted only by the actual parameters of the original sample. Statistical techniques are successfully applied for

solving various engineering tasks [23, 24, 25].

In this vein, the article describes an original multilevel approach elaborated by the authors. This technique allows constructing multivariate statistical models of determining the BHFP of wells in operation. These models are generalized for vast territorial entities.

2. Materials and methods

2.1. Characteristics of original data

To solve the formulated task, we used the data from $n = 4,145$ investigations in which the measurements were taken synchronously at the mouth and bottom of the flowing wells of the oil deposits confined to the Solikamsk depression, a major oil-production area in the Perm krai. The deposits were Un'vinskoye, Gagarinskoye, Siberian, Magovskoye, Ozeroye, Chashkinskoye, Shershnevskoye, and the one named after Sukharev. To create multivariate multilevel models, we used the data of the downhole BHFP measurements in the wells equipped with measuring devices at the suction manifold; in addition, such well performance characteristics were used as liquid rate (Q_l , m^3/day), oil rate (Q_o , t/day), water cut (w , %), fluid pressure in the annulus between tubing and casing (P_{an} , MPa), dynamic fluid level depth (H_{dfl} , m), pump measured depth (H_{pump} , m), distance from the pump hinge point to the dynamic fluid level depth (H_{pd}); oil-water contact (OWC) measured depth, and reservoir pressure (P_r). The BHFP recalculated from the suction manifold pressure was used as actual (P_{bh} , MPa).

2.2. Research features

A distinct feature of our approach is the different degree of differentiating the facilities for which the models are constructed; that is, the approach is multilevel. The levels for which the models are constructed are exposed below:

- Level one involves using the entire sample, without separating deposits and production targets (occurrences).
- Level two consists in considering each deposit without separating occurrences.

- Level three consists in considering each occurrence in a generalized sense, without taking the deposits into account.
- Level four implies constructing the model specifically for each occurrence within a deposit.

We also constructed for practical application the multivariate model covering the BHFP determination results for all of the levels.

3. Literature review

The investigation was conducted using available mathematical statistics tools [26].

In the preliminary phase, we analyzed the correlations between the BHFP and the well operation properties that might affect the BHFP level. To achieve this purpose, we calculated not only coefficient r of the correlation between the input parameters and the BHFP but also coefficient r of the mutual correlation among the input parameters. The calculated results are exposed as correlation matrices, correlation fields, and equations of regression between the BHFP and the well performance indicators. We conducted the investigations in question for the models from all of the levels.

The next phase of the investigation is the construction of multivariate multilevel models using the original cumulative sample technique. According to this approach, the initial data are all tentatively graded against the range of BHFP levels from minimal to maximal. The first model is constructed according to the first three graded data (the amount of data per sample is $n = 3$). Then the model is constructed for $n = 4$. Thus multivariate models are consecutively constructed until all of the available data are used. In these variants, the multivariate models are constructed by step-by-step regression analysis. The dependent feature is P_{bh} , whereas the values of the rest of the above specified well performance indicators are used as independent factors. Step-by-step regression analysis was used not only to derive the equation of regression, defined as a multivariate statistical model, but also to identify the existence

and kind of influence of the independent factors on the dependent variable.

The regression coefficients in the elaborated models were calculated by the least square technique. The functionality of each model was assessed in several ways. First of all, such statistical characteristics of the models were calculated in each case, as multiple correlation coefficient R , significance level p , and standard error S_0 . The applicability limits of each model are defined. To assess the model of each level for functionality, it was used to calculate the BHFP also compared with the actual levels. The comparison was made while visually analyzing the correlation fields, analyzing the equations of regression between the actual and the calculated BHFP levels, and with the help of such other mathematical statistics means as Student's t-test and Pearson's chi-squared test [27, 28].

Thus the models were constructed for all of the four distinguished levels. Four domes are distinguished within the confines of the Un'vinskoe deposit. To take this peculiarity into account, we additionally constructed the models of each of the domes as part of modeling level four.

The analysis of constructed models should be considered a major part of any kind of modeling, multivariate modeling included. In this case, the analysis of all of the models involved examining the succession and frequency of including each of the input parameters in the equations of regression. It is considered that, the more often is a particular parameter found in multivariate models and the earlier it is used in model construction, the larger will be its effect on the BHFP. Thus this analysis allowed distinguishing the factors that had the biggest effect on the BHFP level registered when the commercial wells of the oil deposits in the considered region were in operation.

The multivariate equations of regression derived at all of the modeling levels are supposed to be used together, for which purpose the generalized multivariate model is constructed. It is proposed for use as the mathematical basis of the technique of measuring the BHFP in well operation.

The final phase was the investigation of the accuracy of determining the BHFP using

the elaborated multivariate models. For that purpose, we calculated the model BHFP levels (PbhM) and then compared them with the actual BHFP levels (Pbh). This phase also involves comparing the elaborated technique, based on using multivariate models, with the currently applied BHFP calculation technique based on using the density of the gas-liquid mixture in the wellbore. For that purpose, we also calculated PbhM using the conventional technique and then compared the results with the actual BHFP levels. The elaborated technique based on multivariate multilevel models and its currently applied counterpart based on calculating the density of the gas-liquid mixture in the wellbore were compared by drawing the correlation fields when investigating the correlation (equations of regression and their characteristics) between the actual and the calculated BHFP for the whole sample and the separate deposit development targets.

It is supposed that the joint consideration of the regression equation parameters and their statistical characteristics will allow evaluating not only the tightness but also the kind of relations between the actual BHFP and the BHFP calculated by the two techniques.

4. Results

4.1. Analyzing correlations between BHFP and well performance indicators

The correlation matrix drawn for the whole examined sample (level one) and characterizing the influence of the well performance indicators on the BHFP level is shown in Table 1.

The correlation matrix of level two is represented for the Un'vinskoye deposit as the largest one in the region and exposed in Table 2.

Similar correlation matrices were drawn for all of the deposits in the considered region. The example of a correlation matrix for investigation level three is presented for the occurrences confined to Bobrikovian sedimentations (Table 3).

The correlation matrices at level four were made up for each occurrence within the deposits. The example exposed in Table 4 is the correlation

matrix for the Bobrikovian occurrence of the Un'vinskoye deposit. The making up of the matrix involved calculating coefficients r using all the data (upper string) and, separately, for the domes (lines two, three, and four, and the bottom line are for the Un'vinskii, Palasherskii, Southeastern, and Bystrovskii dome, respectively).

Similar correlation matrices were made up for all of the occurrences of all the deposits. All in all, 29 correlation matrices were made up for different modeling levels; the matrices included 1305 values of r .

4.2. Multivariate BHFP determination models

Then, multivariate models were constructed for all of the levels. The model for level one is recorded as is shown in equation (1)

$$\begin{aligned}
 P_{bh}^{M1} = & 1,163 + 0,0042 \cdot H_{bound} \\
 & + 0,009 \cdot H_{owc} - 0,009 \cdot H_{pump} \\
 & + 0,022 \cdot w + 0,601 \cdot P_{an} \\
 & + 0,097 \cdot Q_l + 0,040 \cdot P_r.
 \end{aligned} \quad (1)$$

The statistical accuracy characteristics determined for the model were multiple correlation coefficient $R = 0,763$, significance level $p < 0,0000$, and standard calculation error $S_0 = 1,76$ MPa. The model forming sequence is provided in the equation of regression. The coefficients describing the reliability of the statistical relations changed as follows: $r = 0,505$; $R = 0,612$; $0,703$; $0,736$; $0,760$; $0,762$; $0,763$.

The multivariate models at level two were constructed separately for the deposits, without distinguishing the occurrences. The model provided as an example was made up for Un'vinskoye deposit as the largest deposit in the region and is recorded in equation (2)

$$\begin{aligned}
 P_{bh}^{M2-U} = & -6,637 - 0,00718 \cdot H_{df1} \\
 & + 0,0086 \cdot H_{owc} + 0,021 \cdot w \\
 & + 0,684 \cdot P_{an} + 0,3054 \cdot P_r \\
 & - 0,0021 \cdot H_{bound},
 \end{aligned} \quad (2)$$

at $R = 0,822$, $p < 0,0000$, $S_0 = 1,52$ MPa. The model was formed in the sequence presented in the

Table 1: Matrix of the correlation between bottom-hole flowing pressures and well performance indicators (level one)

	P_{bh}	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
P_{bh}	1	-0,45*	0,15*	0,19*	0,38*	0,28*	0,48*	0,2*	0,51*	0,32*
H_{dfl}		1	0,13*	0,05*	-0,26*	-0,29*	0,04*	0,07*	-0,87*	-0,02
P_{an}			1	-0,03	0,21*	0,19*	0,07*	0,06*	-0,09*	0,09*
W				1	0,17*	-0,25*	0,02	0,04*	-0,05*	0,33*
Q_l					1	0,83*	0,3*	0,27*	0,34*	0,36*
Q_o						1	0,27*	0,27*	0,38*	0,26*
H_{owc}							1	0,75*	0,3*	0,54*
H_{pump}								1	0,39*	0,53*
H_{pd}									1	0,2*
P_r										1

Note: *-significant correlations

Table 2: Matrix of the correlation between bottom-hole flowing pressures and well performance indicators (level two, Un’vinskoye deposit)

	P_{bh}	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
P_{bh}	1	-0,72*	0,09	-0,02	0,3*	0,26*	0,42*	-0,07	0,65*	0,17*
H_{dfl}		1	0,14*	0,21*	-0,24*	-0,31*	-0,24*	0,02	-0,92*	-0,13*
P_{an}			1	-0,1*	0,31*	0,13	-0,03	-0,03	-0,15	-0,07
W				1	0,03	-0,43*	-0,28*	-0,12*	-0,24*	0,22*
Q_l					1	0,85*	0,2*	0	0,22*	0,09*
Q_o						1	0,24*	0,05	0,31*	0,01
H_{owc}							1	0,31*	0,34*	-0,25*
H_{pump}								1	0,36*	-0,14*
H_{pd}									1	0,07
P_r										1

Note: *-significant correlations

Table 3: Matrix of the correlation between bottom-hole flowing pressures and well performance indicators (level three, Bobrikovian sedimentations)

	P_{bh}	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
P_{bh}	1	-0,51*	0,02	0,25*	0,21*	0,09*	0,13*	-0,26*	0,37*	0,26*
H_{dfl}		1	0,22	0,15	-0,18*	-0,29*	-0,01	0,02	-0,92*	-0,04
P_{an}			1	-0,05	0,28*	0,28*	0,18*	-0,03	-0,21*	0,08
W				1	0,19*	-0,37*	0,02	-0,11*	-0,18	0,22*
Q_l					1	0,77*	0,2*	0,01	0,17*	0,16*
Q_o						1	0,11*	0,07*	0,3*	0
H_{owc}							1	0,2*	0,09	0,59*
H_{pump}								1	0,38*	0,04
H_{pd}									1	0,05
P_r										1

Note: *-significant correlations

equation of regression. The coefficients describing the reliability of statistical relations varied as follows: $R = 0,717; 0,762; 0,790; 0,808; 0,815; 0,822$. Similar models were also constructed for the Siberian, Shershnevskoye, Gagarinskoye, Ozernoye, Magovskoye, and Chashkinskoye deposit. Level three implies constructing multivariate models for the main occurrences distinguished in particular deposits within the considered

region. The model for the Bobrikovian occurrence developed at almost all of the deposits is recorded in equation (3)

$$P_{bh}^{M3-bb} = 19,684 - 0,00404 \cdot H_{dfl} + 0,025 \cdot w - 0,0037 \cdot H_{pump}, \quad (3)$$

at $R = 0,644, p < 0,0000, S_0 = 1,9$ MPa. The model was formed in the sequence exposed in the

Table 4: Correlation matrix of the influence of performance indicators (level four, Un’vinskoye deposit and Bobrikovian sedimentations)

	P_{bh}	Q_l	Q_o	B	P_{an}	H_{dfl}	H_{pump}	H_{owc}	P_r
	1	0,22*	0,16*	0,05	0,21*	-0,44*	-0,19*	0,23*	0,07
	1	0,11*	0,03	0,17*	0,08	-0,46*	-0,11*	-0,01	-0,05
P_{bh}	1	0,28*	0,28*	-0,06	0,17*	-0,65*	-0,33*	0,45*	0,27*
	1	-0,15	0,57*	-0,39	-0,23	-0,48*	-0,12	-0,33	0,23
	1	-0,71	-0,72*	0,79*	0,81*	0,52	-0,72*	-	-
Q_l		1	0,77*	0,05	0,33*	-0,25*	-0,04	-0,17*	-0,02
		1	0,76*	0,08	0,14*	-0,36*	0,09	-0,39*	-0,06
		1	0,76*	0,12*	0,37*	-0,28*	-0,22*	-0,11	-0,12
		1	-0,15	0,54*	-0,09	-0,26	0,88*	-0,58*	-0,9*
		1	0,97*	-0,78*	-0,59	-0,08	0,75*	-	-
Q_o			1	-0,5*	0,31*	-0,23*	-0,05	-0,1	-0,04
			1	-0,48*	0,22*	-0,33*	0,09	-0,29*	-0,04
			1	-0,51*	0,48*	-0,27*	-0,23*	-0,04	0,01
			1	-0,72*	-0,11	-0,30	-0,4	-0,37	0,07
			1	-0,88*	0,53	-0,01	0,86*	-	-
B				1	0,14*	0	0	-0,07	-0,18*
				1	0,02	0,05	-0,04	0,01	-0,26*
				1	-0,22*	0,08	0,08	-0,17*	-0,27*
				1	-0,06	0,07	0,69*	0,23	-0,32*
				1	0,48	0,05	-0,99*	-	-
P_{an}					1	0,18*	-0,01	0,19*	0,04
					1	0,12*	0,01	0,05	-0,37*
					1	0	-0,03	0,01	-0,14
					1	0,5*	-0,22	0,23	-0,06
					1	0,73*	-0,38	-	-
H_{dfl}						1	0,09*	0,27*	0,13*
						1	0,03	0,4*	-0,09
						1	0,25*	-0,16*	0,33*
						1	-0,3	0,5*	0,18
						1	0,07	-	-
H_{pump}							1	-0,29*	0,08
							1	-0,2*	0,05
							1	-0,5*	0,17*
							1	-0,54*	-0,64*
							1	-	-
H_{owc}								1	0,14*
								1	-0,09
								1	-0,31*
								1	0,52*
								1	-
P_r									1
									1
									1
									1
									1

Note: *-significant correlations

equation of regression. The values of coefficient R changed as follows: 0.512; 0.608; 0.644.

The model for all of the occurrences was derived similarly.

Then we built multivariate level-four models specifically for the occurrences within the deposits.

For example, the presented level-four model is recorded for the Bobrikovian occurrence of the Un’vinskoye deposit is shown in equation (4)

$$P_{bh}^{M4-U-bb} = 17,902 - 0,0053 \cdot H_{dfl} + 0,0257 \cdot w + 0,5906 \cdot P_{an} - 0,0026 \cdot H_{pump}, \quad (4)$$

at $R = 0,763$, $p < 0,0000$, $S_0 = 1,62$ MPa. The model was formed in the sequence exposed in the equation of regression. The values of coefficient

R changed as follows: 0,683; 0,723; 0,748; 0,763. For the ranges, in which all of the above recorded models can be used, see Table 5.

The multivariate models for each of the domes of the Un'vinskoye deposit are recorded as follows:

a) Un'vinskiy dome:

$$P_{bh}^{M4-1} = 4,430 - 0,0024 \cdot H_{dfl} + 0,003 \cdot H_{owc} + 0,011 \cdot w + 0,449 \cdot P_{an} - 0,0055 \cdot Q_l \quad (5)$$

at $R = 0,552$, $p < 0,0000$, $S_0 = 1,44$ MPa. The model was formed in the sequence exposed in the equation of regression (5). The values of coefficient R changed as follows: 0,460; 0,497; 0,532; 0,549; 0,552.

b) Palasherskiy dome:

$$P_{bh}^{M4-2} = -57,862 - 0,0045 \cdot H_{dfl} + 0,0075 \cdot H_{owc} + 0,7014 \cdot P_{an} + 0,0172 \cdot Q_l - 0,0177 \cdot Q_o \quad (6)$$

at $R = 0,770$, $p < 0,0000$, $S_0 = 1,37$ MPa. The model was formed in the sequence exposed in the equation of regression (6). The values of coefficient R changed as follows: 0,648; 0,734; 0,573; 0,759; 0,764; 0,770.

c) Southeastern dome:

$$P_{bh}^{M4-3} = -201,705 + 0,375 \cdot Q_o - 0,001 \cdot H_{dfl} + 0,104 \cdot H_{owc} + 0,004 \cdot H_{pump} \quad (7)$$

at $R = 0,729$, $p < 0,0121$, $S_0 = 0,91$ MPa. The model was formed in the sequence exposed in the equation of regression (7). The values of coefficient R changed as follows: 0,566; 0,652; 0,706; 0,729.

d) Bystrovskiy dome:

$$P_{bh}^{M4-4} = -6,775 + 5,321 \cdot P_{an} + 0,22617 \cdot w + 0,0041 \cdot H_{dfl} \quad (8)$$

at $R = 0,943$, $p < 0,02155$, $S_0 = 0,94$ MPa. The model was formed in the sequence exposed in the equation of regression (8). The values of

coefficient R changed as follows: 0,805; 0,927; 0,943.

The value ranges of the indices, in which the models derived for each elevation can be used, are exposed in Table 6.

The correlation of the calculated and the actual BHFP levels was examined by making up equations of regression and exposed in Table 7.

The frequency of inclusion (incidence) of the indices in the models of all the levels was calculated during their analysis and is shown in Table 8.

The generalized model is recorded as shown in equation (9)

$$P_{bh}^{MM} = -0,089 + 0,689 \cdot P_{bh}^{M-3} + 0,361 \cdot P_{bh}^{M-2} - 0,039 \cdot P_{bh}^{M-1} \quad (9)$$

at $R = 0,941$, $p < 0,0000$ and the mean standard error is 0,45 MPa. The model was formed in the sequence exposed in the equation of regression. The values of coefficient R describing the strength of the statistical relations changed as follows: 0,848; 0,851; 0,941.

The data of the results of comparing which are exposed below were obtained by two techniques, the new one based on applying multivariate models and the conventional one based on calculating the gas-liquid mixture density.

The correlation fields for the sample in general are presented in Figure 1.

The equations of regression between the actual BHFP and the BHFP calculated by the two techniques are shown in Table 9 as well as the statistical characteristics of these equations, including the correlation coefficient, the significance level, and the standard calculation error. The equations are compared at two levels, i.e., using all the data (level one) and, separately, for the occurrences of all of the deposits (level two).

In addition, to comparatively evaluate the functionality of the techniques as applied to specific occurrences, the respective correlation fields were drawn by the example of Un'vinskoye as the region's most representative deposit (Figure 2).

Table 5: Applicability ranges of the models of four levels

Model's index	Applicability range of model of levels			
	1	2	3	4
H_{dfl} , m	-	93,9-1946,5	206,9-1946,5	304,9-1946,5
P_{an} , 10^6 Pa	0-8,5	0,06-6,05	-	0,06-6,05
w , %	0,0-99,9	0,0-98	0-99,9	0,0-97,6
Q_l , m^3/day	-	-	-	-
Q_o , t/day	0-119,3	-	-	-
H_{owc} , m	1426,0-2332,7	1837,7-2246,7	-	-
H_{pump} , m	1201,4-2287,6	-	1439,3-2177,2	1445,3-2083,1
H_{pd} , m	3,8 -1720,5	17,1-1554	-	-
P_r , 10^6 Pa	6,4-23,3	13,9-20,4	-	-

Table 6: Applicability ranges of the models of the domes of the Un’vinskoye deposit (Bobrikovian occurrence)

Model's index	Applicability range of the models derived for domes			
	Un’vinskiy	Palasherskiy	Southeastern	Bystrovskiy
H_{dfl} , m	210 – 1935	513 – 2076	916 – 1416	1 194 – 1744
P_{an} , MPa	0,14 – 4,79	0,37 – 3,71		0,94 – 1,3
w , %	1 – 98,4			5-20
Q_l , m^3/day	0,8 – 98,7	0,5 – 107,2		
Q_o , t/day		0,2 – 71,2	0,11 – 6,36	
H_{owc} , m	2088 – 2519,5	2152 – 2553,2		
H_{pump} , m			1715 – 1970	
H_{pd} , m				
P_r , MPa				

Table 7: Equations of regression between the actual BHFP levels and the BHFP values calculated according to the level-four model, considering the discrimination of domes within the Un’vinskoye deposit

Dome	Equation of regression	r	p
Un’vinskiy	$P_{bh}^{M^4-1} = 5,416 + 0,303 \cdot P_{bh}$	0,549	0,0000
Palasherskiy	$P_{bh}^{M^4-2} = 3,484 + 0,598 \cdot P_{bh}$	0,766	0,0000
Southeastern	$P_{bh}^{M^4-3} = 3,769 + 0,532 \cdot P_{bh}$	0,729	0,0002
Bystrovskiy	$P_{bh}^{M^4-4} = 0,818 + 0,890 \cdot P_{bh}$	0,943	0,0004

Table 8: Incidence rate of the indices in the models of all the levels

Modeling level	Models input parameters								
	H_{dfl}	P_{an}	w	Q_l	Q_o	H_{owc}	H_{pump}	H_{pd}	P_r
One		0,142	0,142		0,142	0,142	0,142	0,142	0,142
Two	0,136	0,136	0,136	0,090	0,113	0,136	0,136	0,045	0,068
Three	0,121	0,121	0,090	0,121	0,090	0,121	0,090	0,121	0,121
Four	0,147	0,132	0,102	0,132	0,102	0,088	0,132	0,102	0,068
Total	0,131	0,131	0,111	0,111	0,105	0,111	0,125	0,095	0,079

5. Discussion

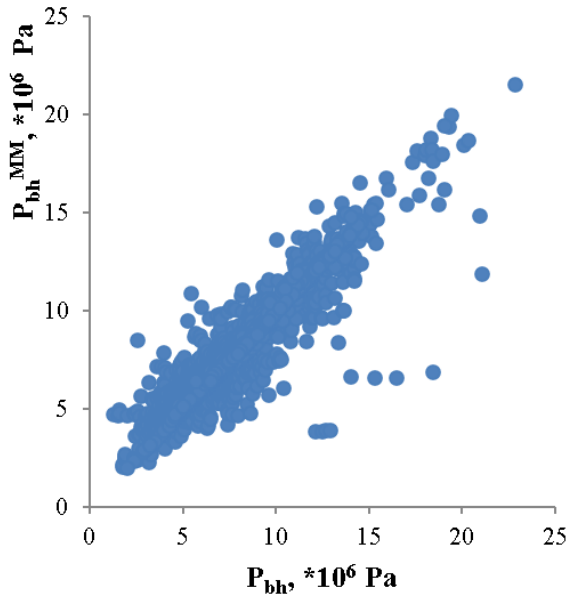
The exposed investigations should be considered the justification of the expediency of applying probabilistic statistical techniques to determining the bottom-hole flowing pressure treated as a major task in the oil extraction industry.

As shown by all of the conducted investigations, the creation of a stable BHFP determination technique is a complex task to solve.

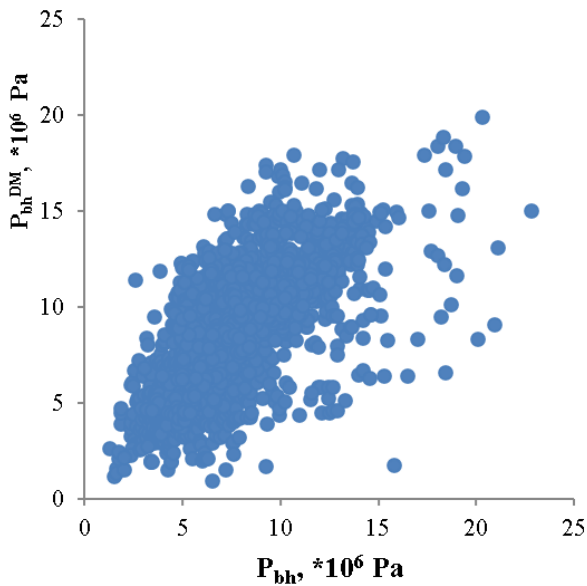
For example, as evidenced by the correlation analysis, in which 29 correlation matrices with 1305 correlation coefficient values were built for the four levels of investigation, the well

Table 9: Equations of regression between the actual BHFP and the BHFP calculated by the two techniques

First Level of comparison			
	Type and amount of data, sedimentation index	Equations of regression for technique	
	Whole sample	$P_{bh}^{MM} = 0,840 + 0,893 \cdot P_{bh}$ $r = 0,945; p < 0,0000$ $S_0 = 0,893 \text{ MPa}$	$P_{bh}^{DM} = 2,309 + 0,791 \cdot P_{bh}$ $r = 0,671; p < 0,0000$ $S_0 = 2,068 \text{ MPa}$
Second Level of comparison			
Un'vinskoye deposit	$C_{1t} - D_3fm$ $n = 174$	$P_{bh}^{MM} = 0,088 + 0,993 \cdot P_{bh}$ $r = 0,998; p < 0,0000$ $S_0 = 0,158 \text{ MPa}$	$P_{bh}^{DM} = 1,633 + 0,856 \cdot P_{bh}$ $r = 0,847; p < 0,0000$ $S_0 = 1,491 \text{ MPa}$
	C_{1v} $n = 889$	$P_{bh}^{MM} = 0,032 + 0,996 \cdot P_{bh}$ $r = 0,997; p < 0,0000$ $S_0 = 0,189 \text{ MPa}$	$P_{bh}^{DM} = 2,561 + 0,823 \cdot P_{bh}$ $r = 0,770; p < 0,0000$ $S_0 = 1,670 \text{ MPa}$
	C_{2b} $n = 143$	$P_{bh}^{MM} = 0,144 + 0,988 \cdot P_{bh}$ $r = 0,997; p < 0,0000$ $S_0 = 0,216 \text{ MPa}$	$P_{bh}^{DM} = 0,739 + 1,042 \cdot P_{bh}$ $r = 0,862; p < 0,0000$ $S_0 = 1,660 \text{ MPa}$
	C_{2vr} $n = 31$	$P_{bh}^{MM} = -0,013 + 0,988 \cdot P_{bh}$ $r = 0,998; p < 0,0000$ $S_0 = 0,045 \text{ MPa}$	$P_{bh}^{DM} = 14,369 - 0,424 \cdot P_{bh}$ $r = 0,374; p < 0,0378$ $S_0 = 0,968 \text{ MPa}$
Chashkinskoye deposit	$C_{1t} - D_3fm$ $n = 89$	$P_{bh}^{MM} = 0,513 + 0,946 \cdot P_{bh}$ $r = 0,972; p < 0,0000$ $S_0 = 0,713 \text{ MPa}$	$P_{bh}^{DM} = 2,704 + 0,710 \cdot P_{bh}$ $r = 0,876; p < 0,0000$ $S_0 = 1,225 \text{ MPa}$
	C_{1v} $n = 161$	$P_{bh}^{MM} = 2,476 + 0,750 \cdot P_{bh}$ $r = 0,865; p < 0,0000$ $S_0 = 0,965 \text{ MPa}$	$P_{bh}^{DM} = 5,955 + 0,467 \cdot P_{bh}$ $r = 0,497; p < 0,0000$ $S_0 = 2,272 \text{ MPa}$
Ozernoye deposit	$C_{1t} - D_3fm$ $n = 579$	$P_{bh}^{MM} = 1,746 + 0,727P_{bh}$ $r = 0,852; p < 0,0000$ $S_0 = 0,655 \text{ MPa}$	$P_{bh}^{DM} = 0,612 + 1,124P_{bh}$ $r = 0,617; p < 0,0000$ $S_0 = 2,107 \text{ MPa}$
	C_{2b} $n = 99$	$P_{bh}^{MM} = 0,241 + 0,958P_{bh}$ $r = 0,979; p < 0,0000$ $S_0 = 0,268 \text{ MPa}$	$P_{bh}^{DM} = -0,226 + 0,932P_{bh}$ $r = 0,964; p < 0,0000$ $S_0 = 0,347 \text{ MPa}$
Magovskoye deposit	$C_{1t} - D_3fm$ $n = 123$	$P_{bh}^{MM} = 1,063 + 0,890P_{bh}$ $r = 0,943; p < 0,0000$ $S_0 = 1,304 \text{ MPa}$	$P_{bh}^{DM} = 5,136 + 0,495P_{bh}$ $r = 0,689; p < 0,0000$ $S_0 = 2,172 \text{ MPa}$
	C_{2b} $n = 33$	$P_{bh}^{MM} = 1,893 + 0,640P_{bh}$ $r = 0,799; p < 0,0000$ $S_0 = 0,681 \text{ MPa}$	$P_{bh}^{DM} = 2,488 + 0,901P_{bh}$ $r = 0,640; p < 0,0000$ $S_0 = 1,532 \text{ MPa}$
Gagarinskoye deposit	$C_{1t} - D_3fm$ $n = 288$	$P_{bh}^{MM} = 3,182 + 0,570P_{bh}$ $r = 0,754; p < 0,0000$ $S_0 = 1,568 \text{ MPa}$	$P_{bh}^{DM} = 3,635 + 0,598P_{bh}$ $r = 0,724; p < 0,0000$ $S_0 = 1,791 \text{ MPa}$
	C_{2b} $n = 193$	$P_{bh}^{MM} = 0,928 + 0,837P_{bh}$ $r = 0,901; p < 0,0000$ $S_0 = 1,053 \text{ MPa}$	$P_{bh}^{DM} = 2,691 + 0,534P_{bh}$ $r = 0,653; p < 0,0000$ $S_0 = 1,680 \text{ MPa}$
Shershnevskoye deposit	$C_{1t} - D_3fm$ $n = 152$	$P_{bh}^{MM} = 2,847 + 0,532P_{bh}$ $r = 0,716; p < 0,0000$ $S_0 = 0,694 \text{ MPa}$	$P_{bh}^{DM} = 1,618 + 0,445P_{bh}$ $r = 0,389; p < 0,0000$ $S_0 = 1,429 \text{ MPa}$
	C_{1v} $n = 112$	$P_{bh}^{MM} = 1,852 + 0,798P_{bh}$ $r = 0,893; p < 0,0000$ $S_0 = 1,344 \text{ MPa}$	$P_{bh}^{DM} = 3,079 + 0,599P_{bh}$ $r = 0,594; p < 0,0000$ $S_0 = 2,716 \text{ MPa}$
Sukharev deposit	$C_{1t} - D_3fm$ $n = 50$	$P_{bh}^{MM} = 0,535 + 0,948P_{bh}$ $r = 0,976; p < 0,0000$ $S_0 = 0,886 \text{ MPa}$	$P_{bh}^{DM} = 0,004 + 0,925P_{bh}$ $r = 0,947; p < 0,0000$ $S_0 = 1,240 \text{ MPa}$
	C_{1v} $n = 61$	$P_{bh}^{MM} = 0,917 + 0,917P_{bh}$ $r = 0,957; p < 0,0000$ $S_0 = 0,651 \text{ MPa}$	$P_{bh}^{DM} = 1,495 + 0,800P_{bh}$ $r = 0,737; p < 0,0000$ $S_0 = 1,728 \text{ MPa}$
	C_{2b} $n = 15$	$P_{bh}^{MM} = 2,013 + 0,802P_{bh}$ $r = 0,895; p < 0,0000$ $S_0 = 1,048 \text{ MPa}$	$P_{bh}^{DM} = 6,005 + 0,305P_{bh}$ $r = 0,358; p < 0,0000$ $S_0 = 2,088 \text{ MPa}$



(a) BHFP calculated by the technique based on multivariate models



(b) BHFP calculated by model based on calculating the gas-liquid mixture density

Figure 1: Field of correlation (whole sample) of the actual BHFP

performance indicators have a complex effect on the BHFP level. The mutual correlations among the indicators vary in a broad range from $-0,87$ to $0,83$. As found out by the correlation analysis one and the same performance indicator may have a different effect on the BHFP level in various conditions, i.e., the input parameters have a differentiated effect on the calculated quantity. For example, coefficient r between P_{bh} and P_{an} changed from -0.08 to 0.51 at various levels of investigation; sometimes, the relation among these parameters is negative, sometimes – positive and statistically significant. On the whole, the identical direction of affecting P_{bh} is followed by such parameters, as H_{dfl} , H_{owc} , H_{pd} , and P_r , whereas the others produce various effects in terms of both, direction and force.

All of this shows that the BHFP is affected by the performance indicators both, together and individually. Thus the correlation analysis has allowed finding out that none of the performance indicators makes it possible to reliably predict BHFP levels. The BHFP formation during well operation follows very complicated and non-stationary laws, which is why the analytical solution making it possible to determine the BHFP in a firm and reliable manner should be considered an extremely complex task to solve.

As mentioned earlier, when analytical techniques do not apply due to the low accuracy of their results, it seems expedient to use statistical (probabilistic statistical) techniques). For this reason, it is not analytical solutions but multivariate statistical models that are proposed in this paper as the techniques for determining the BHFP level during well operation. Therefore, the rest of the investigation is dedicated to constructing multivariate statistical models.

It should be noted that these models were built using the original approach, i.e., by a pre-ordered distribution. Not only does this approach allow deriving the model, which allows reliably determining the BHFP as the target parameter, but it also allows distinguishing all the regularities of its formation in various conditions.

The conclusions derived from analyzing the resulting models are exposed below.

Models A. Procedure based on multivariate models

Models B. Procedure based on calculating gas-liquid mixture density

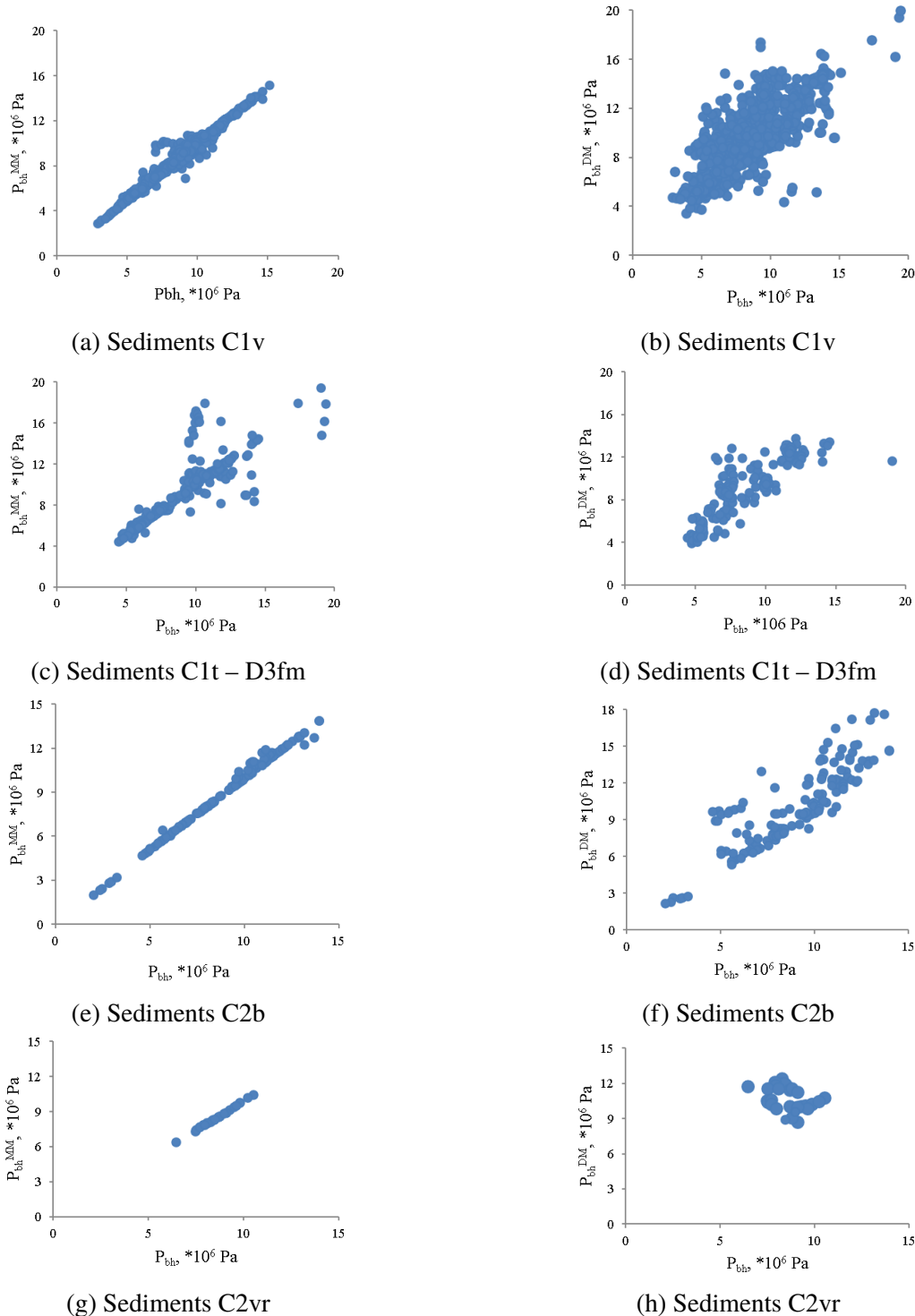


Figure 2: Actual-to-calculated BHFP correlation fields by the example of the Un’vinskoye deposit occurrences

Specifically, the original approach consisting in using a pre-ordered distribution allowed making

a detailed analysis of the frequency with which the input parameters were included in the resulting

equations of regression. As found out by the analysis, the sole indicator with the prevalent effect on the BHFP level could not be distinguished for any of the modeling levels. This conclusion reveals the complexity of the law, according to which the BHFP is formed during well operation, the existence of an integrated effect of the input parameters on the BHFP, and the expediency of using these parameters by applying the built multivariate models of all the levels.

Another characteristic feature of this paper is that the models were constructed for several levels, with different degrees of differentiating modeled objects. The chosen approach is evaluated for correctness by analyzing the values of multiple correlation coefficient R , one of the major indicators characterizing the functionality of a derived model. The model for level one has $R = 0,763$. The average R for the models of levels two, three, and four is 0,790, 0,801, and 0,822, respectively. That is, the model of each subsequent level has a higher degree of functionality.

To consider the established complex effect of the input parameters on the predicted BHFP level, it is proposed to use all of the modeling levels in an integrated manner, by constructing a general multivariate model. This model has very high performance capabilities, for example, multiple correlation coefficient R of 0,941. Thus using all of the distinguished levels of multivariate mathematical modeling has allowed deriving a functional multivariate model the application of which as the mathematical basis of the respective technique will allow determining the BHFP during well operation with a high degree of certainty.

The certainty of determining the BHFP by the technique based on applying the elaborated multilevel multivariate mathematical models was evaluated by a special detailed analysis.

It included comparing the actual (measured) BHFP with the BHFP calculated by the elaborated technique, and the BHFP calculated by the technique currently applied in the region and based on determining the density of the gas-liquid mixture in the wellbore. To derive the most valid conclusions, we made the comparison at different differentiation levels and by means of various tools,

including visual analysis of the correlation fields, and derivation and analysis of the equations of regression among the actual BHFP and the BHFP calculated by the two techniques.

The comparative visual analysis of the fields of correlation between the actual and the calculated BHFP allows concluding that the elaborated technique based on applying multilevel multivariate models is characterized by essentially higher accuracy of calculation: the correlation field for the elaborated technique has a much tighter and even shape, its points group around the line with a slope close to one. The derived regularity is typical of both, the correlation fields drawn for the sample in general and for the fields drawn individually for the occurrences from separate deposits.

The visual analysis of the correlation fields has allowed evaluating, by comparison, the certainty of determining the BHFP by the two techniques at the level of quality. The quantitative comparison of these fields has made it necessary to derive the equations of regression between the actual and the calculated BHFP for the two levels of investigation as well, i.e., for the sample in general (level one) and, individually, for occurrences in separate deposits (level two). The sign of the higher certainty of a particular technique is that the value of the free term in the equation of regression is close to zero and the angular ratio close to one. In addition to the regression equation parameters, their statistical characteristics have been calculated as well, including the correlation coefficient, significance level, and standard calculation error. As shown by analyzing the statistical characteristics of the exploitation targets from all of the deposits, not only the values of correlation coefficient r differ in all of the four cases but the equation of regression themselves. In all of the cases, the coefficients at PBHFP found by the multilevel technique are higher than those found by the density-based technique. The standard errors calculated according to the multivariate models for all of the exploitation targets are much lower than the standard errors calculated by the density-based technique. As shown by comparing the average reduced characteristics values against the t criterion, statistical differences

are found in each case. The enumerated facts convincingly prove that the BHFP determination by the elaborated technique based on applying multivariate multilevel (statistical) models has a higher degree of certainty.

6. Conclusion

Thus this article validates the technique of determining the BHFP during oil-production well operation by means of the constructed multivariate multilevel models.

The models were constructed proceeding from the significant accumulated experience in parallel depth and estuarine measurements conducted when servicing the commercial wells in the Perm krai.

The constructed models have high statistical capability characteristics. A distinct feature of the models is that the parameters they use as the sole original data are easy to determine in practice. This fact should be considered the main strength of the developed technique as compared with its multiple density-based counterparts.

The high capability of the constructed multivariate BHFP models is mainly stipulated by the original approach to making them that consists in creating a model according to a sequence of pre-ordered original data.

Not only has the construction of multivariate mathematical models allowed determining the BHFP in practice but it has also allowed identifying the regularities of its formation and in-operation behavior individually for each considered exploitation target.

The BHFP determination technique based on the developed multivariate models is much more functional than its conventional density-based counterparts.

It is worth noting individually that the new technique should not be considered an alternative to density-based methods. The joint application of these techniques is supposed to ensure a reliable BHFP control during production well operation.

Acknowledgments

We thank OOO LUKOIL-PERM and Sergey Cherepanov and Irina Chernykh in person for

assistance in the pursuance of the research

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

7. References

- [1] L. Nazarova and E. Nechaeva, "The analysis of influence of bottomhole pressure decrease under saturation pressure on the oil recovery," *Oil Industry*, no. 1, pp. 83–85, 2014.
- [2] Y. Kashnikov and S. Yakimov, "Geomechanical and hydrodynamic estimation of the bottom-hole pressure influence on the well performance," *Oil and Gas Business*, vol. 11, pp. 111–115, 2019.
- [3] T. Bikmukhametov and J. Jäschke, "Oil Production Monitoring using Gradient Boosting Machine Learning Algorithm," *IFAC-Papers online*, vol. 52, no. 1, pp. 514–519, 2019.
- [4] S. Natarajan, K. Ghosh, and R. Srinivasan, "Collaborative Multi - Agent based Process Monitoring System for Offshore Oil and Gas Production. Comput," *Computed Aided Chemical Engineering*, vol. 27, pp. 1227–1232, 2009.
- [5] D. Martyushev, "Determination of the rational bottomhole pressure of producing wells in the development of carbonate reservoirs," *Drilling and oil*, no. 11, pp. 22–24, 2014.
- [6] D. Martyushev and V. Mordvinov, "Change in the flow rate of wells in an oil and gas condensate field with a decrease in reservoir and bottomhole pressures," *Oil Industry*, no. 167–69, 2014.
- [7] A. Cudas, B. Ordoñez, and U. Moreno, "Sucker-Rod Pumping System Fault Detection and Isolation Method Using Bottom Hole Pressure Measurement," *IFAC Proceedings Volumes*, vol. 42, no. 8, pp. 1031–1036, 2009.
- [8] V. Mordvinov, A. Lekomtsev, and D. Martyushev, "Determination of pressure at the intake of electric centrifugal pumps when pumping out low-foam carbonated oil," *Oil industry*, no. 61–63, 2014.
- [9] M. Ali Ahmadi, M. Galedarzadeh, and S. Reza-Shadizadeh, "Low parameter model to monitor bottom hole pressure in vertical multiphase flow in oil production wells," *Petroleum*, vol. 2, no. 3, pp. 258–266, 2016.
- [10] E. Levitina, "Influence of density gas liquid phase on parameters of well pressure," *Oil and Gas Stud*, no. 1, pp. 35–40, 2010.
- [11] H. Yang, J. Li, G. Liu, X. Xing, H. Jiang, and C. Wang, "The effect of interfacial mass transfer of slip-rising gas bubbles on two-phase flow in the vertical

- wellbore/pipeline,” *International Journal of heat and mass transfer*, vol. 150, no. 119326, 2020.
- [12] A. Hasan and C. Kabir, “A Study of Multiphase Flow Behavior in Vertical Wells,” *SPE Production Engineering*, vol. 3, no. 2, pp. 263–272, 1988.
- [13] J. Brill and H. Mukherjee, *Multiphase Flow in Wells*. Richardson: SPE, 1999.
- [14] A. Lekomtsev, E. Zhelanov, and I. Chernykh, “Statistical approach to the evaluation of bottomhole pressure in producing wells,” *Oil and Gas Business*, vol. 10, pp. 98–101, 2016.
- [15] A. Lekomtsev and D. Martyushev, “Comparative analysis of methods for determining BHP during well test,” *Oil Industry*, vol. 6, 2014.
- [16] S. Bikbulatov and A. Pashali, “Analisis and selection of methods for calculating the pressure gradient in the wellbore,” *Oil and Gas Business*, vol. 21, no. 2, pp. 1–12, 2005.
- [17] I. Chernykh, V. Galkin, and I. Ponomareva, “Comparative analysis of the methods for defining bottom hole pressure at well operation of Shershnevsky field,” *A Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, vol. 328, no. 8, pp. 41–47, 2017.
- [18] V. Galkin, I. Ponomareva, I. Chernykh, and E. Filippov, “Methodology for estimating downhole pressure using multivariate model,” *Oil Industry*, vol. 1, pp. 40–43, 2019.
- [19] D. Martyushev and Y. Slushkina A, “Assessment of informative value in determination of reservoir filtration parameters based on interpretation of pressure stabilization curves,” *Bulletin of the Tomsk Politechnic University*, vol. 330, no. 9, pp. 26–32, 2019.
- [20] M. Nait Amar, N. Zeraibi, and K. Redouane, “Bottom hole pressure estimation using hybridization neural networks and grey wolves optimization,” *Petroleum*, vol. 4, no. 4, pp. 419–429, 2018.
- [21] W. Chen, Q. Di, F. Ye, J. Zhang, and W. Wang, “Flowing bottomhole pressure prediction for gas wells based on support vector machine and random samples selection,” *International Journal of Hydrogen Energy*, vol. 42, no. 29, pp. 18 333–18 342, 2017.
- [22] A. Sánchez-Fernández, F. Baldán, G. Sainz-Palmero, J. Benítez, and M. Fuente, “Fault detection based on time series modeling and multivariate statistical process control,” *Chemometrics and Intelligent Laboratory Systems*, vol. 118, pp. 57–69, 2018.
- [23] C. Kumar Singha, A. Kumar, S. Shashtrib, A. Kumar, P. Kumar, and J. Mallick, “Multivariate statistical analysis and geochemical modeling for geochemical assessment of groundwater of Delhi, India,” *Journal of Geochemical Exploration*, vol. 175, pp. 59–71, 2017.
- [24] D. Cecconet, S. Bolognesi, S. Daneshgar, A. Callegari, and A. Capodaglio, “Improved process understanding and optimization by multivariate statistical analysis of Microbial Fuel Cells operation,” *International Journal of Hydrogen Energy*, vol. 43, no. 34, pp. 16 719–16 727, 2018.
- [25] J. Hua, J. Li, M. Ouyang, L. Lu, and L. Xu, “Proton exchange membrane fuel cell system diagnosis based on the multivariate statistical method,” *International Journal of Hydrogen Energy*, vol. 36, no. 16, pp. 9896–9905, 2011.
- [26] S. Barkovskiy, V. Zakharov, A. Lukashov, and A. Nurutdinova, *Multidimensional Data Analysis by Applied Statistics Techniques*. Kazan: KSTU, 2010.
- [27] R. Kissell and J. Poserina, *Optimal Sports Math, Statistics, and Fantasy*. Academic Press, 2017.
- [28] E. Wentzel, *Operations Research*. Moscow: Mir Publishers, 1983.

Boron Carbide Composites produced by Self-Propagating High-Temperature Synthesis

Roza G. Abdulkarimova ^{*,a} , Aizhan J. Seidualiyeva ^a , Aisulu N. Batkal ^b ,
Sanat Tolendiuly ^b , Sergey M. Fomenko ^b 

^a*Al-Farabi Kazakh National University. Almaty, Kazakhstan*

^b*Institute of Combustion Problems. Almaty, Kazakhstan*

Abstract.- This work is a study of the conditions for obtaining boron carbide, magnesia, and aluminum oxide composites by self-propagating high temperatures synthesis (SHS). The substances used to synthesize the composites were boron oxide, magnesia, and aluminum oxide. The SHS with the reduction stage was conducted as the following aggregate reaction: $2 B_2O_3 + 6 Mg + xC \longrightarrow ByCx + 6 MgO$, $2 B_2O_3 + 4 Al + xC = 2 Al_2O_3 + ByCx$. The metallothermic SHS using boron oxide seems more attractive than the current carbide production methods characterized by long-time and multistage physico-chemical processes requiring increased material, energy, and financial costs. The thermodynamic calculations of the phase composition of the products and the adiabatic combustion temperature for systems $B_2O_3 - Mg - C$, $B_2O_3 - Al - C$ were made in the FastStage program. The calculations and tests allowed identifying the optimal conditions of SHS. The products of SHS were examined by X-ray phase analysis and the SEM method. The compositions of the $B_2O_3 - Mg - C$ system obtained by SHS were represented mainly by boron carbide, magnesia, magnesium borate, and $Mg_3B_2O_6$. The products of the $B_2O_3 - Al - C$ system obtained by SHS contained boron carbide, alumina, and aluminum boride.

Keywords: composite; boron carbide; self-propagating high-temperature synthesis; thermal dynamic analysis.

Materiales Compuestos de Boro Producidos por Síntesis Autopropagante a Temperaturas Elevadas

Resumen.- Este trabajo es un estudio de las condiciones para la obtención de compuestos de carburo de boro, magnesia y óxido de aluminio por medio de síntesis autopropagante a temperaturas elevadas (SAPTE). Las sustancias utilizadas para sintetizar los compuestos incluyeron óxido de boro, magnesia y óxido de aluminio. Se llevó a cabo la SAPTE con la fase de reducción como la siguiente reacción agregada: $2 B_2O_3 + 6 Mg + xC \longrightarrow ByCx + 6 MgO$, $2 B_2O_3 + 4 Al + xC = 2 Al_2O_3 + ByCx$. La SAPTE térmica metálica que utiliza óxido de boro parece más atractiva que los métodos actuales de producción de carburo que se caracterizan por procesos fisicoquímicos de larga duración y múltiples etapas que requieren mayores costos de materiales, energía y financieros. Los cálculos termodinámicos de la composición de fase de los productos y la temperatura de combustión adiabática para los sistemas $B_2O_3 - Mg - C$, $B_2O_3 - Al - C$ se realizaron en el programa FastStage. Los cálculos y pruebas permitieron identificar las condiciones óptimas de SAPTE. Los productos de SAPTE se examinaron mediante análisis de fase de rayos X y el método SEM. Las composiciones del sistema $B_2O_3 - Mg - C$ obtenido por SAPTE estuvieron representadas principalmente por carburo de boro, magnesia, borato de magnesio y $Mg_3B_2O_6$. Los productos del sistema $B_2O_3 - Al - C$ obtenido por SAPTE contenían carburo de boro, alúmina y boruro de aluminio.

Palabras clave: material compuesto; carburo de boro; síntesis autopropagante a temperaturas elevadas; análisis termodinámico.

Received: January 18, 2021.

Accepted: April 01, 2021.

1. Introduction

Boron carbide is the basis of high-refractory materials with a unique combination of a lot of properties, such as high thermal conduction, high electrical resistance, hardness, chemical inertia, high thermal shock resistance, high corrosion

* Correspondence author:

e-mail:abdulkarimovaroza@mail.ru (R. Abdulkarimova)

resistance, and others [1, 2]. The main method of making ceramic B_4C items is hot pressing at about 2300 K [3, 4]. Another popular method of making them is the sintering in flash plasma [5]. The synthesis of amorphous boron and carbon is conducted in vacuum or inert atmosphere at temperatures above 1500 °C. The synthesis is a slow process despite the reaction's strong exothermal effect due to the difficulties with the hard-phase diffusion of the reacting particles through the layers of synthesized B_4C . The high price of pure boron used as a reactant makes the process economically unviable but for cases when the object of synthesis is highly pure B_4C is intended for experimentation purposes. The method is indispensable in synthesizing ^{10}B -enriched B_4C for nuclear engineering purposes [6].

However, despite its high hardness and wear resistance, B_4C has low strength and cracking resistance. To improve the strength of ceramics based on B_4C , various kinds of additives are used in the course of sintering that make the material more ductile and resistant to cracking [3]. For this reason, a lot of attention is currently being paid to the technology of making composites based on borides, carbides, and nitrides together with more ductile materials used as binders. Self-propagating high-temperature synthesis provides ample opportunities for using magnesium or aluminum as reducing agents in synthesizing materials by combustion. Thus the aluminum oxide or magnesia contained in synthesized composites can function as high-temperature binders and fillers [7, 8, 9]. The SHS method is used at high temperatures in the condensed phase and fully uses the chemical energy of the system but not of external sources of heat. In the course of SHS strongly exothermal chemical interactions are observed among the reactants, which are attended by the release of heat in large amounts. High thermal effect values lead to high temperatures, whereas the maximal temperature achieved during SHS is called the combustion temperature. It plays a key role in the development of SHS and the formation of the final products of combustion. This is why, a lot of attention is being paid to

both, experimental temperature measurements and theoretical temperature calculation [10, 11, 12, 13, 14, 15].

Together with the high speed of the process, SHS allows achieving increased output at minimal costs. This synthesis is energy-saving, self-controlled, and ecofriendly. In addition, SHS allows adjusting the size of the particles and structure and synthesized composites. In these processes, the wasteless chemical synthesis of high-refractory boron, carbide, and other compounds occurs. These compounds are the basis of a broad range of modern nonorganic materials usable in extreme conditions [16, 17, 18, 19, 20, 21, 22].

An important role in obtaining materials by SHS is played by preliminary mechanochemical activation (MCA) that allows attaining a high level of particle dispersion, changing the structure and energy intensity of the material, and, therefore ensures its high reactivity [17, 18, 19, 20, 21, 22].

The pulping of components reduces the kinetic obstacles to chemical reactions in heterogeneous media. The MCA of SHS is characterized by high heat release levels and speed, low initial reaction temperatures, and high chemical and structural heterogeneity of synthesized products [17]. The mechanical activation of reactants before SHS is a very important stage. It can make heterogeneous processes far more intensive and increases the chemical activity of processed particles by increasing their defect rate and/or expanding their reactive surface through their reduction in size. Thus MCA prepares the particles for response activity. Mechanical activation may affect both, the speed and conditions of combustion wave front propagation (macrokinetic effect) and the shape and size of the crystallites and structural porosity (structural effect) [10].

The reason, why the inclusion of SHS and MCA in the process flow of creating new kinds of material is of interest, is that the accumulation and release of energy is observed in each of them and exothermal effects advance the synthesis. These processes allow saving energy, which is especially important in making materials with a high melting point. Their synthesis by conventional technologies, not by SHS, requires using a lot of energy and

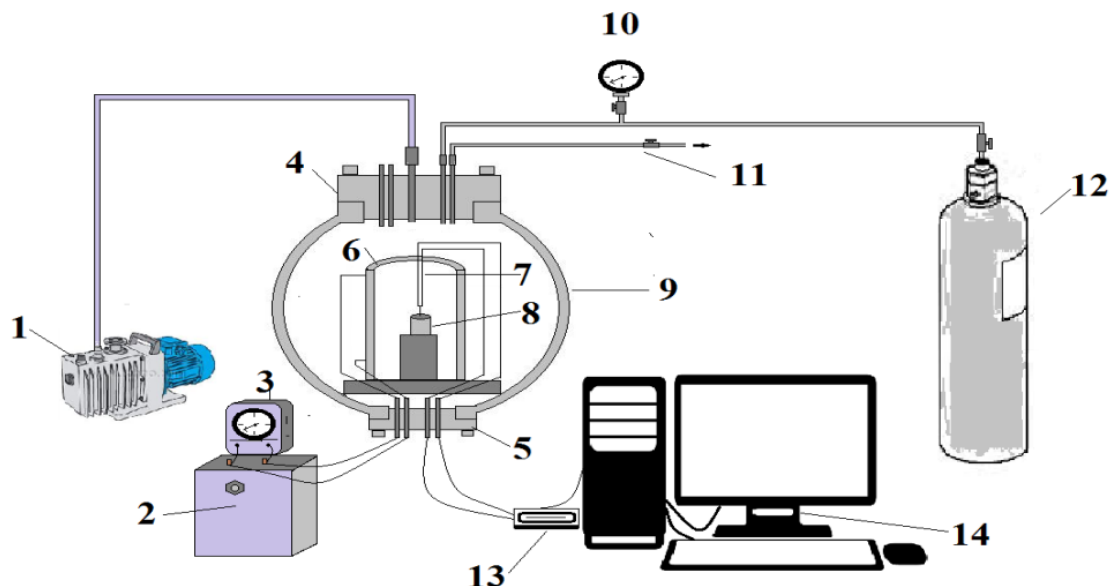


Figure 1: High-pressure reactor: 1 is the vacuum pump, 2 is the transformer, 3 is the amperemeter, 4 is the upper reactor head; 5 is the bottom reactor head; 6 is the tubular heating furnace, 7 is the thermocouple, 8 is the specimen, 9 is the reactor body, 10 is the pressure gage, 11 are the inlet and outlet gate valves, 12 is the argon bulb, 13 is the unit of the LTR-U-1 data collection system, 14 is the computer

sophisticated equipment; first of all, this bears on nitride, boride, and carbide ceramic materials [8, 10, 11].

2. Materials and methods

The specimens were made from a charged material with magnesium powder (100-200 μm particles), boron oxide (50-100 μm particles), and graphite (0,5-1,0 μm particles) with a carbon content of 60-99,9 %. The specimens were molded on a Carver laboratory extruder at an effort of five tons and their sizes were diameter $d = 20$ mm, and height $h = 25$ mm. The specimens were placed in a SNOL electrical laboratory furnace, and the synthesis was conducted at 700 °C. The combustion temperature was measured using an Ircan Ultrimax Plus UX10P pyrometric temperature gage with a working range from 600 to 3000 °C. The respective error in measuring temperatures up to 1500 °C and from 1500 to 2000 °C was $\pm 0,5$ and ± 1 % respectively.

In addition, the SHS of the specimens was conducted in an inert medium in a constant-

pressure reactor. The prepared specimens were placed there for synthesizing high-refractory boron-containing composites at an argon pressure of 1 MPa. The tubular baking furnace placed in the reactor serves to expand the concentration limits of synthesis and allows heating the test specimen up to 1000 °C. The layout of the pressure of the reactor is shown in Figure 1.

After the combustion process was initiated, the specimen temperature was recorded with the help of a computer and special-purpose software that read the data in real-time mode from BP5/20 tungsten rhenium thermocouples with a brazing seam thickness of 200 μm .

The source components were mechanically activated in a high-energy planetary centrifugal mill Pulverizette 5. The back plate rotation speed was 380 rpm. The relation of the ground powder mass to the grinding balls was 1:2.

The X-ray phase analysis of the specimens was made on a DRON-4M diffractometer on exposure to cobalt Ka radiation in range $2\theta = 10^\circ - 70^\circ$. The morphology of the specimens was analyzed by scanning electron microscopy (SEM) on a US-

made microscope FEI QUANTA 3D 200i.

3. Results and discussions

An important parameter of SHS systems is combustion temperature. It eventually affects the quality of synthesized products and must be high enough to ensure an intensive heterogeneous reaction. The adiabatic temperature of combustion should be higher than the melting point of at least one of the components [22, 23].

Thermodynamic analysis allows calculating the adiabatic combustion temperature of exothermal mixtures (T_{ad}) and the phase composition of the equilibrium products of SHS. The temperature and the equilibrium composition of SHS products are calculated with the help of software algorithms of searching for systems minimal free energy [15, 23].

The calculation of adiabatic temperatures and SHS products proper for the B_2O_3-Al-C system (where C is for graphite) was made in FactSage [15].

For the dependence of the adiabatic combustion temperature on the content of aluminum in the B_2O_3-Mg-C system at $T_0 = 700^\circ C$ as is shown in Figure 1. It is seen from Figure 1 that at a magnesium content of 20-22 % wt. in the charged material the maximal SHS temperature is $1767^\circ C$. At a magnesium content in the system of 22-40 % wt; this temperature does not change and is equal to $1764,4^\circ C$.

The system's theoretic combustion temperature of $1767^\circ C$ allows concluding that SHS is possible in this system. The phase formation and structuring of material occur exactly at the maximal temperature attained in SHS systems. The thermodynamic calculation provides an idea of the actual combustion temperature [13, 14, 15].

For the dependence of the potential phase composition of SHS products on the content of magnesium in the B_2O_3-Mg-C at $T_0 = 700^\circ C$ see Figure 2.

It follows from the diagram and the calculations in the FactStage software that the content of boron carbide in the synthesized products increases with an increase in the amount of magnesium and reaches the maximum of 18 % wt. Magnesium

oxide is also the main phase in SHS products and has a melting point of $2800^\circ C$. An increase in the content of magnesium in the source components results in the high amount of magnesium oxide in SHS products. In addition, $Mg_2B_2O_6$ is formed whereas different magnesium borate varieties may form at $800-850^\circ C$ [6]. It should be noted that small amounts of CO form at a magnesium content of 26 to 40 % wt. in the charged material; however, there is still some amount of unreacted carbon that may enter into reaction as follows:

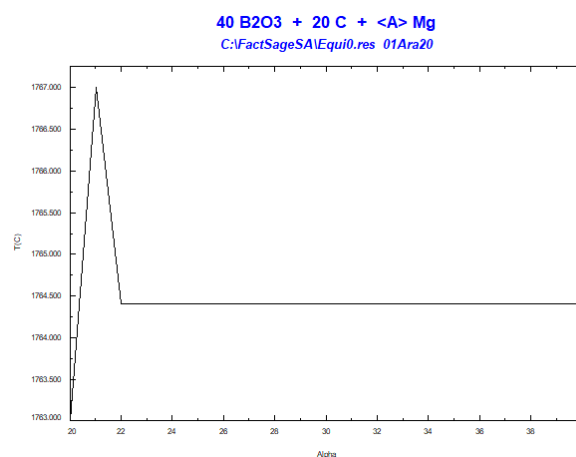
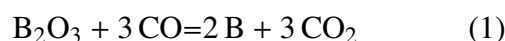


Figure 2: Dependence of the adiabatic combustion temperature on the content of magnesium in the B_2O_3-Mg-C system at $T_0 = 700^\circ C$

The SHS tests were performed in the air at the initial room temperature, in the muffle roaster at $T_0 = 700^\circ C$, and in argon in the pressure reactor. The room-temperature SHS in the B_2O_3-Mg-C and B_2O_3-Al-C systems did not provide any results. It is known from Merzhanov's empirical criterion that the adiabatic temperature (T_{ad}) must be high enough to ensure a reaction in combustion mode without preheating [11].

The optimal proportion of the source components, defined by experiment, taking into account the calculated tentative equilibrium composition of the specimens and the adiabatic temperature, was 25 % of Mg to 30 % of B_2O_3 to 45 % of C. When the SHS test was conducted in the muffle roaster at

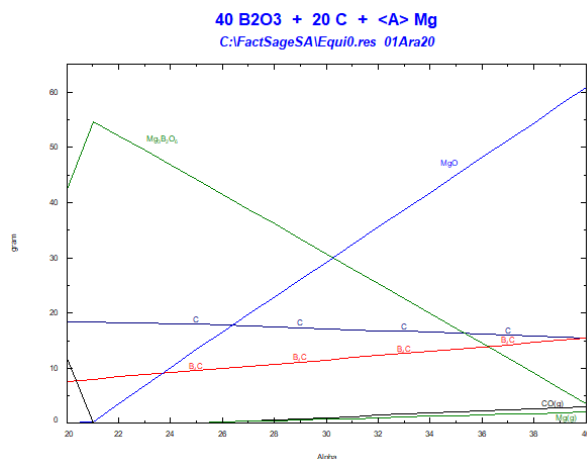


Figure 3: Dependence of the possible phase composition of SHS products on the content of magnesium in the B_2O_3 –Mg–C system at $T_0 = 700\text{ °C}$

various initial heating temperatures, the specimens would not spontaneously ignite at $T < 700\text{ °C}$, whereas the magnesium-induced combustion made the specimens flowed, probably, because of high temperatures. The optimal temperature for starting SHS in the muffler roaster is 700 °C at which the synthesis was performed in its entirety while maintaining the initial shape of the specimens.

As seen from Figure 4 with the thermogram of the SHS of the Mg– B_2O_3 –C system, the maximal combustion temperature is 1600 °C .

According to thermodynamic calculations, the adiabatic temperature of combustion for Mg– B_2O_3 –C is 1760 °C . Actually, the combustion temperatures measured by a specific method are usually lower than the calculated temperatures. This is due to the fact that, in actual conditions, SHS specimens exhibit an intense heat and mass exchange with the environment. In addition, an SHS system often contains synthesized products with phase transitions to Tad. These transitions usually are usually attended by energy costs, which is a thermally destabilizing factor for SHS systems.

The synthesis by combustion usually begins with the melting of boron oxide at 450 °C , followed by the melting of magnesium at 650 °C , which corresponds to the phase transition plateau in the

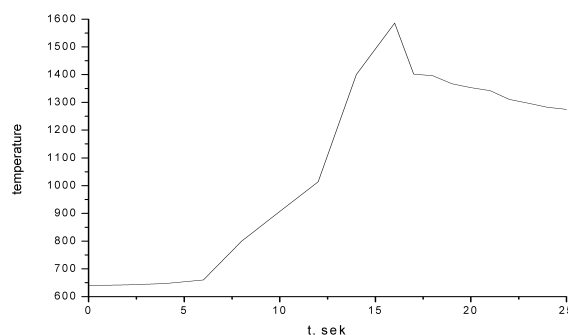
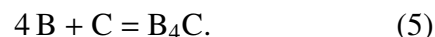
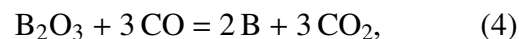
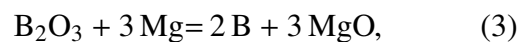


Figure 4: Thermogram of the Mg– B_2O_3 –C system

Figure. The rising temperature is related to the exothermal reactions of SHS (Figure 4). The possible stages of the reactions involved in the formation of B_4C –MgO are recorded as



It is known that the MA of charged materials before SHS in the planetary mill significantly reduces the initial temperature of the exothermal interactions among the components of the mixture, cuts the period of synthesizing the final product, and leads to a more complete extent of reaction. Mechanical activation allows accelerating chemical reactions by improving the mixture reactivity [17, 22]. To study the influence of the preliminary MA of the source components on the phase composition of SHS products, the mechanical activation was carried out for five, seven, and ten minutes. It was established that it is inexpedient to continue the MA for more than ten minutes because the extended MA results in the loss of the system’s accumulated energy due to the agglomeration of the powder mixture particles.

For the percentages of the products of the SHS of the Mg– B_2O_3 –C system in argon in the reactor at the Mg– B_2O_3 –C ratio of 25 to 30 to 45 % see Table 1.

It is seen from Table 1 that the preliminary MA increases the content of boron carbide in the reaction products. The formed boron carbide,

Table 1: Phase composition of the products of the SHS system with the respective content of Mg, B₂O₃, and C of 25, 30, and 45 % at T₀ = 700 °C

MCA duration (min)	Content of SHS products (%)			
	MgO	Mg ₃ B ₂ O ₆	B ₄ C	Mg ₂ C ₃
10	51,9	9,3	37,3	1,5
7	72	7	17,4	3,5
5	75,4	7,8	13,5	3,3
-	69,9	9,5	10,6	10

magnesium oxide, and magnesium carbide are high-temperature bondings.

To compare the conditions of synthesizing boron carbide ceramic materials by combustion, an SHS was carried out in an inert medium in a high-pressure plant (Figure 1). In the earlier study of the influence of the argon pressure on the parameters of SHS the argon pressure in the reactor was varied from 0,5 to 3 MPa. The tests showed that the optimal pressure for SHS was 1 MPa [24, 25]. It is obvious that a rise in the reactor pressure increases the heat losses during SHS and reduces the temperature in the system, which affects the final product yield.

For the results of the XRF of the products of the SHS of the Mg–B₂O₃–C system in argon in the reactor at a pressure of 1 MPa see Table 2. The furnace was heated up to 850 °C. The components ratio in the system was 25 to 30 to 45 %.

Table 2: Phase composition of the products of the SHS system with the respective content of Mg, B₂O₃, and C of 25, 30, and 45 %

MCA duration (min)	Content of SHS products (%)				
	B ₄ C	Mg ₃ B ₂ O ₆	MgO	B ₂ O ₃	C
10	57,2	26,2	16,6	-	-
7	41,1	24,7	34,2	-	-
5	38,8	37,5	19,6	3,4	0,7
-	29,7	39,1	26,1	4,2	0,9

It follows from Table 2 that the synthesis in the inert medium encourages an increase in the output of boron carbide in SHS products. That said, preliminary MA increases the boron carbide content in the reaction products. The optimal duration of preliminary MA is 10 minutes. As a result of the preliminary MA of the charged material, the source boron oxide and graphite in the reaction products remain unreacted, which is

due to an expansion of the unit surface area and an increase in the reactivity of the mixtures.

In addition, the conditions for the SHS of the composites in the B₂O₃–Al–C system were studied. The thermodynamic calculations of the products of the SHS were made for this system.

For the dependence of the adiabatic combustion system on the content of Al in the B₂O₃–Al–C at T₀ = 700 °C see Figure 5.

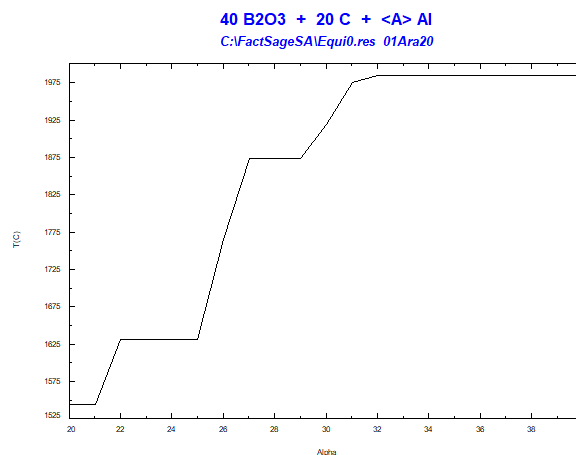


Figure 5: Dependence of the adiabatic combustion system on the content of Al in the B₂O₃–Al–C at T₀ = 700 °C

It is seen from Figure 5 that the maximal SHS temperature at an aluminum content of 32 to 40 % wt. in the charged material is 1980 °C. The adiabatic temperature growth curve exhibits plateaus at the content of aluminum from 22-25 to 27-29 % wt., which can be explained by phase transitions during SHS. The dependence of the phase composition of SHS products on the content of Al in the B₂O₃–Al–C at T₀ = 700 °C is shown in Figure 6.

According to the diagram in Figure 6 and the data of the calculations in FactStage, the boron carbide content in the synthesized products increases with an increase in the content of aluminum and reaches 17 % wt. Aluminum carbide forms at an Al content of more than 31 % wt. in the charged material. An increase in the content of aluminum in the source charged material increases the amount of aluminum oxide in SHS products. It should be noted that, if the Al content in the charged material ranges from

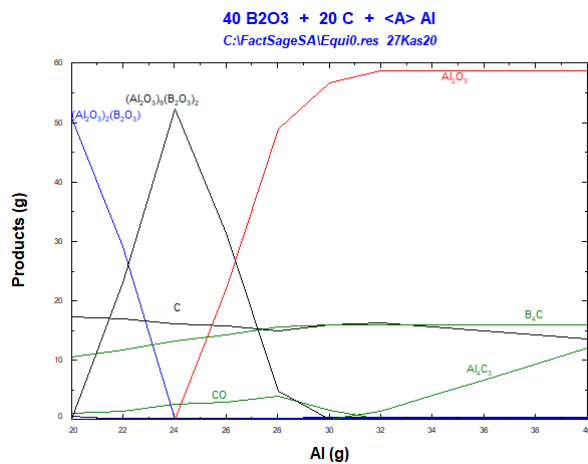


Figure 6: Dependence of the phase composition of SHS products on the content of Al in the B_2O_3-Al-C

20 to 31 % wt., the products of SHS may appear to contain small amounts of carbon dioxide.

It is established that the optimal composition of the source components includes 30 % of B_2O_3 , 30 % of Al, and 40 % of C. The optimal fraction of Al in the charge material is 30 % wt. because more than 35 % of Al in the reaction products remain unreacted and the specimen melts away and does not hold shape. According to the thermodynamic calculation data, the highest boron carbide content is also assumed at a boron carbide content of 30 % wt. The initial baking furnace temperature at which the SHS preserves the specimen shape is $800\text{ }^\circ\text{C}$, which is higher than for the B_2O_3-Mg-C with a more active reducing agent.

Table 3: Phase composition of the SHS products of the B_2O_3-Al-C system at $T_0 = 800\text{ }^\circ\text{C}$

MCA duration (min)	Content of SHS products (%)				
	Al_2O_3	B_4C	Al	$AlBO_5$	AlB_{10}
5	72,6	8,7	-	13,4	8,7
7	84,1	10,3	-	6,6	7,4
10	80,6	13,1	-	5,6	6,3
-	73,7	5,6	2,5	6,2	15,5

For the phase composition of the products of the SHS of B_2O_3-Al-C in the muffle roaster at initial temperature $T_0 = 800\text{ }^\circ\text{C}$ see Table 3. In the course of the SHS, aluminum reduces boron oxide, which results in the formation of boron carbide, alumina, and aluminum boride. In this case, the content

of boron carbide in combustion products increases from using preliminary MA.

For the information on the products of the SHS performed for the $Al + B_2O_3 + C$ system in argon in the reactor at a pressure of 1 MPa see Table 4. During the SHS the baking roaster heated to $850\text{ }^\circ\text{C}$.

Table 4: Combustion products of the B_2O_3-Al-C

MCA duration (min)	Content of SHS products (%)			
	B_4C	Al_2O_3	AlB_{10}	C
10	45,4	42,5	12,1	-
5	44,1	38,7	15,6	0,6
-	37,2	50,4	11,2	1,2

The results from Table 4 allow concluding that there was an increase in the boron content in the combustion synthesis products. The preliminary MA of the charged material significantly increases the boron carbide output.

The formation of the phase and chemical composition as well as of the crystalline structure and microstructure of SHS products are complex interrelated processes. It is very important to study them during SHS for the properties of a material are determined by its structure. The electron microscopic study of the morphology of the SHS products was conducted in an inert medium for systems $Mg-B_2O_3-C$ (Figure 7) and B_2O_3-Al-C (Figure 8).

The morphology of the combustion products is represented by crystalline structures with an average size of 1-10 μm . The contrast among the parts of the image indicates the presence of at least three compounds. The high content of boron and carbon showed by the microanalysis confirms the results of the X-ray phase analysis of the SHS products.

4. Conclusion

The thermodynamic calculation of the supposed phase composition of SHS products and the adiabatic combustion temperature of the B_2O_3-Mg-C and B_2O_3-Al-C systems have been made. The tests have helped determine the optimal conditions for SHS. The influence of the initial temperature on the complete course of SHS has been determined. It has been established that the initial room

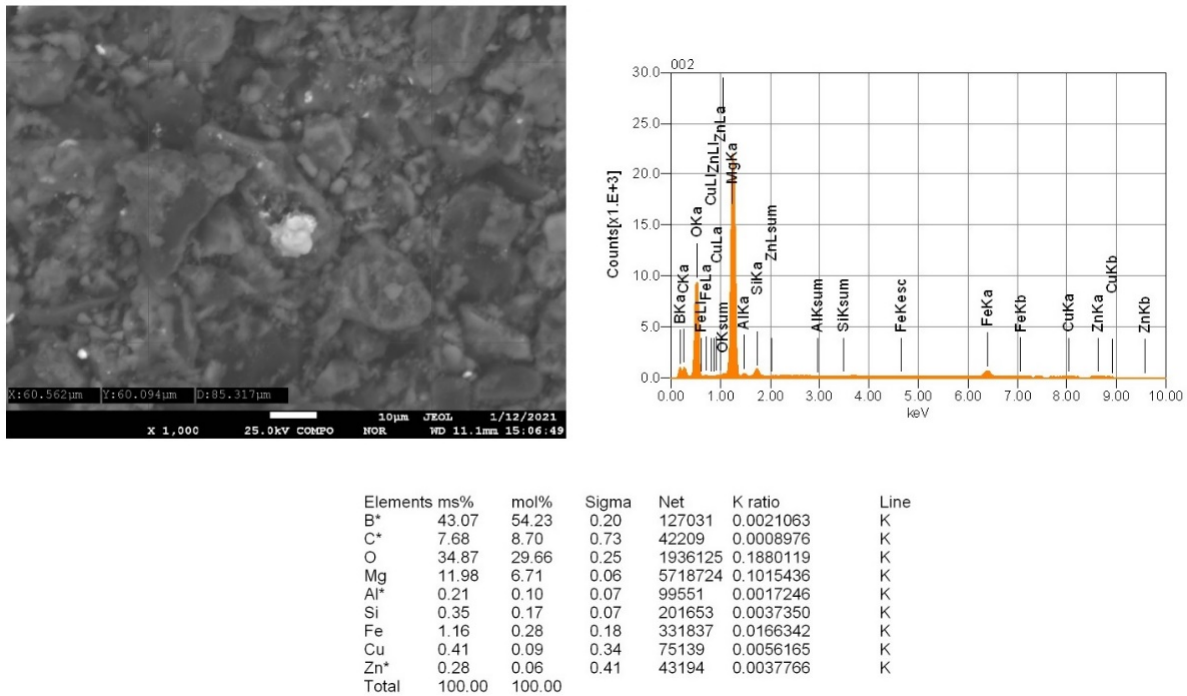


Figure 7: SEM photos of the SHS products of the Mg – B₂O₃ – C system

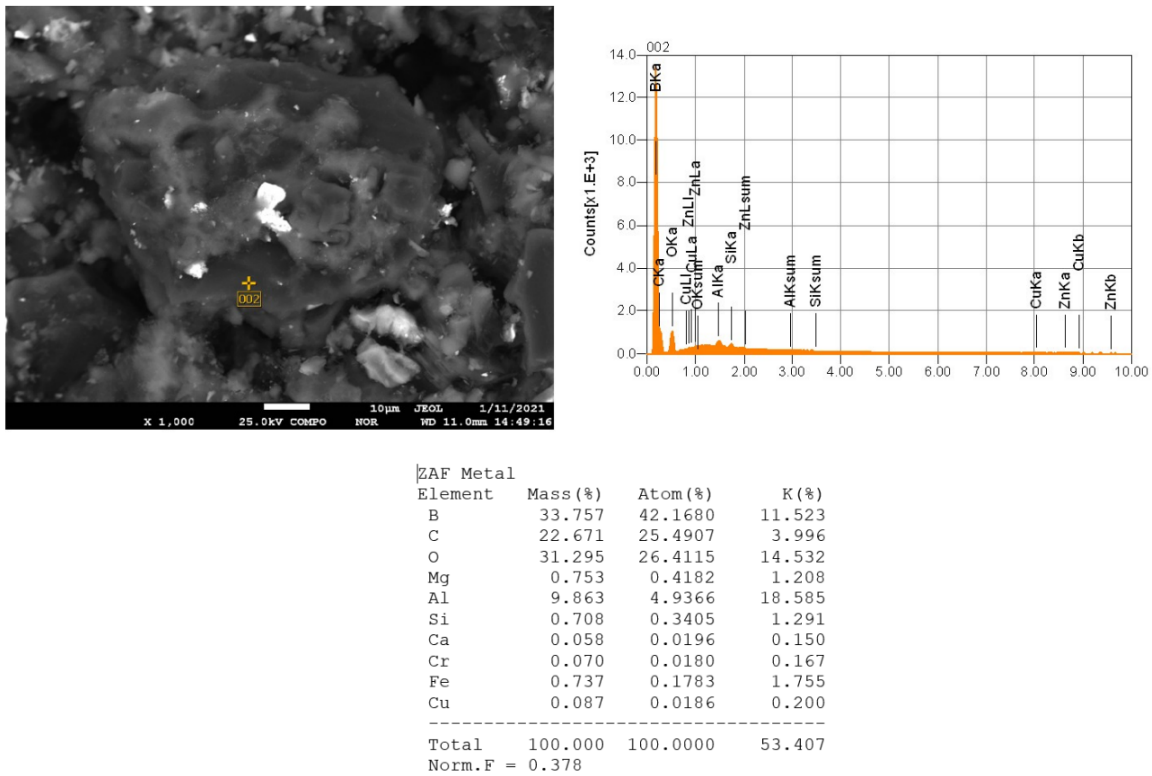


Figure 8: SEM photos of the SHS products of the B₂O₃ – Al – C

temperature makes the synthesis by combustion impossible. When placed to the muffle roaster at

$T = 700$ and $800\text{ }^{\circ}\text{C}$, the $\text{B}_2\text{O}_3\text{-Mg-C}$ and the $\text{B}_2\text{O}_3\text{-Al-C}$ systems self-ignite and retain their initial form after the advancing of the combustion wave. In the case of SHS, the respective specimens also retain their form in the inert medium at a reactor argon pressure of 1 MPa.

In addition, the optimal initial proportions of the SHS products in the systems have been determined: 25 % of Mg to 30 % of B_2O_3 to 45 % of C and 30 % of B_2O_3 to 30 % of Al to 40 % of C, respectively. The influence of the duration of preliminary mechanical activation (MA) of the reactive mixture on the phase composition of ceramic composites has been studied. It has been shown that this activation increases the boron carbide content in the combustion products increases.

According to the X-ray phase analysis of the $\text{B}_2\text{O}_3\text{-Mg-C}$ system, the SHS products contain mostly the phases of boron carbide and MgO and unreduced B_2O_3 reacts with MgO to form $\text{Mg}_3\text{B}_2\text{O}_6$. The composition of the SHS in the $\text{B}_2\text{O}_3\text{-Al-C}$ system is represented by boron carbide, alumina, and aluminum boride. The morphology of carbide impurities in the sintered composites is shown.

Thus it is proven that boron carbide composites with high-refractory compounds, such as alumina and magnesia, can be made by SHS in the muffle roaster at $T = 700$ and $800\text{ }^{\circ}\text{C}$ and in the reactor argon at $P = 1$ MPa. It has been established that to ensure the maximal content of boron carbide in SHS products, it is more preferable to make them with magnesium as a reducing agent, not with aluminum. It has also been established that the preliminary MCA of the source components and the implementation of SHS in the reactor at an argon pressure of 1 MPa allow achieving the highest boron carbide output in the SHS products.

5. References

- [1] P. S. Kislyi and M. A. Kuznetsova, *Boron carbide*. Kyiv: Naukova Dumka, 1988.
- [2] R. A. Andrievskii, "Synthesis, Structure, and Properties of Microsize and Nanosize Boron Carbide [Mikro-i nanorazmernyy karbid bora: sintez, struktura i svoystva]," *Russian Chemical Reviews (Uspekhi Khimii)*, vol. 81, no. 6, pp. 549–559, 2012.
- [3] N. S. Hosmane, *Boron Science: New Technologies and Applications*. Boca Raton (Florida, US): CRC Press (Taylor and Francis Group), 2012.
- [4] D. Radev and E. Ampaw, "Classical and contemporary synthesis methods of boron carbide powders," *Comptes rendus de l'Académie bulgare des sciences: sciences mathématiques et naturelles*, vol. 68, no. 8, pp. 945–956, 2015.
- [5] G. Goller, C. Toy, A. Tekin, and C. K. Gupta, "The production of boron carbide by carbothermic reduction," *High Temperature Materials and Process*, vol. 15, no. 1–2, pp. 117–122, 1996.
- [6] F. Thevenot, "Boron carbide—a comprehensive review," *Journal of the European Ceramic Society*, vol. 6, no. 4, pp. 205–225, 1990.
- [7] D. A. Ovsyannikov, M. Y. Popov, S. A. Perfilov, V. M. Prokhorov, B. A. Kul'nitskiy, I. A. Perezhogin, and V. D. Blank, "Highly Rigid Ceramics Based on Boron Carbide and Fullerite Products [Vysokotverdaya keramika na osnove karbida bora i proizvodnykh fullerita]," *Semiconductors/Physics of the Solid State*, vol. 59, no. 2, pp. 318–321, 2017.
- [8] A. G. Merzhanov and I. P. Borovinskaya, "Historical Retrospective of SHS: An Autoreview," *International Journal of Self-Propagating High-Temperature Synthesis*, vol. 17, no. 4, pp. 242–265, 2008.
- [9] R. G. Abdulkarimova, A. J. Seidualiyeva, and K. Kamunur, "Synthesis of Composite Materials Based on Borides of Metals and Aluminum Oxide under the Combustion Conditions," *Journal of Materials Science and Engineering B*, vol. 8, no. 3–4, pp. 56–65, 2018.
- [10] E. A. Levashov, A. S. Rogachev, V. V. Kurbatkina, Y. M. Maximov, and V. I. Yuhkvid, *Promising materials and technologies of self-propagating high temperature synthesis [Perspektivnyye materialy i tekhnologii samorasprostranyayushchegosya vysokotemperaturnogo sinteza]*. Moscow: NUST MISIS, 2011.
- [11] A. G. Merzhanov and A. S. Mukasyan, *Solid Flame Combustion [Tverdoplammennoye gorenije]*. Moscow: Torus Press, 2007.
- [12] J. H. Lee, W. Won, S. M. Joo, D. Y. Maeng, and H. Kim, "Preparation of B_4C powder from B_2O_3 oxide by SHS process," *Journal of Materials Science Letters*, vol. 19, no. 11, pp. 951–954, 2000.
- [13] V. A. Shcherbakov, A. N. Gryadunov, M. I. Alymov, and N. V. Sachkova, "Combustion synthesis and consolidation $\text{B}_4\text{C-TiB}_2$ composites [SVS-kompaktirovaniye kompozita $\text{B}_4\text{C-TiB}_2$]," *Letters on materials*, vol. 6, no. 3, pp. 217–220, 2016.
- [14] Z. A. Mansurov, S. M. Fomenko, A. N. Alipbaev, R. G. Abdulkarimova, and V. E. Zarko, "Aluminothermic Combustion of Chromium Oxide Based Systems under High Nitrogen Pressure," *Combustion Explosion, and Shock Waves*, vol. 52, no. 2, pp. 184–192, 2016.
- [15] I. H. Jung and M. A. Van Ende, "Computational

- Thermodynamic Calculations: FactSage from CALPHAD Thermodynamic Database to Virtual Process Simulation,” *Metallurgical and Materials Transactions B*, vol. 51, no. 5, pp. 1851–1874, 2020.
- [16] S. Mishra and L. C. Pathak, “Self-Propagating High-Temperature Synthesis (SHS) of Advanced High-Temperature Ceramics,” *Key Engineering Materials*, vol. 395, pp. 15–38, 2008.
- [17] M. A. Korchagin, T. F. Grigorieva, B. B. Bohonov, A. P. Sharafutdinov, B. B. Barinova, and N. Z. Lyakhov, “Solid-State Combustion in Mechanically Activated SHS Systems. II. Effect of Mechanical Activation Conditions on Process Parameters and Combustion Product Composition,” *Combustion, Explosion, and Shock Waves*, vol. 39, no. 1, pp. 51–58, 2003.
- [18] M. A. Korchagin and D. V. Dudin, “The use of self-propagating high temperature synthesis and mechanical activation for production of nanocomposites [Ispol’zovaniye samorasprostranyayushchegosya vysokotemperaturnogo sinteza i mekhanicheskoy aktivatsii dlya polucheniya nanokompozitov],” *Combustion, Explosion, and Shock Waves*, vol. 43, no. 2, pp. 176–187, 2007.
- [19] S. Tolendiuly, S. M. Fomenko, R. G. Abdulkarimova, and A. Akishev, “Synthesis and superconducting properties of the $\text{MgB}_2@ \text{BaO}$ composites,” *Inorganic and Nano-Metal Chemistry*, vol. 50, no. 5, pp. 349–353, 2020.
- [20] S. Tolendiuly, S. M. Fomenko, Z. A. Mansurov, G. Dannangoda, and K. S. Martirosyan, “Self-propagating high temperature synthesis of MgB_2 superconductor in high-pressure of argon condition,” *Eurasian Chemico-Technological Journal*, vol. 19, no. 2, pp. 177–181, 2017.
- [21] F. Bernard and E. Gaffet, “Mechanical alloying in the SHS research,” *International Journal of Self-Propagating High-Temperature Synthesis*, vol. 10, no. 2, pp. 109–132, 2001.
- [22] Z. V. Eremeeva, S. Vorotilo, V. S. Panov, L. V. Myakisheva, A. I. Lizunov, A. A. Nepapushev, D. A. Sidorenko, and D. Y. Mishunin, “Structure and Properties of Boron Carbide Produced by SHS and Mechanochemical Synthesis: a Comparative Study,” in *Fifteenth International Symposium on Self-Propagating High-Temperature Synthesis*, Moscow, Russia, 2019, pp. 96–98.
- [23] N. N.P., B. I.P., and M. A.G., *Combustion Processes in Chemical Technology and Metallurgy [Protsessy goreniya v khimicheskoy tekhnologii i metallurgii]*. Joint Institute of Chemical Physics [Russian], 1975, ch. Thermal Dynamic Analysis of SHS Reactions [Termodinamicheskiy analiz reaktsiy SVS].
- [24] D. S. Raimkhanova, S. M. Fomenko, R. G. Abdulkarimova, and Z. A. Mansurov, “Effect of Argon Pressure and Aluminum Content (in $\text{TiO}_2 - \text{H}_3\text{BO}_3 - \text{Al}$ Mix) on Combustion and Formation of Chemical Composition in Combustion Products,” *Advanced Materials Research*, vol. 746, pp. 62–67, 2013.
- [25] Z. A. Mansurov, D. S. Abdulkarimova, O. Odawara, A. V. Gubarevich, A. S. Rogachev, N. Shkodich, and N. N. Kochetov, “Peculiarities of Self-Propagating High-Temperature Synthesis and Structure Formation of $\text{TiB}_2 - \text{Al}_2\text{O}_3$ and $\text{CrB}_2 - \text{Al}_2\text{O}_3$ Composites,” *Eurasian Chemico-Technological Journal*, vol. 13, no. 3–4, pp. 161–168, 2011.

Tech Note. Embrittlement of duplex stainless steels SAF 2205 y SAF 2507 subjected to artificial aging heat treatments

Laura Sáenz ^{*,a,b} , Krishna Aguilera ^c , Reinaldo Heredia ^c, Trino Romero ^d ,
Caterine Donoso ^b , Marcos Singaña ^b 

^aCentro de Investigación de Materiales CIM, Facultad de Ingeniería Universidad de Carabobo, Valencia, Venezuela.

^bUniversidad de las Fuerzas Armadas ESPE, Sangolquí, Quito, API715231B, Ecuador.

^cEscuela de Ingeniería Mecánica, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela.

^dEscuela de Ingeniería Química, Facultad de Ingeniería, Universidad Central de Venezuela, Caracas, Venezuela.

Abstract.- The research studies the embrittlement of SAF 2205 and SAF 2507 duplex stainless steels subjected to artificial aging treatments. Artificial aging treatments were performed Charpy V-shaped tensile and impact specimens, the treatment consisted of two phases, first the solubilization of the steels was carried at temperatures between 1100 °C and 1200 °C with times between 1 h and 2 h, and second, the artificial aging was carried out at temperatures of 475 °C, 700 °C, 850 °C, 870 °C, and 900 °C. There are several factors that affect the mechanical characteristics of stainless steels, for example, the existence of micro-segregations and the precipitation of secondary phases, which can take place during solidification, shaping, heat treatment, welding operations, etc. The study carried out allowed the construction of curves that show the changed in the mechanical behavior obtained from the SAF 2205 and SAF 2507 duplex steels, where it is observed that as aging time increases, ductility, mechanical strength, and impact toughness are reduced. In the microstructure was observed the precipitation of sigma and Chi secondary phases.

Keywords: microstructure; annealing; ultimate tensile stress; elongation; toughness.

Nota Técnica. Fragilización de aceros inoxidable dúplex SAF 2205 y SAF 2507 sometidos a tratamientos térmicos de envejecimiento artificial

Resumen.- La investigación estudia la fragilización de los aceros inoxidable dúplex SAF 2205 y SAF 2507 sometidos a tratamientos de envejecimiento artificial. Las probetas de tracción e impacto Charpy en V fueron sometidas a tratamiento de envejecimiento artificial que constó de dos fases, primero se realizó la solubilización a temperaturas entre 1100 °C y 1200 °C con tiempos entre 1 h y 2 h, y segundo el envejecimiento artificial fue realizado a temperaturas de 475 °C, 700 °C, 850 °C, 870 °C y 900 °C. Existen varios factores que afectan a las características mecánicas de los aceros inoxidable, por ejemplo, la existencia de micro segregaciones y la precipitación de fases secundarias, que pueden tener lugar durante la solidificación, conformado, tratamiento térmico, operaciones de soldadura, entre otros. El estudio realizado permitió la construcción de curvas que muestran los cambios en el comportamiento mecánico de los aceros dúplex SAF 2205 y SAF 2507, donde se observa que a medida que incrementa el tiempo de envejecimiento, se reduce la ductilidad, resistencia mecánica y tenacidad de impacto. En la microestructura se observó la precipitación de fases secundarias sigma y Chi.

Palabras clave: microestructura; recocido; esfuerzo máximo; ductilidad; tenacidad.

Recibido: 19 de enero, 2021.

Aceptado: 16 de marzo, 2021.

1. Introducción

Los primeros aceros inoxidable dúplex que aparecieron en la historia estaban compuestos

básicamente por cromo y níquel, y estos han sido denominados aceros de primera generación, se produjeron en Suecia en el año 1930, para ser usados en la industria del papel con la finalidad de reducir los problemas de corrosión que se presentaban regularmente es esa industria. De acuerdo con la Asociación Internacional de Molibdeno [1], el molibdeno ha pasado a formar parte de los componentes básicos incrementando

* Autor para correspondencia:

Correo-e:lsaenz29@gmail.com (L. Sáenz)

así la resistencia, estos han sido denominados aceros de última generación. Investigaciones de Nilsson [2] y Charles [3] encontraron que los aceros dúplex poseen una notable combinación de propiedades mecánicas y resistencia a la corrosión, producto de su estructura bifásica austenita-ferrita. La ferrita aporta alta resistencia mecánica y a la corrosión, mientras que la austenita ductilidad y resistencia a la corrosión uniforme. La resistencia a la corrosión es superior a la de los aceros austeníticos convencionales, sin que se experimente una pérdida en la ductilidad considerable de los mismos. Sin embargo, las propiedades de estos son sensibles a variaciones en su composición química, especialmente de cromo, molibdeno, níquel y nitrógeno. En [4] y [5] se determinó que los aceros dúplex además de poseer las fases de ferrita y austenita, también pueden aparecer otras fases secundarias debido a variaciones en la temperatura a la que se encuentren expuestos, por lo general producto de algún tratamiento térmico. La formación de estas fases modifica la microestructura original del acero, lo cual influye directamente en sus propiedades mecánicas, variando su dureza, resistencia mecánica, tenacidad a la fractura, entre otras. Por lo que pudiera destinarse el acero a aplicaciones específicas en función del comportamiento mecánico que presente a una determinada temperatura. Pohl [4], Llorca-Isern [6], Kim [7] y Ferro [8], han reportado que la precipitación de fases intermetálicas en aceros inoxidables dúplex (SSD) y aceros inoxidables súper dúplex (SSSD) ha sido producto de los tratamientos isotérmicos realizados al acero, verificándose precipitación suficiente para deteriorar las propiedades mecánicas y de resistencia a la corrosión en tiempos mayores de 100 segundos a temperaturas en el rango entre 850 °C y 950 °C. Durante el enfriamiento de los SSD, la precipitación de la austenita está acompañada de precipitación de otras fases (M_7C_3 , $M_{23}C_6$, σ , Chi χ , Laves, α') dependiendo de la temperatura máxima alcanzada y del tiempo de permanencia en el intervalo de precipitación. Estas fases intermetálicas precipitan por nucleación en sitios de alta energía como puntos triples de granos,

límite de grano, inclusiones, dislocaciones y vacancias y es favorecida por la distribución preferencial de los elementos aleantes en la ferrita y en la austenita. Así el molibdeno, cromo y silicio se disuelven preferencialmente en la ferrita favoreciendo la precipitación de fases intermetálicas, mientras que el manganeso, níquel y nitrógeno se difunden preferencialmente en la austenita. En [4] han indicado que la austenita secundaria, γ_2 : aparece cuando se realiza un recocido a altas temperaturas, próximas o incluso incluidas en el dominio monofásico de ferrita, tras el enfriamiento, la ferrita se encuentra enriquecida de elementos gammágenos. Un posterior calentamiento a temperaturas comprendidas entre 600 °C y 900 °C produce la transformación de la ferrita en austenita, y es esta nueva austenita formada lo que se denomina austenita secundaria. Pohl [4], Llorca-Isern [6], Kim [7] y Ferro [8] indicaron que la fase σ es la más perjudicial de todas las fases que puedan precipitar en los aceros dúplex, influye directamente en la tenacidad y resistencia a la corrosión, afecta la ductilidad en caliente y frío, favorece la aparición de grietas muy finas. El intervalo de aparición en el que se evidencia esta fase es de aproximadamente (550 °C–900 °C) precipitando en las uniones triples o en los límites de grano ferrita/austenita, coexiste con fases como los nitruros de cromo, la austenita secundaria, la fase Chi, la fase R, la fase Pi, la fase Tao, las cuales afectan su nucleación y crecimiento, concordando con los estudios de Pohl [4], Llorca-Isern [6], Kim [7] y Ferro [8]. Los estudios realizados por Pohl [4], Ghosh y Mondal [9], Kim [7] y Ferro [8], han mostrado que la fase Chi, χ es una fase intermetálica, se forma en los límites de las fases ferrita-austenita, creciendo hacia el interior de la ferrita y puede ser identificada usando difracción de electrones en el microscopio electrónico de transmisión. Esta fase suele aparecer en el rango de temperatura de 700 °C a 900 °C y en cantidades menores que la fase σ , además de que ocurre en tiempos menores, y es considerada metaestable ya que actúa como facilitadora para la formación de fase σ , descomponiéndose completamente en ella después de largos tiempos de exposición. El efecto de esta fase sobre la tenacidad y el comportamiento

contra la corrosión son perjudiciales, y se dificulta separar de la fase sigma. Pohl [4] define la Fase R, como una fase intermetálica trigonal que puede precipitar en el intervalo de temperaturas de 550 °C a 800 °C, influyen en la fragilización y disminución de las propiedades anticorrosivas. Esta fase presenta una composición aproximada de (16-30) %Cr, (3-5) %Ni, y (25-40) %Mo. Pohl [4], Kim [7] y Ferro [8] definen la Fase pi, π como una estructura cristalina cúbica y al igual que la fase R, contribuye a la fragilización y corrosión por picadura cuando el material ha sido sometido a un tratamiento isotérmico de envejecimiento a 600 °C. Esta fase contiene aproximadamente 28 %Fe, 35 %Cr, 3 %Ni y 34 %Mo. Pohl [4], Kim [7] y Ferro [8] han encontrado que la precipitación de nitruros de cromo (Cr_2N) se presenta en el intervalo de temperaturas entre 700 °C y 900 °C por el enfriamiento rápido desde altas temperaturas de solubilización como consecuencia de una sobresaturación de nitrógeno en la ferrita. Ocurre preferencialmente en los límites de grano ferrita/ferrita, presentándose como partículas alargadas que precipitan intergranularmente. Pohl [4], Kim [7] y Ferro [8] han encontrado que a temperaturas inferiores a 500 °C las reacciones de precipitación de las fases mencionadas son relativamente lentas, lo que produce un fenómeno denominado fragilización, exactamente a una temperatura de 475 °C. La formación de estas fases influye directamente en las propiedades mecánicas y anticorrosivas de los aceros inoxidable dúplex. Topolska y Lawanoski [10], indicaron que el endurecimiento por precipitación es un poderoso método para controlar el endurecimiento óptimo por dispersión, es decir para aumentar la dureza y resistencia de las aleaciones. El endurecimiento por precipitación se basa en los cambios de la solubilidad de sólido con la temperatura para producir partículas finas de una impureza de fase, que impiden el movimiento de dislocaciones o defectos a través de la estructura del cristal. Dado que las dislocaciones son a menudo los operadores dominantes de la plasticidad, esto sirve para endurecer el material. Las impurezas desempeñan la misma función que los refuerzos en los materiales compuestos

reforzados. Así como la formación de hielo en el aire puede producir nubes, nieve o granizo, dependiendo de la historia térmica de una porción dada de la atmósfera, la precipitación de los sólidos puede producir diferentes tamaños de partículas, que tienen propiedades radicalmente diferentes. A diferencia del temple ordinario, las aleaciones deben mantenerse a temperatura elevada durante horas para permitir que la precipitación tenga lugar, este retardo de tiempo se denomina envejecimiento. Topolska y Lawanoski [10] encontraron que cuando los aceros inoxidable dúplex son expuestos a altas temperaturas de operación incluso durante pequeños períodos de tiempo, precipitan diversas fases secundarias. Resulta de interés, por lo tanto, el estudio del efecto de la formación de estos precipitados, sobre el comportamiento mecánico de los aceros dúplex, para determinar si influyen favorablemente o no sobre las propiedades mecánicas de los mismos. Sin embargo, esta familia de aceros sufre fragilización cuando son expuestos a un rango de temperaturas, lo que limita sus aplicaciones. Topolska S. y Lawanoski J. [10], indican que la severa caída en la tenacidad de impacto y resistencia mecánica y ductilidad que puede ocurrir en el material hace necesaria una mayor investigación sobre su respuesta al envejecimiento a temperaturas intermedias y altas con el fin de precisar la confiabilidad de estos cuando son sometidos a tales condiciones en servicio. La presente investigación tiene como objetivo evaluar la fragilización de los aceros dúplex SAF 2205 y SAF 2507 sometidos a tratamientos térmicos de envejecimiento.

2. Procedimiento experimental

2.1. Material

Se utilizaron muestras de aceros inoxidable dúplex SAF 2205, y SAF 2507, la composición química se muestra en la Tabla 1.

2.2. Tratamientos térmicos

Se realizaron tratamientos térmicos de solubilización con posterior envejecimiento artificial a probetas de aceros dúplex de tracción y Charpy en V de acuerdo con las normas para ensayos de

Tabla 1: Composición química de los aceros inoxidables dúplex

% peso	SAF 2205	SAF 2507
%C (máx.)	0,03	0,02
%Cr	21 – 23	25
%Ni	4,5 – 6,5	7,0
%Mo	2,5 – 3,5	4
%N	0,08 – 0,2	0,27

tensión e impacto ASTM A370-17 [11] y ASTM E23-17 [12] respectivamente, a las condiciones del tratamiento especificadas en la Tabla 2.

Tabla 2: Especificaciones de los tratamientos térmicos realizados a las probetas de tracción y Charpy en V de aceros inoxidables dúplex SAF2205 y SAF 2507

Aceros dúplex	Condiciones Solubilización	Condiciones Envejecimiento artificial
SAF2205 Sáenz [13]	1120 °C por 2 horas	475 °C por 1, 2, 6,5, 12, 24, 40 y 120 horas.
SAF2205 Griman y Marcano [14]	1020 °C por 1 hora	475 °C por 14, 24, 72 y 120 horas.
SAF2205 Grippi y Ocando [15]	1100 °C por 1 hora	850 °C por 48, 54 y 60 minutos.
SAF2205 Plaza [16]	1100 °C por 1 hora	850 °C por 6, 12 y 24 horas.
SAF2507 Mosquera y Tiamo [17]	1100 °C por 90 minutos	700 °C por 24, 48, 72 y horas.
SAF2205 Sáenz et al. [18]	1010 °C por 2 horas	850 °C por 0,5; 1, 5, 24 y 48 horas.
SAF2507 Mejias et al. [19]	1120 °C por 2 horas	850 °C por 6, 12 y 24 horas.
SAF2507 Arruti [20]	1100 °C por 90 minutos	870 °C por 2, 4, 6 y 10 minutos.
SAF2507 Mejias [21]	1100 °C por 90 minutos	900 °C por 2, 4, 6 y 10 minutos.

2.3. Ensayos mecánicos

Los ensayos de tracción se realizaron en una máquina Universal marca Galdabini de 20 ton de capacidad, de acuerdo con las especificaciones de la norma ASTM A – 370-17 [11] se utilizaron probetas planas y cilíndricas. Los ensayos de impacto Charpy en V, se realizaron con un péndulo marca Karl Frank, con un intervalo de 0 hasta 30 Kp.m, acorde con la Norma ASTM E-23-17 [12].

3. Resultados

3.1. Resistencia mecánica

En la Figura 1, se muestran las curvas de resistencia mecánica en función del tiempo de envejecimiento para los aceros inoxidables dúplex SAF 2205 y 2507 para las condiciones especificadas en la Tabla 2, donde se observa que el esfuerzo de fluencia del acero inoxidable dúplex SAF 2205 genera una familia de curvas de comportamiento similar, a pesar de arrojar valores distintos del esfuerzo a la misma temperatura del tratamiento, como es el caso de 475 °C, reportado por Sáenz [13] y Griman y Marcano [14], se asume fue debido al tamaño de las probetas utilizadas por Sáenz [13], probetas cilíndricas tipo sub-size, los posibles elementos aleantes o impurezas que componen las muestras de estudio o al proceso de mecanizado que recibió cada probeta, debido a que eran de dimensiones y forma diferente. Puede notarse que para tiempos prolongados de envejecimiento y una temperatura inferior, el esfuerzo de fluencia se incrementa consecuentemente, por lo que se puede concluir que mientras mayor sea el tiempo de envejecimiento del acero dúplex, la resistencia mecánica se incrementa y provoca el endurecimiento al acero dúplex. En la Figura 1, se muestran resultados del estudio de envejecimiento realizados por Grippi y Ocando [15] y Plaza [16] al acero inoxidable dúplex SAF 2205 a 850 °C, se encontró que, para periodos cortos de exposición y altas temperaturas (850 °C), inicialmente el esfuerzo disminuye considerablemente y no es sino a partir de una hora que empieza a incrementar la resistencia mecánica, se presume que es debido a la formación de fase sigma lo cual será estudiado posteriormente con el estudio de la microestructura. En la Figura 1, Arruti [20] encontró que para el acero SAF 2507 mientras mayor es la temperatura del tratamiento térmico de envejecimiento mayor es la resistencia mecánica del acero, esto sucede para las temperaturas de 870 °C, en Mejias [21] a 900 °C, pero el estudio realizado por Mosquera y Tiamo [17] a 700 °C se encuentra un punto de inflexión a 72 horas y luego disminuye la resistencia mecánica. Por otra

parte, el comportamiento mecánico es inverso para Cepeda [22] a la temperatura de 685 °C, donde se observa que a medida que aumenta el tiempo de envejecimiento la resistencia mecánica disminuye. Las investigaciones realizadas por Llorca-Isern [6], Kim [7] y Ferro [8], Arruti [20], Cepeda [22], Mejias [19] y Sáenz [18] han encontrado que los comportamientos de la resistencia mecánica de los aceros inoxidable dúplex están asociados con los precipitados (carburos, nitruros) o las partículas de segunda fase (σ , χ , α' , γ_2) que se han formado para cada una de las condiciones de las variables temperatura y tiempo en las que se han realizado los tratamientos térmicos de envejecimiento. En la Figura 1, se puede observar que el esfuerzo de fluencia tiende a incrementar a medida que incrementa el tiempo de envejecimiento del acero, generando así un leve incremento en la resistencia mecánica del acero inoxidable dúplex 2507 que puede ser provocado a la descomposición de la fase ferrita por la difusión de sus elementos aleantes (cromo y molibdeno), a elevadas temperaturas, lo que trae como consecuencia la formación de las fases sigma (σ), Chi (χ), y austenita secundaria (γ_2) como se observa en la Figura 1. En los resultados obtenidos en las investigaciones realizadas por Llorca-Isern [6], Kim [7] y Arruti [20], los precipitados se presentan en forma de partículas finas que se oponen al movimiento de las dislocaciones obligándolas a pasar a través de estos.

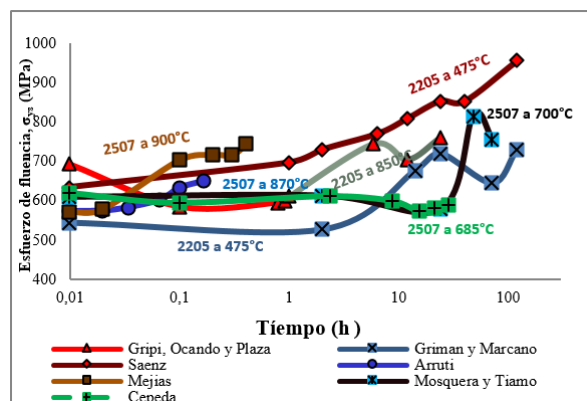


Figura 1: Curva de Resistencia Mecánica ($\sigma_{y,s}$) en función del tiempo de envejecimiento de los aceros inoxidable dúplex SAF 2205 y 2507 sometidos a tratamientos de envejecimiento artificial

El esfuerzo de fluencia $\sigma_{y,s}$ se incrementa con el tiempo de envejecimiento, y provoca la fragilización del acero dúplex, debido a la formación de fases inestables a diferentes temperaturas y a la alta concentración de precipitados que favorece la formación de microgrietas y da lugar a superficies de fractura frágil, con una escasa reducción en el diámetro de la zona calibrada de las probetas ensayadas.

3.2. Ductilidad

En la Figura 2 se muestra el comportamiento de la ductilidad evaluado a través del porcentaje de deformación de Ingeniería (%e), se puede observar que a medida que aumenta el tiempo de envejecimiento, el %e se mantiene en rangos de valores de poca fluctuación y a partir de cierto tiempo comienzan a disminuir, lo que indica que el material se endurece por efecto del tratamiento térmico de envejecimiento por la formación de precipitados o a la reducción de la cantidad de dislocaciones que dificultan y limitan la capacidad de deformación a medida que aumenta el tiempo del tratamiento térmico de envejecimiento. Sáenz [13] y Griman y Marcano [14] encontraron que para el acero 2205 a 475 °C, aumenta la ductilidad a medida que incremento el tiempo de envejecimiento (Figura 2), este comportamiento es similar al reportado por Grippi y Ocando [15] y Plaza [16] para la temperatura de 850 °C.

Cepeda [22], y Mosquera y Tiamo [17] han reportado que para el acero 2507 a 685 °C y 700 °C respectivamente, existe endurecimiento del material hasta 28 h y 70 h de tratamiento respectivamente, mientras que Arruti [20] a 870 °C, y Mejias [21] y Mejias [19] a 900 °C, encontraron que el endurecimiento se evidencia en tiempos muy cortos de 0,16 h y 0,4 h. Este comportamiento evidencia la formación de fases secundarias, como la fase sigma en cortos periodos de exposición del tratamiento térmico de envejecimiento en este rango de temperaturas (850 °C y 900 °C). En las investigaciones realizadas por Llorca-Isern [6], Kim [7], Arruti [20], Cepeda [22], Sáenz [18] y Mejias [19], se observa claramente que, aunque se produce un ligero aumento en la resistencia del acero dúplex 2507, ocurre una

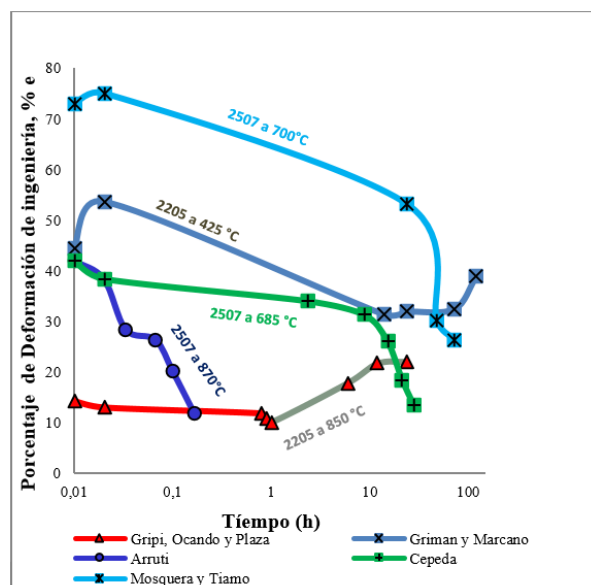


Figura 2: Curva de Ductilidad en función del tiempo de envejecimiento de los aceros inoxidables dúplex SAF

importante reducción de la ductilidad, producida por la formación de precipitados como la fase Chi (χ) y la fase sigma (σ) las cuales se pueden observar en la Figura 3. En la Figura 2 se muestra como el aumento en la temperatura tiene una influencia negativa en la ductilidad del material, aun cuando el tiempo de exposición de las probetas dentro del horno es considerablemente reducido con respecto a otros. Arruti [20], encontró que para la temperatura de 870 °C, se reflejó que con sólo 2 minutos de tratamiento el porcentaje de reducción de área disminuye un 45% y fue disminuyendo progresivamente entre 4 y 10 minutos de envejecimiento, incluso el tratamiento de solubilización produce un efecto adverso en la ductilidad del acero dúplex. Las investigaciones de Arruti [20], Cepeda [22], Mosquera y Tiamo [17], y Mejias [21] arrojaron que el comportamiento mecánico de la resistencia y la ductilidad del acero dúplex estudiado 2507, están en concordancia con la investigación previa de Ghosh y Mondal [9], en la cual se determinó que la formación de precipitados como la fase Chi (χ) hacia el interior de los granos de ferrita y de fase Sigma (σ) a lo largo de los límites de las fases austenita/ferrita, produce en un acero dúplex un pequeño incremento en

la resistencia del material pero una significativa reducción en los valores de su ductilidad, cuando es sometido a altas temperaturas de envejecimiento, aun si el tiempo del tratamiento es corto.

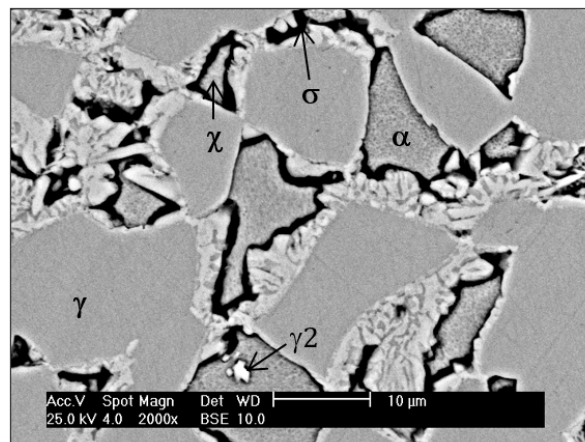


Figura 3: Microestructura del acero inoxidable dúplex SAF 2507 observado mediante SEM (X2000), en su condición original, solubilizada a 1100°C por 90 min y envejecida isotérmicamente a 870°C por 10 min., se observan las fases ferrita y austenita y las fases secundarias sigma, Chi y austenita secundaria (σ , χ , γ_2) [20].

3.3. Tenacidad de impacto

De los estudios realizados por Sáenz [18], Griman y Marcano [14], Grippi y Ocando [15], Plaza [16], Cepeda [22], Mosquera y Tiamo [17], Arruti [20] y Mejias [21] se pueden resumir en la Figura 4 donde se observa que para todos los aceros inoxidables dúplex 2205 y 2507 sometidos al tratamiento de envejecimiento realizado a las temperaturas de 475 °C, 850 °C y para el 2507 a las temperaturas de 685 °C, 700 °C, 870 °C y 900 °C, a medida que se incrementa el tiempo de envejecimiento la tenacidad de impacto cae bruscamente, siendo más crítica a temperaturas altas (870 °C y 900 °C) donde la fragilización del acero es inmediata, antes de una hora de tratamiento, lo que indica la formación de precipitados de segunda fase como lo son la fase Sigma (σ) y Chi (χ) que provocan que el acero se fragilice a medida que aumenta el tiempo de envejecimiento, es decir, pierde su capacidad de absorber energía, los resultados

se observan en la microestructura reportados en la Figura 3 [20]. La tenacidad de impacto evaluada a través de la energía absorbida por impacto disminuyó considerablemente para los estudios en comparación, puede observarse en la Figura 4 que el par tiempo-temperatura va a influir directamente en el descenso de los valores de la tenacidad de impacto. Esto se debe a la formación de precipitados (como las fases: α' , σ , χ , carburos y nitruros), tal como lo demuestran los estudios realizados por Llorca-Isern [6], Kim [7], Arruti [20], Cepeda [22] y Mejías [21]. Por otra parte, Nilsson [2] observó que, en la estructura original del material, estos impiden el movimiento de las dislocaciones durante la deformación, forzando a que éstas pasen a través de las partículas del precipitado, cortándolas, lo que incrementa la posibilidad de iniciación de microgrietas, conduciendo a un comportamiento frágil del material. Nilsson [2] encontró que tan sólo una fracción de volumen del uno por ciento de precipitados puede ocasionar que la energía absorbida por impacto del material alcance un valor equivalente a un tercio de la presentada en su condición original. Se presume que la disminución drástica en la energía absorbida por impacto del material se deba principalmente a la formación de fase Sigma (σ) en los límites de las fases austenita/ferrita y ferrita/ferrita. La fase Sigma (σ) es una fase secundaria que fragiliza a los aceros dúplex incluso cuando aparece en pequeñas cantidades. Una vez que precipita, una pequeña deformación en el material puede producir fracturas frágiles con una estructura fina y transcristalina de partículas de fase Sigma (σ). Las grietas tienden a crecer a través de las partículas de esta fase por largas distancias, favoreciendo al comportamiento frágil del material. Investigaciones realizadas por Nilsson [2], han demostrado que los aceros dúplex toleran hasta un contenido de fase Sigma (σ) del 2 al 5% antes de alcanzar una energía absorbida por impacto crítica de 32 J encontrada por Hau *et al.* [23]. El acero inoxidable dúplex 2205 a 475 °C cumple con esta relación encontrada por Hau [23], mientras que el resto de los aceros fragiliza bruscamente por lo que limita el uso del

acero inoxidable dúplex 2507 a temperaturas altas (685 °C, 700 °C, 850 °C y 900 °C) donde la energía absorbida por impacto es inferior a 32 J.

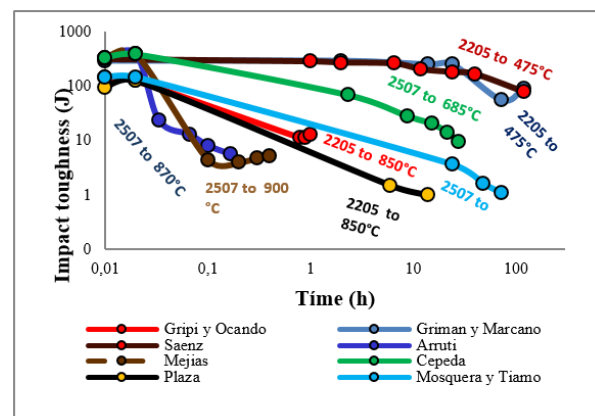


Figura 4: Curva de Tenacidad de impacto en función del tiempo de envejecimiento de los aceros inoxidables dúplex SAF 2205 y 2507 sometidos a tratamientos de envejecimiento artificial

4. Conclusiones

El estudio realizado permitió la construcción de curvas que muestran el comportamiento mecánico de los aceros inoxidables dúplex SAF 2205 y SAF 2507 sometidos a tratamiento térmico de envejecimiento artificial realizado a las temperaturas de 475 °C, 700 °C, 850 °C, 870 °C y 900 °C. Los tratamientos de envejecimiento artificial al cual fueron sometidos los aceros inoxidable dúplex SAF 2205 y SAF 2507, ocasionaron el aumento de resistencia mecánica, así como la pérdida de ductilidad y tenacidad de impacto a medida que se incrementó el tiempo de envejecimiento, ocasionada por la presencia de fases secundarias como lo son la fase Sigma, Chi y austenita secundaria en los aceros inoxidables duplex estudiados. El acero duplex 2205 para altas y bajas temperaturas mantiene la ductilidad o tienen menor tendencia a endurecer o formar precipitados, mientras que el 2507 incrementa la resistencia mecánica a tiempos de envejecimiento muy cortos, endurece tempranamente provocando la fragilización del acero, limitando su uso.

5. Referencias

- [1] Asociación Internacional de Molibdeno, *Directrices prácticas para la fabricación de los aceros inoxidable dúplex*. Londres, Reino Unido: International Molybdenum Association(IMOA), 2012.
- [2] J. Nilsson, “Super Duplex Stainless Steels,” *Materials Science and Technology*, vol. 8, no. 8, pp. 685–700, 1992.
- [3] J. Charles, “Superduplex, stainless steels: structure and properties,” in *Duplex stainless steels’91*, vol. 1. Les Ulis, France: Les Editions de Physique, 1991.
- [4] M. Pohl, O. Storz, and T. Glogowski, “Effect of intermetallic precipitations on the properties of duplex stainless steel,” *Materials Characterization*, vol. 58, no. 1, pp. 65–71, 2007.
- [5] F. M. Elsabbagh, R. M. Hamouda, and M. A. Taha, “On microstructure and microhardness of isothermally aged UNS S32760 and the effect on toughness and corrosion behavior,” *Journal of Materials Engineering and Performance*, vol. 25, pp. 275–284, 2014.
- [6] N. Llorca-Isern, I. López-Jiménez, H. López-Luque, M. V. Biezma, and A. Roca, “Study of the precipitation of secondary phases in duplex and superduplex stainless steel,” in *Materials Science Forum*, vol. 879. Trans Tech Publications, 2017.
- [7] S.-K. Kim, K.-Y. Kang, M.-S. Kim, and J.-M. Lee, “Low-Temperature Mechanical Behavior of Super Duplex Stainless Steel with Sigma Precipitation,” *Metals*, vol. 5, no. 3, pp. 1732–1745, 2015.
- [8] P. Ferro, F. Bonollo, and G. Timelli, “Sigma phase precipitation modelling in a UNS S32760 superduplex stainless steel,” *La Metallurgia Italiana*, vol. 104, no. 5, pp. 7–12, 2012.
- [9] S. K. Ghosh and S. Mondal, “High temperature ageing behavior of a duplex stainless-steel material,” *Materials Characterization*, vol. 59, no. 12, pp. 1776–1783, 2008.
- [10] S. Topolska and J. Lawanoski, “Effect of microstructure on impact toughness of duplex and super duplex stainless steel,” *Journal of Achievements in Materials and Manufacturing Engineering*, vol. 36, no. 2, 2009.
- [11] ASTM A 370, *Standard Test Methods and Definitions for Mechanical Testing of Steel Products*, American Society for Testing and Materials, USA, 2018.
- [12] ASTM E 23, *Test Methods for Notched Bar Impact Testing of Metallic Materials*, American Society for Testing and Materials, USA, 2014.
- [13] L. Sáenz, “Naturaleza fractal de la deformación plástica de un acero inoxidable dúplex envejecido,” *Revista Ingeniería y Sociedad*, 2011.
- [14] J. Griman and J. Marcano, “Evaluación de la tenacidad de impacto del acero inoxidable dúplex SAF 2205 sometido a un tratamiento térmico de envejecimiento 475 °C,” Trabajo de Grado de Ingeniería Mecánica, Facultad de Ingeniería, Universidad de Carabobo, Venezuela, 2005.
- [15] L. Grippi and A. Ocando, “Evaluación de la tenacidad de impacto del acero inoxidable dúplex SAF 2205 sometido a un tratamiento térmico de envejecimiento 850 °C con unos tiempos de 48, 54 y 60 minutos,” Trabajo de Grado de Ingeniería Mecánica, Facultad de Ingeniería, Universidad de Carabobo, Venezuela, 2006.
- [16] L. Plaza, “Evaluación de la tenacidad de impacto de un acero inoxidable dúplex SAF 2205 sometido a un tratamiento térmico de envejecimiento a una temperatura de 850 °C, con unos tiempos de 6,12 y 24 horas,” Trabajo de Grado de Ingeniería Mecánica, Facultad de Ingeniería, Universidad de Carabobo, Venezuela, 2007.
- [17] M. Mosquera and R. Tiamo, “Evaluación de la tenacidad de impacto de un acero inoxidable duplex SAF 2507 sometido a un tratamiento térmico de envejecimiento a una temperatura de 700 °C,” Trabajo de Grado de Ingeniería Mecánica, Facultad de Ingeniería, Universidad de Carabobo, Venezuela, 2008.
- [18] L. Sáenz, O. Hilders, and H. R., “Estudio del acero inoxidable dúplex SAF 2507 sometido a un tratamiento térmico de envejecimiento a una temperatura de 850 °C por tiempos entre 1 hora y 48 horas,” *Revista Facultad de Ingeniería UCV*, vol. 22, no. 3, pp. 21–30, 2007.
- [19] A. Mejías, J. Arruti, J. Cepeda, and L. Sáenz, “Influencia del tratamiento térmico de envejecimiento sobre las propiedades mecánicas y la microestructura de un acero inoxidable dúplex 2507,” *Revista Ingeniería UC*, vol. 17, no. 3, pp. 28–39, 2010.
- [20] J. Arruti, “Determinación de la influencia del tratamiento térmico realizado a una temperatura de 870 °C durante tiempos comprendidos entre 2 y 10 minutos sobre las propiedades mecánicas y la microestructura de un acero inoxidable dúplex,” Trabajo de Grado de Maestría, Universidad de Carabobo, Venezuela, 2009.
- [21] A. Mejías, “Evaluación de la tenacidad de impacto de un acero inoxidable dúplex SAF 2507 sometido a un tratamiento térmico de envejecimiento a una temperatura de 900 °C,” Trabajo de Grado de Maestría, Universidad de Carabobo, Venezuela, 2009.
- [22] J. Cepeda, “Determinación de la influencia del tratamiento térmico realizado a una temperatura de 685 °C sobre las propiedades mecánicas y la microestructura de un acero inoxidable dúplex SAF 2507,” Trabajo de Grado de Maestría, Universidad de Carabobo, Venezuela, 2009.
- [23] J. Hau and A. Seijas, Eds., *Sigma Phase Embrittlement of stainless steel in FCC Service. NACE 61st Annual Conference & Exposition*, no. 6578. NACE, 2006.

**Jornada de Investigación de la Escuela de
Ingeniería Eléctrica “Prof. César Rodolfo
Ruiz” Octubre, 2020.**

Design of a 34,5 kV line for backup auxiliary power services of the La Arenosa 765 kV Substation

Santiago Abreu , Antonio Otero , Eva Monagas * , César Rodolfo Ruiz ^(†),
Juan C. Ataya , Yonder Guevara 

Departamento de Potencia, Escuela de Ingeniería Eléctrica, Universidad de Carabobo, Valencia, Venezuela

La selección de este artículo fue realizada en el marco de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica “Prof. César Rodolfo Ruiz” Octubre 2020, siendo la evaluación, arbitraje, aceptación y edición a cargo de Revista Ingeniería UC.

Abstract.- The purpose of this paper is to present the new design of the 34,5 kV line, originally designed to back up the auxiliary services of the La Arenosa 765 kV Substation, once it was out of service due to a failure, lack of maintenance and later due to vandalism. The route planned for its recovery and commissioning corresponds to the same as the original design located between S/E La Arenosa 765 kV and S/E La Arenosa 400 kV. The design criteria are presented based on the diagnosis of the existing infrastructure, the electrical and mechanical calculations of conductors and structures. The SISTRANS software, which was prepared and validated by the School of Electrical Engineering of the University of Carabobo, was used for the calculations in order to comply with current national regulations. This design contributes to the applied electrical engineering research line of the Power Department.

Keywords: substations; auxiliary services; backup systems; design.

Diseño de línea 34,5 kV para respaldo de energía de los servicios auxiliares de la S/E La Arenosa 765 kV

Resumen.- El propósito de este artículo es presentar el estudio de reconstrucción de la línea 34,5 kV, originalmente diseñada para el respaldo de energía de los servicios auxiliares de la S/E La Arenosa 765 kV, una vez que quedó fuera de servicio por una falla, falta de mantenimiento y posteriormente por vandalismo. La ruta prevista para su recuperación y puesta en servicio corresponde a la misma del diseño original ubicada entre S/E La Arenosa 765 kV y S/E La Arenosa 400 kV. Se presentan los criterios de diseño con base en el diagnóstico de la infraestructura existente, los cálculos eléctricos y mecánicos de conductores y estructuras. Para los cálculos se utilizó el software SISTRANS, elaborado y validado por la Escuela de Ingeniería Eléctrica de la Universidad de Carabobo, con el fin de cumplir de las normativas nacionales vigentes. Este diseño aporta a la línea de investigación de ingeniería eléctrica aplicada del Departamento de Potencia.

Palabras clave: subestaciones; servicios auxiliares; sistemas de respaldo; diseño.

Recibido: 22 de octubre, 2020.

Aceptado: 18 de enero, 2021.

1. Introducción

La subestación La Arenosa es una importante vía para el transporte de energía desde el complejo hidroeléctrico del Guri hacia el occidente de Venezuela, se conecta con la subestación San

Gerónimo y es la vía para interconectar con otros centros de generación. Ella se conecta con la subestación La Horqueta, Yaracuy y Planta Centro por lo cual recibe mensualmente un aproximado de 800.000 MWh, cuenta con un patio de 765 kV y un patio de 400 kV tal como se muestra en la Figura 1.

La importancia de los servicios auxiliares en subestaciones de gran potencia, y especialmente en las subestaciones del Sistema Eléctrico Nacional, ha sido analizada en diversas investigaciones realizadas en distintas universidades venezolanas que

* Autor para correspondencia:

Correo-e: evamonagas@gmail.com (E. Monagas)

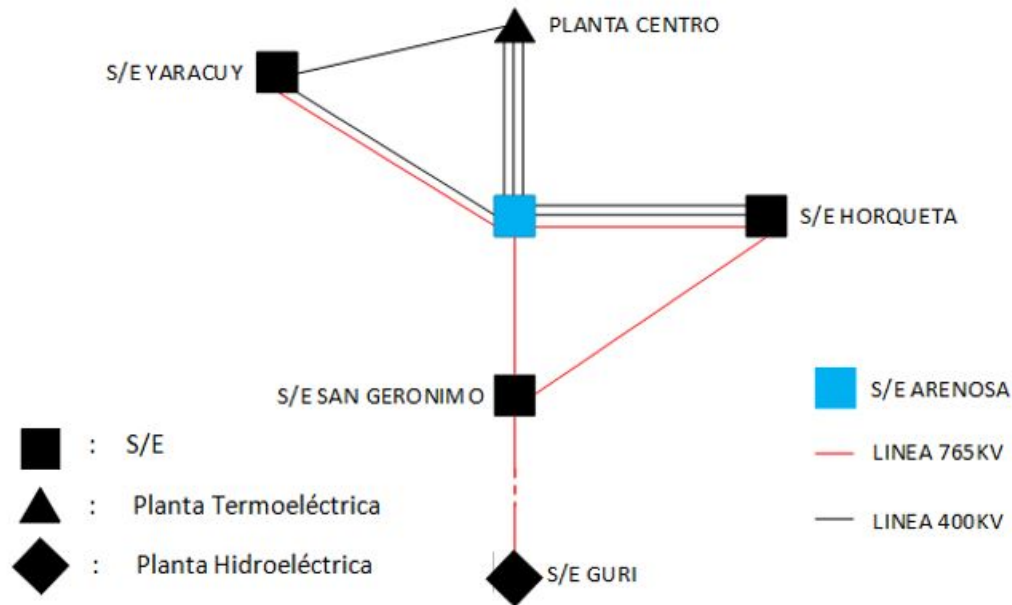


Figura 1: Sistema de anillo S/E Arenosa-S/E Horqueta-S/E Yaracuy [1]

han contribuido con el desarrollo de metodologías para el diseño, análisis del funcionamiento, la confiabilidad de los sistemas de alimentación y desarrollo de programas de mantenimiento.

Dada la importancia de la subestación La Arenosa, el sistema de alimentación de servicios auxiliares (AC y DC) fue diseñado con dos sistemas de respaldo de energía para cubrir los distintos elementos correspondientes a cargas críticas tales como: sistemas de enfriamiento para transformadores, cambiadores de tomas, circuitos de interruptores (control, compresores y motores para accionar), equipos de calentamiento, relés de protección, alarmas, sistema contra incendios, equipos de comunicación, luces de seguridad y cargadores de baterías [2]. También incluye cargas no críticas tales como: iluminación de exteriores e interiores, tomas de uso general en exteriores e interiores, cargas destinadas al mantenimiento (equipos para tratamiento de aceite, equipos de transporte para gas SF₆, entre otros) y cargas destinadas a la construcción (soldadoras, taladros, remolques, entre otros) [2].

En este caso, los servicios auxiliares del patio de 765 kV contaban con alimentación principal por medio de los terciarios (20 kV) de 2 bancos de

autotransformadores, y dos sistemas de respaldo: 1) una línea de 34,5 kV (aproximadamente 600 m) alimentada desde el patio de 400 kV y 2) un generador diésel de 440 V con capacidad de 1.500 kVA.

La línea de 34,5 kV alimentada del patio de 400 kV se deriva de una barra alimentada por los terciarios (34,5 kV) de los 2 bancos de autotransformadores (ATR1 y ATR2) y un generador diésel de 1.500 kVA también como respaldo de acuerdo a la actualización de planos en sitio que se muestra en la Figura 2 [1]. La línea de interconexión entre patios funciona mediante la conexión por un supervisor de voltaje y un seccionador en el patio de 765 kV, mientras que el otro extremo de la línea se encuentra normalmente abierto por medio de un interruptor de bajo volumen de aceite a la barra de servicios auxiliares (en adelante SSAA), lo cual permite una ausencia de flujo de potencia cuando se opera con los otros medios de alimentación.

En la fecha del 06/10/2006 se deja fuera de servicio a la línea de SSAA que conecta los patios de 765 kV y 400 kV, debido a una falla en el autotransformador N°2 de 400 kV / 230 kV / 34,5 kV por una sobretensión que ocasionó el

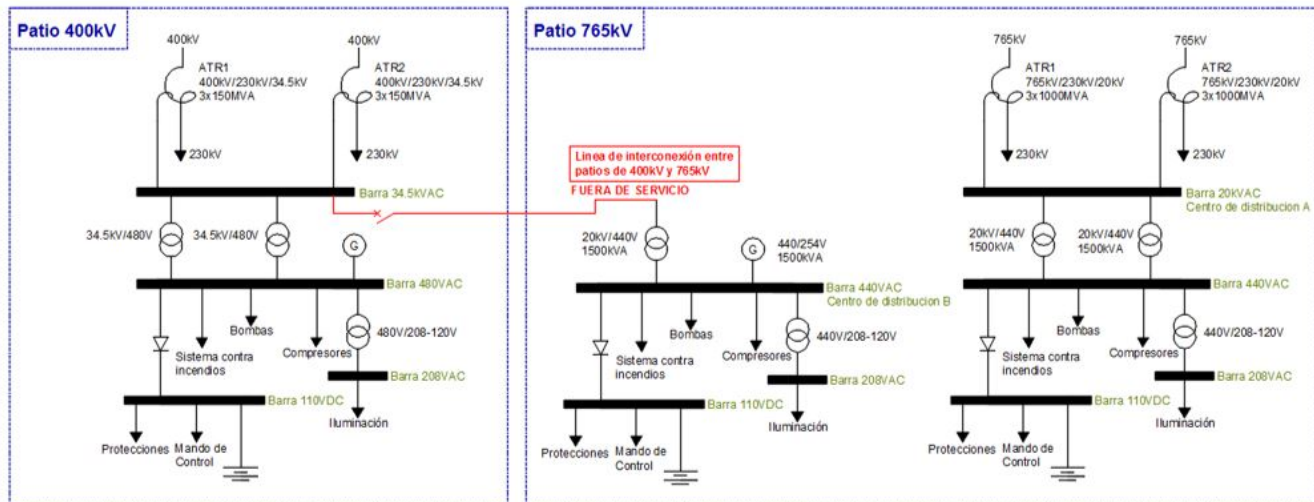


Figura 2: Diagrama de alimentación para SSAA de los patios de 765 kV y 400 kV [1]

deterioro de diversos elementos producto de un tiempo largo de despeje de la falla (30 ciclos) [3]. Una vez despejada la falla, los conductores y varios elementos (copas terminales, aisladores de vidrio, pararrayos, fusibles, conectores y contactos de seccionadores) presentaron daños.

La empresa también informó que realizó una prueba de aislamiento al transformador 34,5 kV / 440 V que conecta la línea con el centro de distribución B de los servicios auxiliares en 765 kV y este arrojó valores normales. Posteriormente, la falta de mantenimiento correctivo y el hurto en secciones subterráneas dejó desmantelada y fuera de servicio la línea.

Por su parte, la Universidad de Carabobo, a través de sus líneas de investigación y desde el año 2002, realiza una aplicación de software denominado SISTRANS, y desarrollado en la Escuela de Ingeniería Eléctrica, para realizar cálculos para el diseño de líneas de transmisión y distribución.

2. Metodología

Para realizar el diseño se siguió el siguiente diseño metodológico:

1. Fijación de criterios de diseño. De acuerdo a la información de los componentes de la red existente, el proyecto original, informes de fallas relacionadas con la puesta fuera de servicio de la línea 34,5 kV de interconexión

de servicios auxiliares de los patios de 765 kV y 400 kV en la S/E Arenosa, perfil longitudinal del terreno, datos climatológicos y la revisión en campo de las condiciones actuales de funcionamiento de equipos y componentes de la línea original se fijan criterios de diseño considerando los componentes que pueden ser reutilizados, los que deben ser desincorporados.

2. Cálculo del conductor. Para la especificación del conductor de la línea de 34,5 kV se consideran las normas de diseño, la potencia de la línea, ampacidad, el cálculo térmico, la caída de tensión, las distancias mínimas, y el cálculo mecánico del conductor.
3. Cálculo de los soportes. Para verificar las estructuras se consideran las normas de diseño, la carga vertical admisible, el vano máximo por carga, la carga longitudinal y transversal, el vano máximo por flecha máxima, se conforman la hoja de localización y de tensado, y se dibuja el tendido de acuerdo al perfil del terreno.
4. Validación de los resultados por medio de herramienta computacional SISTRANS.

Aunque el diseño y modelación de líneas de distribución en Venezuela está regida por las Normas de la empresa estatal distribuidora

Corpoelec, que establecen la metodología, los análisis, las técnicas, los materiales, entre otros, es importante considerar que aún el tema del diseño de líneas sigue vigente y abierto a propuestas de innovación.

En este caso, las decisiones del aprovechamiento de postes existentes tratando de reducir los esfuerzos mecánicos involucra la revisión de cómo organizar las líneas en los postes, cómo reducir las distancias sin afectar la seguridad, cuáles son las nuevas tecnologías en herrajes y arreglos, cómo mejorar las técnicas de mantenimiento con líneas energizadas, cómo aumentar los ciclos de vida de los activos, entre otros [4].

En este proyecto, aunque las limitaciones normativas se imponen para el diseño, el mercado de nuevos productos y la preparación técnica de los líderes de proyectos en la industria eléctrica [5], son determinantes en el desempeño de la misma.

3. Resultados

3.1. Criterios de diseño

Los criterios para el nuevo diseño de línea 34,5 kV para respaldo de energía de los servicios auxiliares de la S/E La Arenosa 765 kV son los siguientes:

1. Cumplimiento de Normas Cadafe – Corpoelec.
2. Aprovechamiento de las estructuras y soportes existentes en buen estado incluyendo los equipos de protección y soporte en el pórtico de 765 kV (Edelca) todo esto debido al mantenimiento realizado de fusibles, seccionadores, pararrayos y soportes dañados.
3. Diseño con base a la potencia del transformador de 1.500 kVA, 34,5 kV / 440 V ubicado en el patio de 765 kV.
4. Cambio del conductor en tramo subterráneo.
5. Cambio del interruptor de potencia.

3.2. Cálculo del conductor

La norma Cadafe Nº 53-87: Normas de diseño para líneas de alimentación y redes distribución. Normalización de calibres primarios y secundarios [6]. Especifica los calibres que pueden utilizarse para las condiciones de trabajo establecidas en este proyecto, se hace referencia al menor calibre permitido (AWG 2) y el mayor (AWG 4/0) así como el material del conductor.

La verificación de caída de tensión en la línea permite asegurar que el conductor seleccionado permita una variación de tensión dentro de los rangos establecidos por las normas pertinentes. Para un sistema de subtransmisión 34,5 kV la norma Cadafe 42-87 de Regulación de tensión en el sistema de distribución establece que la caída de tensión permitida no debe ser mayor al 3,5 % en media tensión [7]. Considerando un factor de potencia 0,8 inductivo en el sistema y una distancia de 700 m de línea. Los resultados del cálculo de caída de tensión, en las condiciones indicadas se muestran en la Tabla 1.

Tabla 1: Caída de tensión para conductores

Calibre AWG	Kd 10-3	%ΔV
2	0,0925	0,097125
1/0	0,0656	0,06888
2/0	0,0557	0,058485
3/0	0,0489	0,051345
4/0	0,0426	0,04473

Longitud = 0,7 km; kVA = 1.500

Ya que todos los calibres cumplen las condiciones de ampacidad y caída de tensión se decide seleccionar el conductor de menor calibre ya que este tendrá un menor costo con respecto a los demás, siendo AAAC 6201 T81 AWG #2.

3.3. Cálculo térmico del conductor

Se realizó mediante una hoja de cálculo en Excel el cálculo térmico del conductor para diferentes potencias transmitiéndose en la línea, dicho cálculo térmico está basado en el método de Westinghouse, el cálculo arrojó que para la potencia nominal de transmisión (1.500 kVA) se tiene una temperatura de operación en el conductor $T = 30,99^{\circ}\text{C}$.

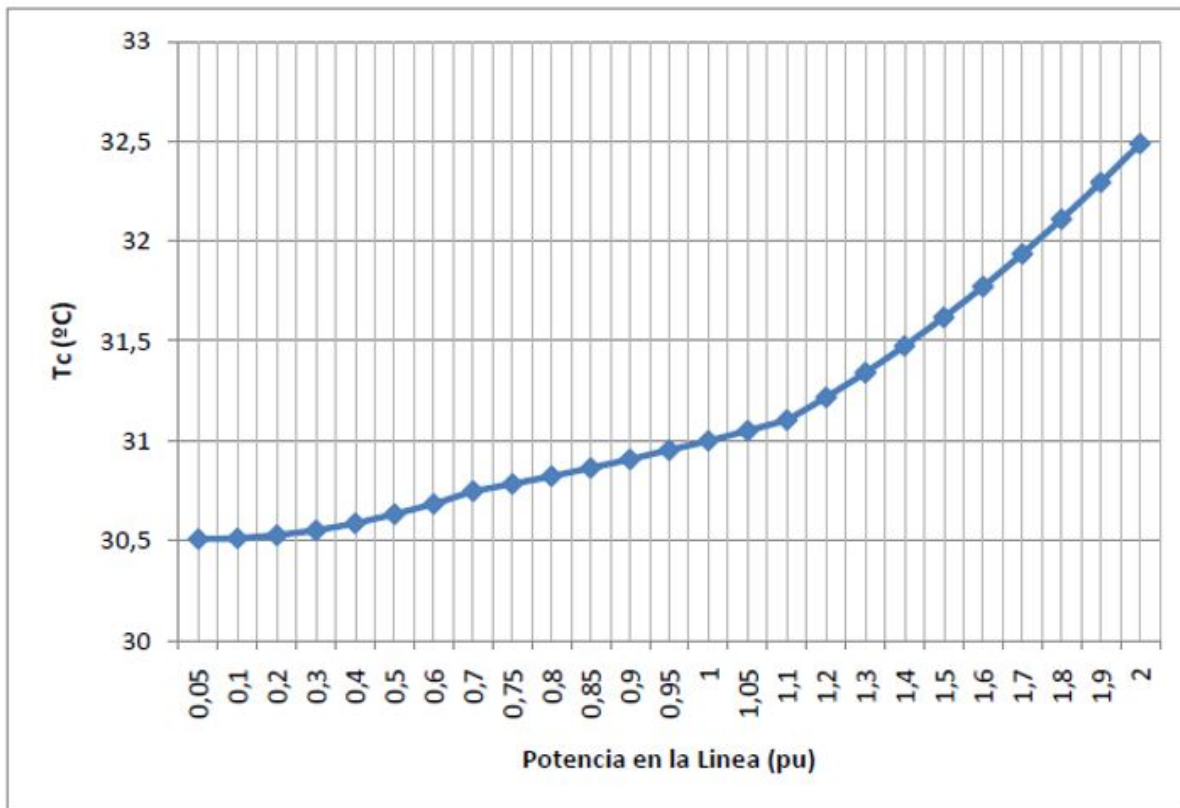


Figura 3: Cálculo térmico del conductor

Todo esto partiendo de la ecuación (1) de equilibrio térmico (Potencia absorbida = Potencia disipada), la ecuación resultante será [8]:

$$R \cdot I^2 + a \cdot S \cdot \left(\frac{A}{2}\right) = (w_r + w_c) \cdot A, \quad (1)$$

donde:

- I*: Corriente del conductor, (A)
- R*: Resistencia efectiva del conductor, (Ω/pie)
- a*: Coeficiente de absorción solar
- S*: Intensidad de radiación, (W/pulg²)
- A*: Área transversal del conductor por unidad de longitud, (pulg²/pie)
- w_r*: Potencia disipada por radiación, (W/pulg²)
- w_c*: Potencia disipada por convección, (W/pulg²)

3.4. Diseño de línea

En la norma de Cadafe NR apartado 2-04-04 se expresan las temperaturas a ser utilizadas en el cálculo mecánico de conductores para redes de distribución y subtransmisión. Como la línea de

34,5 kVA de interconexión de servicios auxiliares se encuentra aproximadamente a una altura de 440 m respecto al nivel del mar, se consideran las siguientes temperaturas: máxima = 60 °C, media = 28 °C y mínima = 10 °C.

La disposición existente en el tramo de estudio es una instalación en forma de triángulo con crucetas de 2,44 m colocada a 10 cm del tope del poste. Lo establecido en la norma de Cadafe 55-87 la separación mínima horizontal entre conductores de un mismo circuito viene dada por la expresión (2), donde se podrá obtener la siguiente flecha máxima:

$$S = 0,762 \cdot V + 3,68 \cdot \sqrt{fmax} \quad (2)$$

$$\rightarrow fmax = 607,6 \text{ cm.}$$

El cálculo mecánico del conductor fue realizado siguiendo lo establecido por la norma de Cadafe 55-87 Normas de diseño para las líneas de alimentación y redes de distribución: Cargas mecánicas [9] y se realiza para verificar el funcionamiento del mismo en condiciones climatológicas extremas y fundamentar la selección de postes y estructuras.

Sin embargo, el criterio de diseño establece usar las estructuras o postes existentes, por lo cual, las características de los postes existentes se presentan en la Tabla 2.

Tabla 2: Dimensiones de postes existentes

Altura	Lo	Diámetro		
		D1	D2	D3
Total (m)	EMP (m)	(mm)	(mm)	(mm)
12,2	1,8	219,1	168,3	139,7
Juntas		Espesor de Sección		
J1 (mm)	J2 (mm)	E1 (mm)	E2 (mm)	E3 (mm)
460	356	7	6,3	5,5
Carga en Cumbre (kg)	Peso (kg)	Longitud de Sección		
		L1 (m)	L2 (m)	L3 (m)
460	356	7	6,3	5,5

El grado de esbeltez mecánico para los postes de acero según lo establecido en la norma de Cadafe 449-05 [10], es mayor a 105. De esta manera, el poste seleccionado es: 40' o 12,2 m con secciones de 8" - 6 5/8" - 5 1/2", Ec. 437 kg, Peso: 355 kg. Esta selección se corresponde con lo instalado actualmente.

Para el cálculo de la fuerza vertical se consideran los soportes de 3 secciones (a una altura total de 12,2m de acuerdo a los postes existentes). Será calculada aplicando la fórmula de Euler (ecuación (3)), la cual permite determinar la fuerza crítica vertical admisible.

$$F_{cvadm} = \frac{K \cdot I_c \cdot E \cdot \pi^2}{\eta \cdot L^2 \cdot 100}, \quad (3)$$

donde:

F_{cv} : Fuerza crítica vertical máxima admisible, [kg]

I_c : Momento de inercia equivalente, [cm⁴]

E : Módulo de elasticidad del acero, [20 300 kg/cm²]

η : Coeficiente de seguridad, [2,5]

K : Coeficiente dependiente del modo de fijación de los extremos

Aunque no se realiza una comprobación vía excavación o calicata de las fundaciones existentes, en las Figuras 4, 5 y 6 se observan los planos de las fundaciones originales que muestra el cumplimiento con lo establecido en las normas.

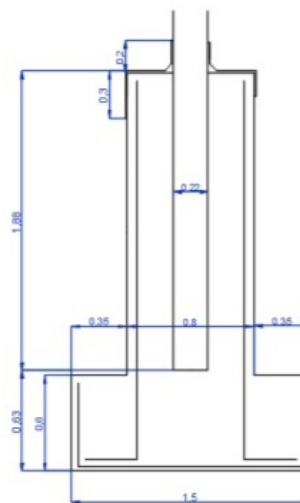


Figura 4: Detalles de postes 2, 3, 5 y 9

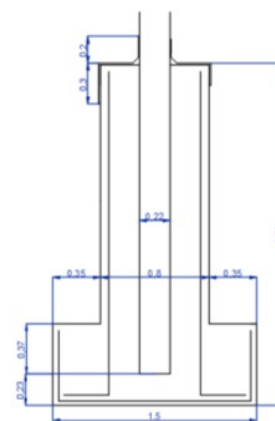


Figura 5: Detalles de postes 3, 6, 7 y 8

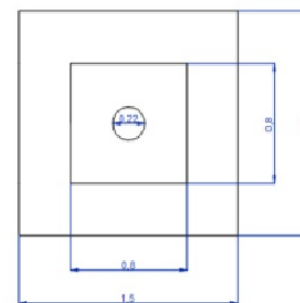


Figura 6: Vista en planta de postes 2, 3, 4, 5, 6, 7, 8 y 9

Se debe comprobar que se cumple la siguiente condición $F_v < F_{cvadm}$, donde F_v se obtiene considerando el peso de crucetas, herrajes, aisladores, poste, una persona, conductor y fuerzas

Tabla 3: Fuerzas vertical crítica admisible según sujeción

Tipo de sujeción	K	L	I1 (cm4)	I2 (cm4)	I3 (cm4)	Ie (cm4)	Fcvadm (kg)
2 Empotrado	4	0,5L	2625,75	1053,42	522,89	1731,05	149130,49
1 empotrado 1 articulado	2	0,7L	2625,75	1053,42	522,89	1731,05	38043,49
2 articulados	1	1L	2625,75	1053,42	522,89	1731,05	9320,66
1 libre 1 empotrado	0,25	1L	2625,75	1053,42	522,89	1731,05	2330,16

del viento. Que para el caso planteado da como resultado $FV = 934,33$ kg, menor que la fuerza vertical crítica admisible que se obtuvo en la Tabla 3.

En los postes de alineación se debe cumplir que la fuerza que ejerce el viento sobre los conductores (Fvc) sea menor o igual al esfuerzo útil disponible en cumbre del poste, donde el esfuerzo útil en la cumbre del poste ($E.U$) es equivalente al esfuerzo cumbre ($E.C$). Una vez conocidas estas condiciones, sabiendo ciertos parámetros constructivos del poste y valores climatológicos del ambiente como es la presión del viento máximo, valor que se encuentra normalizado por Cadafe como 55 kg/m^2 es posible calcular cuál será el vano máximo permitido, en la Tabla 4 se muestran los parámetros obtenidos en el proceso de cálculo y el vano máximo resultante.

Tabla 4: Vano máximo por carga del viento

Pvc (kg/m)	55
E.C (kg)	437
Fvp1 (kg)	126,184
Fvp2 (kg)	45,986
Fvp3 (kg)	38,172
Fvp (kg)	137,71
E.U (kg)	299,29
Smax	244,458

El vano máximo permitido por carga de viento en el conductor obtenido fue $S_{max} = 244$ m, se puede observar comparando con los vanos reales que se encuentran físicamente en el tramo de línea de estudio que todos los vanos cumplen esta condición, teniendo un vano máximo de 61 m.

Las cargas mecánicas longitudinales ocurren en la aparición de la rotura de un conductor, generando un desbalance en el equilibrio de tensiones sobre

el soporte. La norma Cadafe 55-87 establece las condiciones que deben cumplirse para considerar el soporte adecuado, los valores de tensión según el tipo de soporte (suspensión o amarre). Para que el soporte no se someta a deformaciones debe cumplir la ecuación (3.4):

$$\text{coef.de seguridad} = \frac{2,5 \cdot E.C}{E.R} \leq 1,3;$$

$$E.R = \sqrt{T^2 + E.C^2},$$

donde:

$E.C$: Esfuerzo en cumbre del poste, [kg]

$E.R$: Esfuerzo resultante en la cumbre, [kg]

T : Tensión reducida, [kg]

Los resultados para cada poste que se encuentran reflejados en la Tabla 5.

Una vez realizados todos los cálculos, se procedió finalmente a dibujar las catenarias que modelarían el tendido eléctrico para las condiciones establecidas, esto tiene como objetivo comprobar que no se irrespete la distancia mínima hacia el suelo (6,75 m). Dicho tramo de la plantilla representaría el segmento de catenaria similar al conductor físicamente una vez fuese tendido, implementando la ecuación (4):

$$y = C \cdot \cosh\left(\frac{x}{C}\right), \quad (4)$$

donde:

C : Parámetro en caliente, (m)

X : distancia en eje horizontal, (m)

Y : distancia en eje vertical, (y)

La distancia que existe en cada tramo se muestra en la Tabla 6, así como en la Figura 4 se encuentra el tendido del conductor, pudiendo observarse como

Tabla 5: Coeficientes de seguridad para cargas longitudinales

Poste	Amarre	Suspensión	E.C (kg)	T (kg)	E.R (kg)	Coef. Seguridad
2	X		437	135,875	457,636	2,387
3	X		437	135,875	457,636	2,387
4		X	437	67,9375	442,249	2,47
5	X		437	135,875	457,636	2,387
6		X	437	67,9375	442,249	2,47
7		X	437	67,9375	442,249	2,47
8		X	437	67,9375	442,249	2,47
9	x		437	135,875	457,636	2,387

en todos los tramos se cumple la condición de distancia mínima al suelo, línea representada de color verde.

Tabla 6: Longitud del tendido aéreo

Tramo	Long. Conductor (m)
S ₁₋₂	50,4
S ₂₋₃	52
S ₃₋₄	64,6
S ₄₋₅	64,4
S ₅₋₆	50,8
S ₆₋₇	50,3
S ₇₋₈	50,5
S ₈₋₉	60,7
S ₉₋₁₀	27,5

La Figura 7 representa el tendido del conductor. Se puede observar como la línea verde hace referencia a la distancia mínima, la cual no debe ser traspasada por el conductor que se encuentra dibujado en azul. En dicha figura se puede observar como el diseño cumple lo establecido con la norma.

En correspondencia con el trazo demarcado en rojo de la Figura 2, el proyecto comprende lo siguiente:

- Instalación de tramo trifásico subterráneo de cable 1P MV-90 con aislamiento XLPE para 36 kV calibre #2 AWG en bancada existente, siguiendo el Std 525-2016 [11].
- Instalación de red trifásica aérea con conductor de aleación 6201 #2 AWG.
- Red de postes de acuerdo a Figura 7 donde se aprovechan los existentes.
- Para el tendido de la línea y en correspondencia con la Tabla 5, se consideran los

herrajes y materiales establecidos para cada tipo de sujeción, ya sea amarre o alineación, establecidos de acuerdo a las Normas Cadafe 274-91 [12], 269-91 [8], 276-91 [13], 265-88 [14], 264-91 [15], 259-91 [16] y 26-01 [17].

- Instalación de seccionador de línea en 34,5 kV de acuerdo a Norma Cadafe 39-87 [18].
- Instalación de aisladores para línea aérea de acuerdo a la Tabla 5, ya sea en amarre, con aisladores de suspensión, o alineación, con aisladores de espiga.
- Instalación de puesta a tierra y pararrayos en poste con seccionador según Norma Cadafe 46-87 [19].
- Suministro e instalación de interruptor de potencia de 36 kV de SSAA.

Se estima que la reconstrucción de la línea requiera una inversión aproximada de \$10.300,00 dólares americanos desde el pórtico de La Arenosa 765 kV hasta el pórtico de La Arenosa 400kV incluyendo el interruptor de potencia en 36kV, y se estima que el tiempo de ejecución de la obra sea de 10 días sin considerar los tiempos de procura de materiales.

4. Discusión de resultados

Partiendo de la revisión del proyecto original y la revisión de campo se fijaron los criterios de diseño de la línea de 34,5 kV para los servicios auxiliares donde el principal fundamento de diseño fue la aplicación de las Normas Cadafe – Corpoelec. Uno de los criterios fundamentales para iniciar los

donde se observa como nunca se excede el esfuerzo cumbre del poste. Con el cálculo mecánico del conductor se verificaron resultados de tensión y flechas para diferentes estados ambientales y condiciones de trabajo, comprobando que se respeta la distancia mínima entre conductor a tierra especificado por Cadafe.

A pesar de consolidar una propuesta técnica y económica, este diseño no soluciona algunos problemas observados que pueden seguir incidiendo en los sistemas de respaldo de energía tales como: la falta de vigilancia y el vandalismo, falta de personal calificado y alta rotación del personal, falta de planificación, poca inversión y falta de mantenimiento.

Para este proyecto los datos suministrados sostienen su validez cualitativa y cuantitativa en la información suministrada por la propia empresa distribuidora y el método aplicado sostiene su validez en las Normas nacionales Cadafe. Aunque el software SISTRANS, desarrollado intramuros, ayuda a los cálculos y validación de resultados, requiere mejoras en la interfaz gráfica, en la disponibilidad y usabilidad para que pueda convertirse en producto masificable.

Finalmente, aunque la investigación se justifica en la pertinencia social que busca resolver problemas del sector eléctrico y su adscripción académica para aportar a las líneas de investigación, no puede garantizarse el impacto de la solución si no logra vincularse a la decisión de la ejecución proyecto.

5. Conclusiones

El propósito del estudio de reconstrucción de la línea 34,5 kV que ofrece el respaldo de energía de los servicios auxiliares de la S/E La Arenosa 765 kV muestra que el conductor seleccionado es aleación 6201 #2 AWG de acuerdo al cálculo por ampacidad, cálculo térmico, por caída de tensión y por cálculo mecánico. A diferencia del establecido en el diseño original (Aleación 6201 4/0 AWG), esta investigación determina que no es necesario tal sobredimensionamiento.

Esto mejora las condiciones para la reutilización del soporte estructural existente determinada en el criterio de diseño sin alteración de la ruta

respecto al diseño original. Desde el punto de vista metodológico los datos utilizados para los cálculos fueron validados y revisados con la empresa distribuidora y la metodología responde a lo establecido en las Normas vigentes.

En este punto, es importante considerar que las innovaciones para el diseño de líneas y, específicamente, el rediseño considerando aprovechamiento de estructuras y soportes existentes, pueden aumentar el abanico de soluciones y costos cuando se incorporan nuevos materiales y el ejercicio del liderazgo para el diseño e implementación de nuevas normas técnicas.

En cuanto al uso del software SISTRANS como técnica de procesamiento y análisis de la información se concluye que requiere nuevos proyectos de mejora en la interfaz gráfica, en la disponibilidad y usabilidad. Aunque esta herramienta fue aplicada en una línea de 34,5 kV; aún queda por validar su uso a otros niveles de tensión.

Finalmente, esta investigación también permitió determinar que la vinculación de la Escuela de Ingeniería Eléctrica de la Universidad de Carabobo con la empresa distribuidora para el aporte a las soluciones técnicas a los problemas del sistema, aún no se consolida de manera sostenida ante la cantidad de problemas observados. En este caso, los proyectos para solucionar problemas de líneas de distribución y subtransmisión pueden convertirse en una línea de investigación específica de la Universidad de Carabobo.

6. Referencias

- [1] S. Abreu and A. Otero, “Diseño de línea de distribución en 34.5 kV de servicios auxiliares entre subestaciones de transmisión. Caso estudio: tramo S/E La Arenosa 765 kV a S/E La Arenosa 400 kV. CORPOELEC,” Trabajo especial de grado para optar al título de Ingeniero Electricista, Facultad de Ingeniería, Universidad de Carabobo, Venezuela, 2018.
- [2] IEEE P1818/D6.2, *IEEE Draft Guide for the Design of Low Voltage Auxiliary Systems for Electric Power Substations*, IEEE Std, Jan. 2017.
- [3] J. Perez, “Falla de la línea Cadafe-Edelca a 34,5 kV, asociada a la alimentación de respaldo de los servicios auxiliares de corriente alterna de S/E La Arenosa, ocurrida el 08-10-2006,” Cadafe-Edelca, Venezuela, Informe interno de Falla, 2006.

- [4] R. G. Rodriguez, “Innovative new live line tool and framing designs,” in *2017 IEEE 37th Central America and Panama Convention (CONCAPAN XXXVII)*, Managua, 2017, pp. 1–6.
- [5] G. Gela, “IEC TC78 “Live working”: Historical overview, new structure, programme of work, and market relevance,” in *2016 IEEE PES 13th International Conference on Transmission & Distribution Construction, Operation & Live-Line Maintenance (ESMO)*, Columbus, OH, USA, 2016, pp. 1–4.
- [6] Cadafe , 53–87, *Normas de diseño para líneas de alimentación y redes de distribución: normalización de calibres primarios y secundarios sistema de distribución*, Cadafe, Venezuela, 1987.
- [7] Cadafe , 42–87, *Normas de diseño para líneas de alimentación y redes de distribución: regulación de tensión en los sistemas de distribución*, Cadafe, Venezuela, 1987.
- [8] Cadafe , 269–91, *Norma herrajes de distribución: asiento para crucetas*, Cadafe, Venezuela, 1991.
- [9] Cadafe , 55–87, *Normas de diseño para líneas de alimentación y redes de distribución: cargas mecánicas*, Cadafe, Venezuela, 1987.
- [10] Cadafe , 449–05, *Norma poste de secciones tubulares de acero*, Cadafe, Venezuela, 2005.
- [11] IEEE Std 525-2016 (Revision of IEEE Std 525-2007), *IEEE Guide for the Design and Installation of Cable Systems in Substations*, IEEE Std, 2016.
- [12] Cadafe , 274–91, *Norma herrajes de distribución: cruceta*, Cadafe, Venezuela, 1991.
- [13] Cadafe , 276–91, *Norma herrajes de distribución: perno roscado a todo lo largo*, Cadafe, Venezuela, 1991.
- [14] Cadafe , 265–88, *Norma herrajes de distribución: abrazadera universal*, Cadafe, Venezuela, 1988.
- [15] Cadafe , 264–91, *Norma de herrajes de distribución: pletina*, Cadafe, Venezuela, 1991.
- [16] Cadafe , 259–91, *Norma herrajes de distribución: grillete*, Cadafe, Venezuela, 1991.
- [17] Cadafe , 26–01, *Norma de herrajes de distribución: espiga para aisladores*, Cadafe, Venezuela, 2001.
- [18] Cadafe , 39–87, *Norma especificación técnica para seccionadores y sus accesorios*, Cadafe, Venezuela, 1987.
- [19] Cadafe , 46–87, *Normas de diseño para líneas de alimentación y redes de distribución: protección contra sobretensiones*, Cadafe, Venezuela, 1987.

Prof. César Rodolfo Ruiz[†]



Ingeniero Electricista de la Universidad Central de Venezuela (1964). Se desempeñó como Jefe Técnico, Jefe de Distribución, Supervisor de Mediciones, Gerente de Producción y Gerente de Región en Cadafe, y también fue Director en la Energía Eléctrica de Barquisimeto (Enelbar).

Fue Profesor de la Universidad de los Andes (1967 a 1969) y de Universidad de Carabobo (1985 a 2019).

En la Universidad de Carabobo fue Profesor del Departamento de Potencia de la Escuela de Ingeniería Eléctrica donde trabajó en las cátedras de Transmisión de Energía, Sistemas de Distribución y Protecciones Eléctricas. Fallece en funciones como Director de la Escuela de Ingeniería Eléctrica el 01/09/2019.

Neuropathology Detector in EEG using Higher Order Statistics and Deep Learning

César Seijas ^{*,a} , Sergio Villazana ^a , Guillermo Montilla ^b , Egilda Pérez ^a ,
Ricardo Montilla ^b 

^aCentro de Procesamiento de Imágenes, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela.

^bYttrium-Technology Corp., Panamá, Panamá.

La selección de este artículo fue realizada en el marco de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica “Prof. César Rodolfo Ruiz” Octubre 2020, siendo la evaluación, arbitraje, aceptación y edición a cargo de Revista Ingeniería UC.

Abstract.- This article presents a neuropathology detector, based on the patient’s electroencephalogram (EEG). Detection is based on HOSA (“High Order Statistical Analysis”) image classification of higher order statistics derived from time series corresponding to EEG of human patients. The classifier is a DL model (“Deep Learning”) with the pretrained CNN (“Convolutional Neural Network”) architecture: Inception. The CNN training and test set are HOSA images of non-linear and non-Gaussian segments, of signals corresponding to the selected channel of the EEG of patients with neuropathologies (specifically, epilepsy) or healthy. The performance of the classifier is very satisfactory, presenting an accuracy of approximately 94 % in the detection of epilepsy.

Keywords: EEG; higher order statistics; deep learning; pre-trained convolutional neural network Inception.

Detector de Neuropatologías en EEG usando Estadísticas de Orden Superior y Aprendizaje Profundo

Resumen.- En el presente artículo se presenta un detector de neuropatologías, a partir del electroencefalograma (EEG) del paciente. La detección se basa en la clasificación de imágenes de HOSA (siglas en inglés para análisis de estadísticas de orden superior o “High Order Statistical Analysis”) derivadas de series de tiempo correspondientes a EEG de pacientes humanos. El clasificador es un modelo de aprendizaje profundo DL (“Deep Learning”) con la arquitectura de la CNN (Red Neuronal Convolutiva o “Convolutional Neural Networks”) pre-entrenada: “Inception”. El conjunto de entrenamiento y prueba de la CNN son imágenes de HOSA, que representan los cumulantes de tercer orden de segmentos no lineales y no gaussianos, de señales correspondientes al canal seleccionado del EEG de pacientes con neuropatologías (específicamente, epilepsia) o sanos. El desempeño del clasificador es muy satisfactorio, presentando una exactitud de aproximadamente 94 % en la detección de epilepsia.

Palabras clave: EEG; estadísticas de orden superior; aprendizaje profundo; red neuronal convolutiva pre-entrenada Inception.

Recibido: 22 de octubre, 2020.

Aceptado: 14 de diciembre, 2020.

1. Introducción

El procesamiento de señales electrofisiológicas, con fines diagnósticos, ha tenido grandes avances

en el campo de la medicina y actualmente la incorporación de la inteligencia artificial (IA), y particularmente aprendizaje profundo DL (siglas en inglés para “Deep Learning”) está produciendo importantes desarrollos en el tratamiento de desórdenes neurológicos [1, 2, 3]. En general, las señales se captan mediante sensores que traducen los fenómenos electrofisiológicos, en débiles potenciales eléctricos, que posterior a un

* Autor para correspondencia:
Correo-e: cseijas@uc.edu.ve (C. Seijas)

adecuado procesamiento analógico, se convierten en señales digitales compatibles con análisis mediante el computador.

El Electroencefalograma (EEG) es el registro de las señales de la actividad electrofisiológica de las células cerebrales, conocidas como neuronas. Esta actividad es el resultado de la superposición de múltiples impulsos eléctricos (potencial de acción) [4, 1], generados en las neuronas, producto del intercambio iónico a través de la membrana celular (proceso de difusión iónica). El EEG se obtiene mediante la colocación de un conjunto de electrodos en el cráneo, sobre la región del cuero cabelludo. El sistema internacional de ubicación de electrodos 10-20 [5] genera un mapa de puntos separados 10 o 20 %, del área total bajo registro, con una nomenclatura asociada al nombre de la región craneal cubierta.

La interpretación correcta del EEG, por parte del experto, permite detectar y diagnosticar diferentes neuropatologías o estados neurológicos tales como: alteraciones del sueño, epilepsia, nivel de anestesia, presencia de actividad cerebral o incluso muerte. La epilepsia es una neuropatía muy común, que se manifiesta en repentinas ráfagas de actividad eléctrica, no controlada, en algunas regiones de la corteza cerebral. Los pacientes epilépticos sufren de ataques recurrentes, con manifestaciones que varían desde leves sensaciones anormales, hasta cambios impredecibles que alcanzan la pérdida inmediata de la consciencia y convulsiones [6]. Esta neuropatología afecta aproximadamente al 1 % de la población mundial. Las Figuras 1a y 1b, generadas con el software para procesamiento de señales electroencefalográficas EEGLAB [7], corresponden a segmentos de señales de los canales de EEG, para un paciente de control (sin la neuropatía) y uno epiléptico, respectivamente.

El objetivo del presente estudio, es desarrollar un nuevo método, para la detección de esta neuropatología, a partir de información derivada del EEG. Lo anterior, en aras de mejorar la calidad de vida de un paciente epiléptico o incluso de personas que desconocen de su predisposición a padecer de este mal, En el presente artículo se describe el desarrollo de un sistema automatizado para detectar epilepsia, a partir del EEG, usando

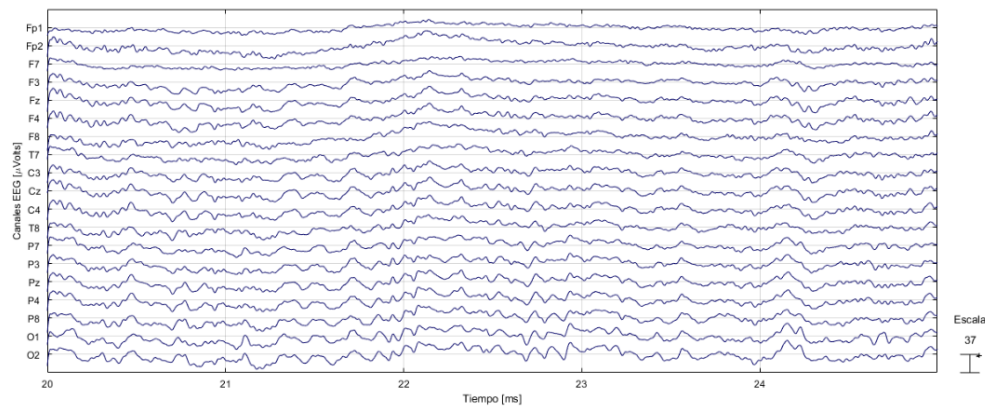
herramientas de IA y HOSA [8].

El sistema basado en DL, usa un clasificador construido con la CNN pre-entrenada Inception [9], pero especializado, para detección de epilepsia en el EEG. El clasificador discrimina rasgos subyacentes en imágenes HOSA, obtenidas a partir de las señales de EEG de pacientes epilépticos y de control. Como se afirma en estudios previos y se verifica, en este estudio, las señales de EEG son series de tiempo, altamente no lineales y no gaussianas, que, en consecuencia, exhiben estadísticas de orden superior. En el caso específico, de las estadísticas: momentos y cumulantes de orden mayor al tercero, estas métricas se manifiestan como señales hiperdimensionales. La estrategia planteada, en el sistema desarrollado, es detectar epilepsia en los EEG, usando estas señales, que son bidimensionales para los cumulantes de tercer orden [8, 10, 11] y debidamente etiquetadas como epilépticas o de control, como las imágenes de entrenamiento del clasificador.

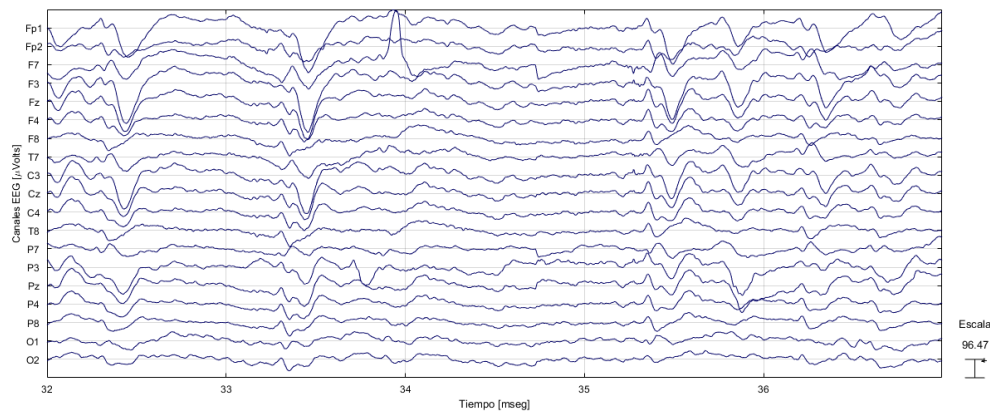
2. Antecedentes

Existen numerosos e interesantes estudios, que han reportado aplicaciones de DL en EEG. A continuación, se mencionan algunos trabajos en el asunto, objeto del presente artículo: Craik *et al.* [12], hacen una revisión de la literatura, sobre aplicaciones con DL, en el proceso de clasificación de EEG. La consulta ubicó y analizó 90 publicaciones, en bases de datos de sitios especializados como son: “Web of Science” y “PubMed”. En su revisión, presentan, para cada documento estudiado, aspectos fundamentales como son: arquitectura de la red empleada (CNN, redes recurrentes, etc.), tipo de tarea abordada por la red, esto es: detección de neuropatías, clasificación de emociones, predicción de ataques, entre otras posibles tareas, y concluyen presentando, para cada caso analizado, los logros y resultados alcanzados en cada publicación consultada.

Los autores, Raghavendra *et al.* [3], también presentan una revisión exhaustiva, pero específicamente, en el tema del diagnóstico automatizado mediante IA, aplicada a señales e imágenes



(a) Canales de EEG para un paciente de control



(b) Canales de EEG para un paciente epiléptico

Figura 1: Canales de EEG, para (a) un paciente de control y (b) un paciente epiléptico

electrofisiológicas, de 5 trastornos neurológicos seleccionados; la revisión abarca las últimas 2 décadas. Las neuropatías consideradas son: epilepsia, enfermedad de Parkinson, enfermedad de Alzheimer, esclerosis múltiple y accidente cerebrovascular isquémico. En el estudio, los autores analizan y comparan, los diferentes métodos de extracción y selección de rasgos, las técnicas de reducción de dimensionalidad usadas, y las técnicas de clasificación.

En [13], Golmohammadi *et al.*, describen el desarrollo de un sistema automatizado de análisis de EEG, de alto rendimiento, basado en principios de aprendizaje de máquinas ML (para “Machine Learning”) y manejo de gran volumen de datos (“big data”). El sistema desarrollado es una arquitectura híbrida, que incorpora contexto temporal y espacial; integrando HMM (para “Hidden Markov Models” o modelos ocultos de

Márkov) para la decodificación secuencial de eventos en el EEG, y postprocesamiento espacial, basado en DL (CNN). Los modelos del sistema, fueron entrenados y evaluados con la base de datos de EEG, TUH (“Temple University Hospital”). Los autores reportan que el sistema registra y clasifica patrones de interés clínico, de la actividad cerebral, como son: picos y ondas agudas, descargas epileptiformes periódicas generalizadas, descargas epileptiformes periódicas lateralizadas; además de detectar movimiento de los ojos, artefactos y actividad de fondo [2, 6]. Según los resultados, presentados en el artículo, el sistema exhibe una sensibilidad por encima del 90 % mientras mantiene una especificidad por debajo del 5 %.

El trabajo de Gegein *et al.*, presentado en [14], también corresponde a una plataforma híbrida (como [13]), pero en este caso usada como clasificador para discriminar entre EEG patológico

y no patológico. La plataforma usa una TCN (siglas en inglés para “Temporal Convolutional Network”) como extractora de rasgos secuenciales y una CNN, para rasgos espaciales; la base de datos usada, similar a [13] también es TUH. La TCN, usa como vector de rasgos, los cálculos derivados de aplicar las operaciones: Transformada Discreta de Fourier, Transformadas Wavelet Continua y Discreta y características de conectividad entre electrodos basadas en la Transformada Hilbert; las operaciones se aplicaron a segmentos de un canal seleccionado del EEG. Los autores manifiestan, que los resultados reportados, colocan el marco de decodificación propuesto, basado en rasgos, al mismo nivel que las redes neuronales profundas de última generación. El rango de precisión, entre ambos, es del 81 al 86 %.

Mahmoodian *et al.* [10], predicen ataques epilépticos, con un modelo a ML, a partir de diez características multivariantes no lineales, extraídas del biespectro cruzado [8]. Las estadísticas mencionadas, se emplearon, como vectores de entrenamiento de una SVM (siglas en inglés para “Support Vector Machines”) [11, 15], que discrimina estados preictales, de los interictales [6]. El predictor a SVM, en referencia, obtuvo una sensibilidad del 100 % y una FPR False Positive Rate” o tasa de falsos positivos) promedio de 0,044; utilizando el conjunto de datos de “Predicción de ataques epilépticos de Frieiburg” (University of Freiburg, 2008) La alta sensibilidad y muy bajo FPR, en comparación con otros estudios, evidencian la excelente capacidad del método, basado en HOSA, para analizar las señales de EEG epilépticas.

En [16], Yean *et al.*, estudiaron la no gaussianidad en la señal de EEG emocional (alegría, desagrado o neutro) de pacientes con accidente cerebrovascular y pacientes de control. La no gaussianidad se determinó, mediante el cálculo de las métricas estadísticas [8]: asimetría (skewness”) y kurtosis. La función de distribución estimada, del EEG emocional, fue simétricamente no gaussiana tanto para los grupos de accidente cerebrovascular como para los de control. En particular, se encontró que, la distribución de EEG no gaussiana, es más frecuente en pacientes de control, que en aquellos

con accidente cerebrovascular.

En [11], el autor clasifica EEG en epilépticos o normales, con un clasificador a SVM, entrenado con vectores conformados por dos rasgos extraídos de la respuesta frecuencial de orden superior: la bicoherencia y la máxima potencia. Los rasgos se obtuvieron, mediante la aplicación de HOSA, a la señal del canal seleccionado, de los EEG, de los pacientes en estudio. El criterio de selección de un canal del EEG, fue elegir el canal con máxima potencia, en toda la banda de actividad electrofisiológica cerebral [2]. La exactitud lograda con el clasificador fue de 92 %.

3. Fundamentos teóricos

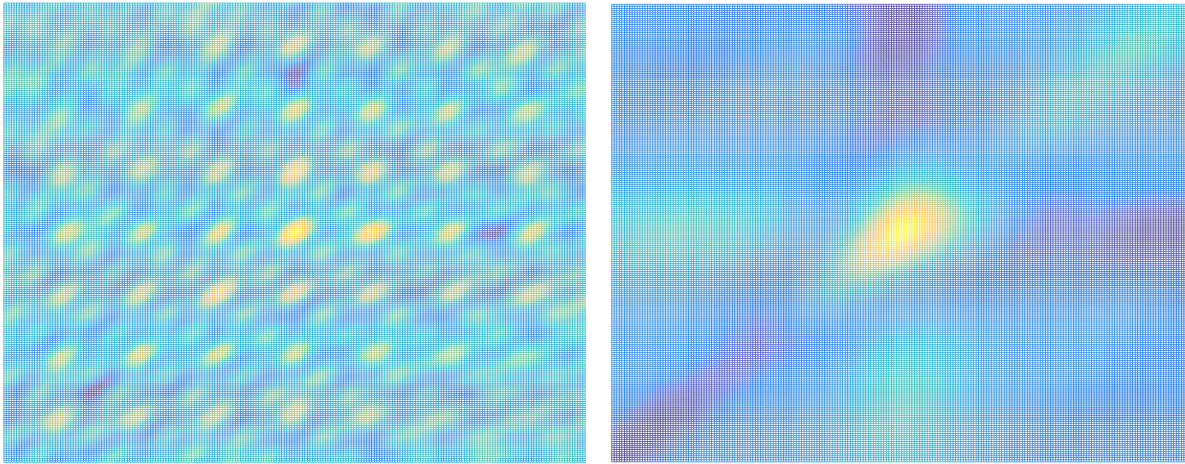
El objetivo de este artículo es el desarrollo de un sistema, que usa IA, en el proceso de detección de neuropatías, específicamente epilepsia, a partir de EEG. El sistema usa un clasificador basado en la CNN pre-entrenada Inception, que se debe entrenar (especializar) en la detección de rasgos epilépticos en un EEG.

El conjunto de datos de entrenamiento corresponde a señales bidimensionales, convertidas a imágenes y obtenidas del procesamiento, basado en HOSA, de segmentos, de la serie de tiempo que representa la actividad eléctrica de un canal (de los múltiples canales) del EEG. El uso de HOSA trae a colación, una discusión importante, en el tema del procesamiento de las señales del EEG, la cual es: deben estas señales, ser estudiadas, con base a una naturaleza lineal y determinística, o deben ser procesadas como sistemas complejos y caóticos, o si, más bien, deben considerarse predominante estocásticas.

Con base a lo último expuesto, es preciso validar si los segmentos usados para la generación de las imágenes de entrenamiento, son efectivamente no gaussianas y no lineales [18, 17]. En esta sección, se discuten los fundamentos teóricos de HOSA, de la prueba de linealidad y gaussianidad de señales y de la arquitectura CNN Inception.

3.1. Análisis no lineal utilizando estadísticas de orden superior

Las señales gaussianas y lineales se caracterizan, completamente, con los momentos estadísticos



(a) Cumulante de tercer orden de un canal de EEG, para un paciente de control (b) Cumulante de tercer orden de un canal de EEG, para un paciente epiléptico

Figura 2: Cumulantes de tercer orden de un canal de EEG

de primer y segundo orden (media y varianza). Las señales de EEG son altamente no lineales y no gaussianas [10, 16, 11, 17]. Los momentos y cumulantes se definen por las siguientes expresiones:

$$C_{2x}(k) = E\{x^*(n)x(n+k)\}, \quad (1)$$

$$C_{3x}(k,l) = E\{x^*(n)x(n+k)x(n+l)\}, \quad (2)$$

$$\begin{aligned} C_{4x}(k,l,m) = & E\{x^*(n)x(n+k)x(n+l)x^*(n+m)\} \\ & - C_{2x}(k)C_{2x}(l-m) \\ & - C_{2x}(l)C_{2x}(k-m) \\ & - M_{2x}^*(m)M_{2x}(k-l), \end{aligned} \quad (3)$$

$$M_{2x}(m) = E\{x(n)x(n+m)\}. \quad (4)$$

Las ecuaciones (1), (2) y (3), corresponden a los cumulantes de segundo, tercer y cuarto orden respectivamente, mientras que la ecuación (4), define al momento de segundo orden $C_{2x}(m)$, que aparece en la ecuación (3); el superíndice * en $x^*(n)$ denota la señal $x(n)$ compleja conjugada, y los índices: k, l, m son retardos en cada dimensión espacial. Obsérvese que $C_{2x}(m)$ es igual a $C_{2x}(k)$ en el caso de series de tiempo reales [8, 10] y,

además, para retardo $k = 0$, $C_{2x}(k)$ corresponde a la varianza de la señal.

En la Figura 2, se puede apreciar, que las estructuras, que representan el campo escalar definido por el cumulante de tercer orden $C_{3x}(k,l)$, tienen comportamientos visiblemente muy diferentes. Mientras que la imagen del cumulante de la señal de un paciente de control, mostrada en la Figura 2a está conformada por una estructura de puntos de máxima energía (color amarillo) distribuidos con alta periodicidad; la estructura, en el caso de la señal de un paciente epiléptico, Figura 2b, es una región de máxima energía, concentrada en el centro del rectángulo de la imagen.

Esta diferencia, en las estructuras de la imagen de $C_{3x}(k,l)$, las percibe la CNN Inception, entrenada como detector de epilepsia, como rasgo discriminador de la neuropatía en un EEG.

3.2. Prueba de linealidad y gaussianidad

Una prueba estadística, para comprobar la linealidad y gaussianidad de señales, es el algoritmo de Hinich [18]. El algoritmo se basa en detectar la condición de oblicuidad (asimetría) diferente de cero en la función de distribución: esta condición se mide con la estadística “skewness” y se fundamenta en, el hecho de, que, en un proceso gaussiano, los cumulantes de orden mayor a dos son

cero y, en consecuencia, también lo son el bispectro y la bicoherencia.

Entonces, existe la hipótesis nula de no gaussiano si se determina que el bispectro es distinto de cero; por otro lado, si además la bicoherencia, no es constante, se debe concluir que el proceso es no lineal. Swami *et al.* [8], construyeron una herramienta de software para uso académico, que denominaron HOSA, desarrollada con el software matemático Matlab, que, entre muchas otras más funciones, implementa el algoritmo de Hinich, con estimaciones consistentes y no sesgadas de la bicoherencia y el bispectro.

3.3. Red Neuronal Convolutiva Inception

La CNN Inception se basa en la aplicación del módulo Inception [9]. La característica clave del módulo Inception, utilizada en la implementación del modelo GoogLeNet (ganador del concurso ILSVRC 2014), se muestra en la Figura 3. El modelo GoogLeNet consta de 22 de estos módulos.

El principio básico detrás de la estructura del módulo Inception se inspira en el hecho, de que, a medida, que aumenta el número de capas de una CNN, así como su ancho, su precisión también aumenta. Sin embargo, la implantación de este principio, se refleja en un incremento creciente de los recursos computacionales, que, por supuesto, son limitados; además, crece proporcionalmente el número de parámetros a entrenar, lo que aumenta la tendencia a sobreajustes (degrada la capacidad de generalización) [9, 15].

El módulo Inception tiene como objetivo satisfacer los requisitos de aumentar las capas convolucionales, horizontalmente (ancho), en lugar de aumentar el tamaño de la red en profundidad, extrayendo características en cada capa a diferentes niveles de resolución (filtros de convolución).

Como puede observarse en las Figuras 3a y 3b, el tensor de entrada, al módulo Inception, es procesado por varias ramas convolucionales paralelas cuyas salidas luego se fusionan en un único tensor de salida. Para evitar problemas de alineación de imágenes de salida, de procesos intermedios, la implementación de la arquitectura Inception, está restringida a tamaños de filtro 1×1 , 3×3 y 5×5 . La salida de un módulo es

un solo tensor de salida, que es una combinación de los bancos de filtros de salida, concatenados y forma la entrada a la siguiente etapa. El módulo de versión de dimensiones reducida (Figura 3b) disminuye la salida del vector concatenado, al colocar en cascada al bloque “pooling” [9, 15], un filtro adicional 1×1 ; el efecto de esta variante, si bien aumenta, el tiempo de proceso con el nuevo filtro, se adiciona una función de activación extra “Relu” [9, 15], en el bloque convolutivo, con los beneficios implícitos, de agregar no linealidad.

En la Figura 4, se muestra la arquitectura completa de la implementación de GoogLeNet, los dos últimos bloques (a la derecha), identificados en la leyenda como: “Fully connected”, de color violeta, y “Softmax”, de color rojo, constituyen, lo que se conoce como etapa clasificadora, mientras que el resto, se le denomina etapa o base convolutiva. La base convolutiva es la responsable de la extracción de rasgos de la imagen; mientras, que la etapa clasificadora es un perceptrón multicapa [15], de dos capas.

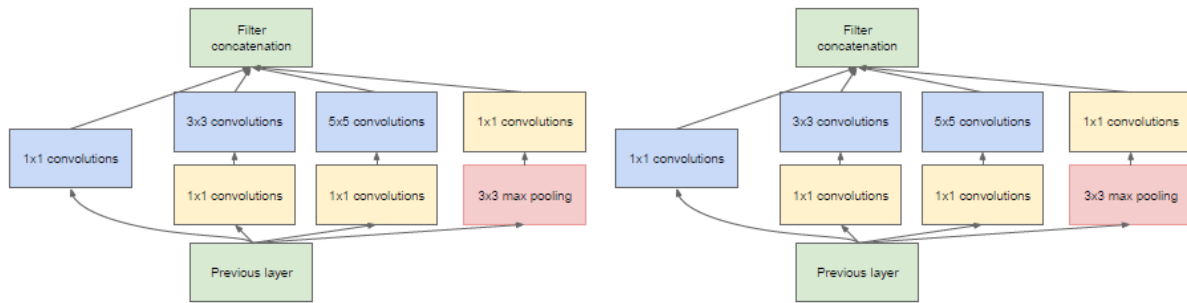
4. Metodología

En la presente sección se describe la metodología, la cual abarca fundamentalmente dos aspectos: procesamiento del EEG y diseño del clasificador.

4.1. Procesamiento del EEG

El procesamiento del EEG se inicia con la obtención de la base de datos de EEG a estudiar, continúa con el procesamiento y selección del canal y finaliza con la generación del conjunto de imágenes de entrenamiento al clasificador. A continuación, se describe cada tarea:

1. Obtención de la base de datos a estudiar, esto es, acopio de EEG de pacientes diagnosticados epilépticos y EEG de pacientes neurológicamente sanos. Con esta tarea, se define una base de datos, de dos clases: EEG epilépticos y EEG de control.
2. Procesamiento y selección del canal: En la fase de procesamiento, a cada EEG, se le aplicó filtrado digital para eliminación y/o reducción de artefactos [2, 5], mediante



(a) versión básica

(b) versión con dimensiones reducidas

Figura 3: Diagrama de bloques del módulo Inception

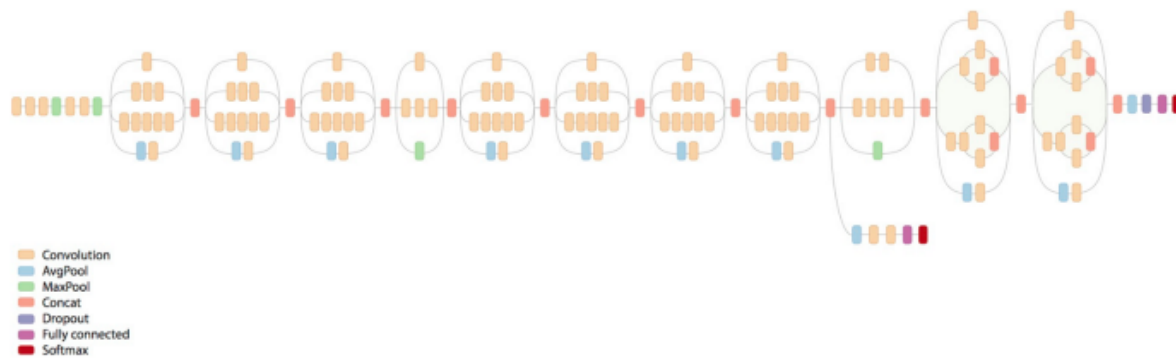
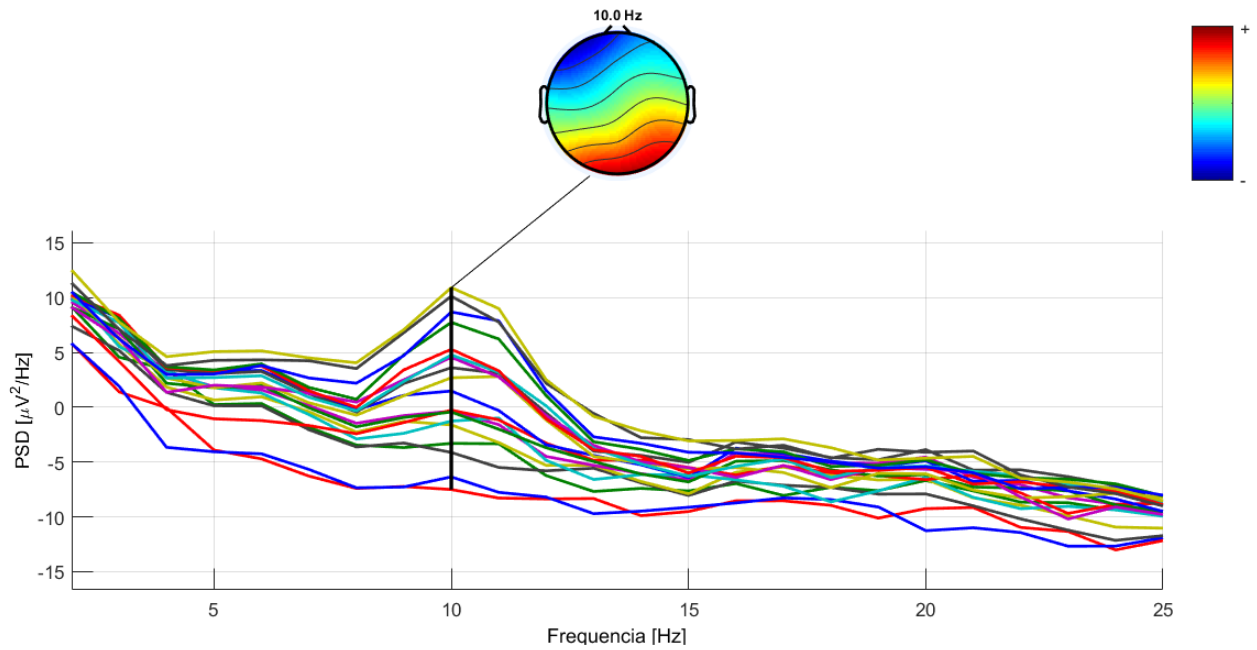


Figura 4: Diagrama de bloques de GoogLeNet

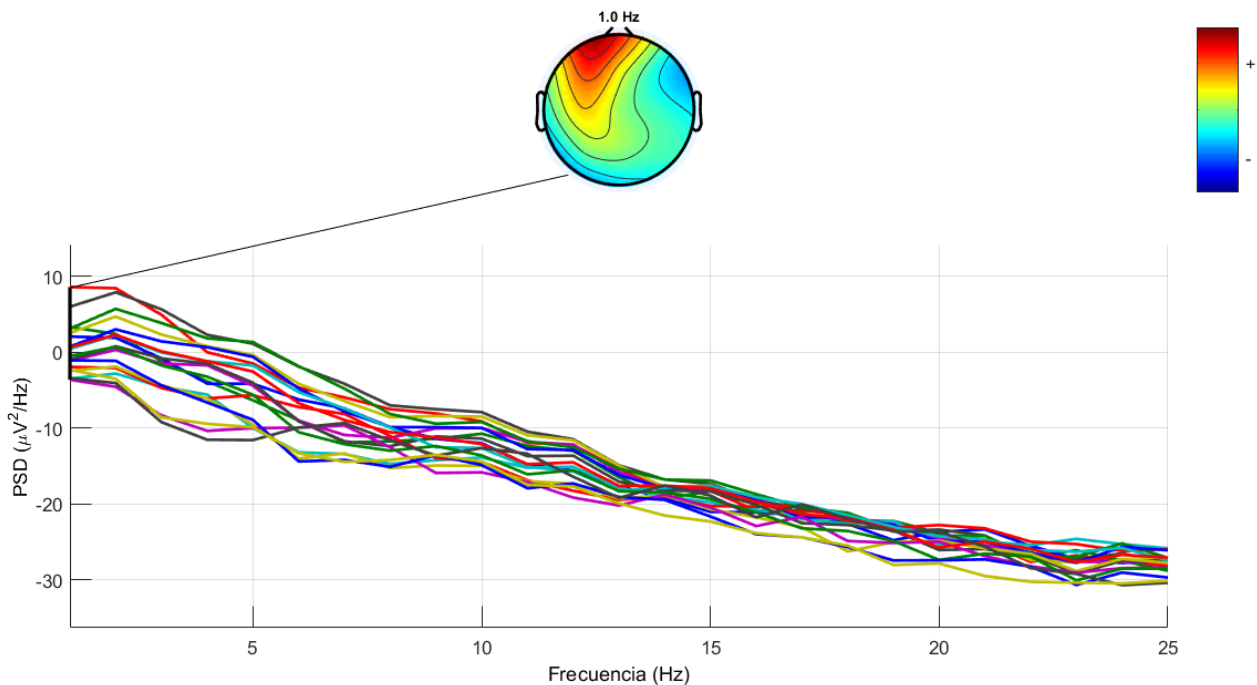
el software EEGLAB [7]. El programa se configuró para filtrado pasa-bandas, entre 0,1 y 100 Hz (banda pasante de EEG [2]), con filtrado adicional de rechazo de ruido de línea de modo común de 60 Hz. Para la selección del canal, se adoptó como criterio, identificar el canal con máxima potencia; estrategia, que funcionó satisfactoriamente en investigaciones previas [11]. EEGLAB permitió obtener la respuesta espectral del EEG y la identificación visual del canal de máxima potencia, como puede observarse en las Figuras 5a y 5b.

3. Generación del conjunto de entrenamiento; la señal del canal seleccionado se segmenta en trozos solapados, que conserven los

atributos de no linealidad y no “gaussianidad” [18, 17]. Este proceso de segmentación con solapamiento, tiene el objetivo de aumentar el tamaño del conjunto de datos acopiado; aspecto clave, para evitar el riesgo de sobreajuste en el entrenamiento [9, 15]. A cada segmento, se le calculó el correspondiente cumulante de tercer orden, mediante HOSA, para generar las señales bidimensionales. Estas señales bidimensionales se guardaron como imágenes, recortadas a dimensiones compatibles con la CNN a usar, de modo de poder integrar el conjunto de datos de entrenamiento, validación y prueba. En el caso actual, la dimensión de cada imagen es 229×229 , que es el formato de imagen



(a) Paciente de control



(b) Paciente epiléptico reducidas

Figura 5: Respuesta espectral del EEG

compatible con la CNN Inception.

4.2. Diseño del clasificador

En esta investigación, se usó la arquitectura de CNN pre-entrenada: Inception. El modelo utilizado en la arquitectura es del formato “No Top”,

lo que significa, que del modelo preentrenado, solo se utilizó la estructura correspondiente a la base convolucional, mientras que, la etapa clasificadora se sustituyó por un perceptrón multicapa especializado en la base de datos de EEG. La técnica de adaptar una red pre-entrenada a

una base de datos, distinta a la usada para entrenarla originalmente, se conoce como aprendizaje por transferencia (“transfer learning” [15]). La CNN Inception mostró buen desempeño, evaluada con el conjunto de datos generado para los experimentos, y en consecuencia, fue la CNN seleccionada para la tarea de clasificador, del sistema.

La Figura 6, muestra la estructura de las cuatro últimas capas del clasificador implementado con Inception. La capa identificada como “avg_pool” es la capa de salida de la base convolucional, tiene un ancho de 2048 neuronas, y se convierte en la dimensión del tensor, que se conecta a la capa de entrada, de la etapa clasificadora final. Las capas a continuación, conforman el perceptrón multicapa (etapa clasificadora), con una capa “dropout” [15], para aumentar la capacidad de generalización del modelo; inmediatamente, una capa intermedia totalmente interconectada, también llamada densa, de 100 neuronas y otra capa densa final de dos neuronas (las dos clases: epiléptico y control). El total de parámetros entrenables es de 205.192, el total de parámetros del detector (base convolucional mas etapa clasificadora) es de 22.007.886, lo cual es una cifra modesta, comparada con otras arquitecturas pre-entrenadas [9].

5. Análisis y discusión de resultados

Siguiendo los pasos definidos en la metodología, la fase experimental produjo los siguientes resultados:

5.1. Procesamiento del EEG

1. Base de datos

La base de datos usada corresponde a los EEG de 10 pacientes correctamente diagnosticados epilépticos y 10 pacientes de control, provenientes del banco de datos del Laboratorio de Electrofisiología del Hospital San Juan de Dios de Mérida [19]. Para cada EEG se registraron 19 canales electroencefalográficos usando el método referencial (diferencias de potencial entre el sitio activo y un punto de referencia dado por ambos lóbulos auriculares cortocircuitados), la frecuencia de muestreo fue 256 Hz. Los registros fueron

realizados tanto para controles como pacientes en estado de vigilia con ojos cerrados durante aproximadamente cinco minutos.

2. Procesamiento y selección del canal

Para cada uno de los EEG de la base de datos, se aplicó filtrado digital pasabandas 0,1-100 Hz) y de línea (60 Hz); así como también, se eliminaron segmentos, que, a simple inspección visual, lucieron afectados por artefactos técnicos, tales como desplazamiento accidental de electrodos [2, 4]. En cada EEG, ya filtrado, se identificó el canal que exhibió mayor potencia, usando la respuesta espectral, mediante EEGLAB (Figuras 5a y 5b). En el caso de EEG de pacientes epilépticos, se detectó gran variabilidad en la identidad del canal de máxima potencia, mientras que, en el caso de pacientes de control, la identificación del canal, se concentró en los canales de la región occipital (15 al 19), lo que es consistente con la condición de relajación y ritmo cerebral alfa [2].

3. Generación de imágenes

Con 20 señales disponibles (10 canales epilépticos y 10 de control); para cada señal se cortaron segmentos de 128 muestras, solapados 50%, con inicio de segmento aleatorio; lo que permitió obtener 2000 segmentos diferentes por clase. A cada uno de los segmentos, se les realizó el siguiente procedimiento:

- Verificación de condición de no gaussianidad y no linealidad.
- Cálculo de cumulante de tercer orden $C_{3x}(k, l)$.
- Recorte a dimensiones 299×299 pixeles, para compatibilidad con Inception.

La base datos fue de 4000 imágenes de 299×299 pixeles (2000 imágenes por clase).

5.2. Entrenamiento del clasificador

La CNN Inception se entrenó con la base de datos disponible de 2000 imágenes de $[299 \times 299]$ por clase. La base de datos organizada como 80% de entrenamiento (1600 por clase) y 20% para validación (200 por clase) y prueba (200 por clase).

```

=====
avg_pool (GlobalAveragePooling2 (None, 2048))      0      mixed10[0][0]
-----
dropout_3 (Dropout)                                (None, 2048)      0      avg_pool[0][0]
-----
dense_4 (Dense)                                    (None, 100)      204900      dropout_3[0][0]
-----
dense_5 (Dense)                                    (None, 2)      202      dense_4[0][0]
=====
Total params: 22,007,886
Trainable params: 205,102
Non-trainable params: 21,802,784
=====

```

Figura 6: Capas finales del clasificador usando Inception

El detector exhibió una exactitud de 96,5 % y 94,5 % para entrenamiento y prueba respectivamente, el número de épocas empleado fue de 20, con una duración de 90 segundos promedio por época, lo que representó un lapso total de tiempo entrenamiento, de aproximadamente 35 minutos. Las Tablas 1 y 2, resumen los resultados experimentales.

Tabla 1: Base de datos de clasificador

Clase	Imágenes por clase	Imágenes entrenamiento	Imágenes validación /prueba
Epiléptico	2000	1600	400
Control	2000	1600	400
Total, clases: 2	Total, clases: 2	entrenamiento: 3200	validación / prueba: 800

Tabla 2: Parámetros de entrenamiento de clasificador a CNN Inception

Modelo	Épocas	Exactitud de entrenamiento %	Exactitud de prueba %
EEG_Incep_01.h	20	96.5	94.5

5.3. Análisis de resultados

La aplicación del algoritmo de Hinich, permitió verificar la naturaleza no lineal y no gaussiana del EEG, tanto para pacientes epilépticos como de control.

La medición de la máxima potencia en la respuesta espectral, de los canales que conforman

el registro del EEG, indicaron que, en los pacientes de control, la mayor actividad cerebral se concentra en la región occipital, lo cual es consistente con el estado de relajación, en el que se encontraban al momento de la captación del EEG, y asociado al estado alfa de los ritmos cerebrales. Por otro lado, la alta variabilidad en la identificación del canal de máxima potencia, en el caso de EEG epiléptico, puede estar correlacionado con el foco epiléptico [2] del paciente, lo que sugiere un tema de investigación a explorar.

Las representaciones bidimensionales de los cumulantes de tercer orden, permitieron discriminar, en forma efectiva, los rasgos subyacentes de condición epiléptica, respecto a la de control, y en consecuencia lograr una alta precisión, en la detección de la neuropatía. La selección de la red Inception fue acertada, ya que tuvo un excelente desempeño en la tarea de clasificación.

6. Conclusiones

En este artículo se presentó el diseño y evaluación de un sistema detector de neuropatías basado en aprendizaje profundo y técnicas de análisis de orden superior.

La neuropatía, usada como clase a predecir, fue epilepsia, por ser una de las neuropatías más comunes y con alta degradación, en la calidad de vida, del enfermo.

Se verificó el comportamiento altamente no lineal y no gaussiano de las series de tiempo derivadas de registros electroencefalográficos y en consecuencia la acertada selección del uso

de métricas estadísticas de orden superior, para entrenar el sistema detector.

El desempeño del detector fue muy satisfactorio, alcanzando una exactitud de 94,5 % en el conjunto de prueba. La selección de la red Inception, en la implementación del clasificador, así, como el uso de la técnica de aprendizaje por transferencia, en su entrenamiento, mostró una alta efectividad.

Los resultados de este estudio sugieren la conveniencia de desarrollar este tipo de herramientas computacionales, de alta precisión, para el diagnóstico de otras neuropatologías.

7. Referencias

- [1] U. K. Patel, A. Anwar, S. Saleem, P. Malik, B. Rasul, K. Patel, R. Yao, A. Seshadri, M. Yousufuddin, and K. Arumathurai, "Artificial intelligence as an emerging technology in the current care of neurological disorders," *Journal of Neurology*, Aug. 2019.
- [2] L. Sörnmo and P. Laguna, *Bioelectrical Signals Processing in Cardiac and Neurological Applications*. Amsterdam: Elsevier Academic Press, 2005.
- [3] U. Raghavendra, U. Acharya, and H. Adeli, "Artificial Intelligence Techniques for Automated Diagnosis of Neurological Disorders," *Europea Neurology*, vol. 82, pp. 41–64, 2019.
- [4] C. Seijas, A. Caralli, and S. Villazana, "Estimation of Action Potential of the Cellular Membrane using Support Vectors Machines," in *Proceedings of the 28th IEEE EMBS Annual International Conference*, 2006, pp. 4200–4204.
- [5] M. Teplan, "Fundamentals of EEG Measurement," *Measurement science review*, vol. 2, pp. 1–11, 2002.
- [6] S. Smith, "EEG in the Diagnosis, Classification, and Management of Patients with Epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. Supplement 2: ii2-ii7, 2005.
- [7] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 132, pp. 9–12, 2004.
- [8] A. Swami, J. Mendel, and C. Nikias, *Higher-Order Spectral Analysis Toolbox: User's Guide*. The MathWorks Inc., 2007.
- [9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, v. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–9.
- [10] N. Mahmoodian, J. Haddadnia, A. Illanes, A. Boese, and M. Friebe, "Seizure prediction with cross-higher-order spectral analysis of EEG signals," *Signal, Image and Video Processing*, vol. 14, pp. 821–828, 2020.
- [11] C. Seijas, A. Caralli, and S. Villazana, "Neuropathology Classifier Based on Higher Order Spectra," *Journal of Computer and Communications*, vol. 1, no. 4, pp. 36–40, 2013.
- [12] A. Craik, y. He, and J. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal Neural Engineering*, vol. 16, no. 3, pp. 1–28, 2019.
- [13] M. Golmohammadi, A. Hossein, H. Nejad, S. Lopez de Diego, I. Obeid, and J. Picon, "Automatic Analysis of EEGs Using Big Data and Hybrid Deep Learning Architectures," *Frontiers in human neuroscience*, 2019.
- [14] L. Gemein, R. Schirrmeyer, P. Chrabaszcz, D. Wilson, J. Boedecker, A. Schulze-Bonhage, F. Hutter, and T. Ball, "Machine-learning-based diagnostics of EEG pathology," *Neuroimage*, vol. 220, p. 117021, 2020.
- [15] F. Chollet, *Deep Learning with Python*. New York, USA.: Manning Publications Co., 2018.
- [16] C. Wen-Yean, M. Murugappan, M. Iqbal, W. Khai-runizam, B. Siao-Zheng, A. Joseph, and Z. Ibrahim, "A Study of Non-Gaussian Properties in Emotional EEG in Stroke Using Higher-Order Statistics," in *Advances in Electrical and Computer Technologies. Select Proceedings of ICAECT 2019*, T. Sengodan, M. Murugappan, and S. Misra, Eds., vol. 672. Springer, 2020, pp. 635–645.
- [17] P. Cauchy, S. Jena, S. Nadar, and N. Pradhan, "Higher-Order Spectrum in Understanding Nonlinearity in EEG Rhythms," *Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine*, vol. 2012, no. ID 206857, 2012.
- [18] M. Hinich, "Testing for Gaussianity and Linearity of a Stationary Time Series," *Journal of Time Series Analysis*, vol. 3, no. 3, 1982.
- [19] P. Guillén, "Procesamiento Digital de EEG," Tesis doctoral, Universidad de los Andes, Mérida, Venezuela., 2005.

Environment Descriptor for Visually Impaired People Implemented on Raspberry Pi Based on Convolutional and Recurrent Artificial Neural Networks

Rafael Chourio ^{*,a} , Wilmer Sanz ^b 

^aMaestría en Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela.

^bLaboratorio de Robótica y Visión Industrial, Escuela de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela

La selección de este artículo fue realizada en el marco de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica "Prof. César Rodolfo Ruiz" Octubre 2020, siendo la evaluación, arbitraje, aceptación y edición a cargo de Revista Ingeniería UC.

Abstract.- Vision problems and blindness are disorders of the human body that according to figures from the World Health Organization affect 217 million people in low-income countries. The quality of life of at least 75 million of them can be improved with the development of systems that allow guiding them safely in their daily tasks; this is where it is important to look for technological alternatives oriented to solve this problem and it is precisely where the idea of this research work was born. The idea presented here is based on the development of an image description system trained with deep learning algorithms based on convolutional and recurrent neural networks, implemented in a single-board computer. This implementation uses a low-cost camera for taken images of the environment and obtains a description of it that can be converted to an audible voice signal through a hearing aid system so that visually impaired people can improve their standard of living by obtaining real-time information from the environment surrounding.

Keywords: Artificial Neural Network; Natural Language Processing; Raspberry Pi; Deep Learning; Features Extraction.

Descriptor del Entorno para Personas con Discapacidad Visual Implementado en Raspberry Pi Basado en Redes Neuronales Artificiales Convolucionales y Recurrentes

Resumen.- Los problemas de visión y la ceguera son desórdenes del cuerpo humano que de acuerdo con cifras de la Organización Mundial de la Salud afectan a 217 millones de personas en países de bajos recursos. La calidad de vida de al menos 75 millones de estas personas puede ser mejorada con el desarrollo de sistemas que permitan guiarlas de manera segura en sus tareas de desplazamiento diario; allí es donde resulta importante buscar alternativas tecnológicas orientadas a solventar esta problemática y es donde precisamente nace la idea de este trabajo de investigación. La idea aquí presentada se basa en el desarrollo de un sistema de descripción de imágenes entrenado con tecnología de aprendizaje profundo, con redes neuronales convolucionales y recurrentes implementado en un computador de placa única, de modo tal que con una cámara de bajo costo se puedan tomar imágenes del entorno y obtener una descripción del mismo que pueda ser convertida a una señal de voz audible a través de un sistema de audífonos, para que personas con deficiencia visual puedan mejorar su nivel de vida a través de la obtención de información en tiempo real del entorno que los rodea.

Palabras clave: Redes Neuronales; Procesamiento Natural de Lenguaje; Raspberry Pi; Aprendizaje Profundo; Extracción de Características.

Recibido: 22 de octubre, 2020.

Aceptado: 29 de noviembre, 2020.

1. Introducción

Una de las discapacidades que mayores limitaciones puede producir al normal desenvolvimiento de un ser humano se refiere a las correspondientes

* Autor para correspondencia:

Correo-e: rafaelchourio@gmail.com (R. Chourio)

al sentido de la vista, dada nuestra alta dependencia de este sentido para la interacción con el medio ambiente (en relación por ejemplo con otros sentidos o mecanismos de interacción como el oído, el habla, el gusto o el olfato). Los problemas de visión y la ceguera son desórdenes del cuerpo humano que generan la pérdida o falta de observación visual, y que en general se relacionan con problemas fisiológicos o neurológicos.

A través de la historia de la humanidad se han ideado un sinnúmero de soluciones basadas en la técnica y la práctica a fin de dar una mayor calidad de vida a las personas con diversos grados de discapacidad visual; todos estos esfuerzos se han visto reflejados en la creación y desarrollo de instrumentos como espejuelos, lentes de contacto o anteojos, pasando por el desarrollo de métodos de lectura para invidentes como el sistema Braille hasta los actuales esfuerzos basados en el uso de sistemas de visión artificial que ayuden a describir el entorno que rodea a las personas con deficiencias visuales.

De acuerdo con la Clasificación Internacional de Enfermedades (CIE-10) actualizado al año 2008 [1], el desempeño de la función visual se puede clasificar en cuatro categorías principales: Visión Normal, Discapacidad Visual Moderada, Discapacidad Visual Grave y Ceguera. Según datos de la Organización Mundial de la salud, para Octubre del año 2017 la cifra estimada a nivel mundial de personas con discapacidad visual era de 253 millones de individuos, de los cuales 36 millones estaban diagnosticados con ceguera permanente y 217 millones con discapacidad visual que iba de moderada a grave.

De estas 217 millones de personas con baja visión a nivel mundial el 90 % de ellas (195 millones aproximadamente) viven en países de ingresos bajos. De este universo se estima que unos 120 millones de ellos padecen dicha afección por errores de refracción no corregidos (problema que en gran medida puede ser resuelto con el uso de lentes correctivos u operaciones quirúrgicas). Por lo tanto la población restante (al menos 75 millones de personas) va a requerir algún tipo de apoyo visual para dar mayor normalidad a sus vidas, ya que para sus afecciones hoy día no existe

solución tecnológica ni médica posible. A todas estas cifras también debemos sumar que el 28 % de las personas con baja visión se encuentran en edad laboral productiva.

La calidad de vida de las personas con impedimento visual puede ser elevada enormemente en cuanto se desarrollen sistemas que permitan guiarlas de manera segura en sus tareas de desplazamiento pedestre diario, siendo este uno de los grandes problemas que confrontan tanto las personas que sufren este tipo de desórdenes como sus familiares y allegados [2]. Para poder ayudar a una persona con discapacidad visual a describir el entorno que le rodea, es necesario que la información generada pueda ser de alguna manera transmitida a dicha persona de forma tal que pueda ser interpretada, para lo cual la forma más simple es a través del uso de alguno de los otros sentidos; para esta labor el sentido que se muestra con mayor capacidad de cumplir con esta función es el oído.

La principal función de la ciencia aplicada es la resolución de problemas que mejoren las condiciones de vida de los seres humanos, bien sea para potenciar las limitaciones inherentes a nuestra forma biológica, para mejorar la dinámica y/o funcionamiento social o para superar las limitaciones de nuestros sentidos producidas por discapacidades que puedan mermar nuestras condiciones de vida. Es por ello que se hace importante buscar alternativas tecnológicas que permitan mejorar la calidad de vida de las personas con desórdenes de baja visión, de modo tal de poder brindarles las herramientas necesarias para que puedan llevar una vida más independiente y productiva. Dada la alta incidencia de este tipo de desórdenes en países de bajos ingresos (lo cual implica elevados niveles de pobreza) se hace importante tomar en cuenta que dichas alternativas, para poder ser de ayuda efectiva a estas personas, deben ser de bajo costo y ser capaces de funcionar sin conexión a internet dedicada.

Es en base a todo ello que el presente trabajo propone la implementación de un sistema de bajo costo orientado a la ayuda de personas con deficiencia visual en países de bajos ingresos económicos. La idea presentada se basa en el desarrollo de un sistema de descripción de

imágenes entrenado con tecnología de aprendizaje profundo e implementado en un computador de placa reducida, de modo tal que con una cámara de bajo costo se puedan tomar imágenes del entorno y obtener una descripción del mismo que sea convertida a una señal de voz audible a través de un sistema de audífonos, para que así las personas con deficiencia visual puedan mejorar su nivel de vida a través de la obtención de información en tiempo real del entorno que los rodea.

1.1. Aprendizaje Automático: Redes Neuronales Artificiales

El Aprendizaje Automático se refiere a la rama de las ciencias de la computación que estudia las técnicas que le permiten a los computadores aprender en lugar de seguir instrucciones programadas; más específicamente trata acerca del desarrollo de programas que le permitan a un computador generalizar comportamientos tomando como base un conjunto de ejemplos dados a partir de los cuales debe aprender [3]. Si se parte de la premisa de que el método de investigación lógico inductivo se basa en la elaboración de conclusiones de corte general partiendo de observaciones particulares, no es difícil concluir que los algoritmos de máquinas de aprendizaje terminan siendo un proceso de inducción del conocimiento.

De acuerdo a la forma que tienen los datos a partir de los cuales se pretende aprender con una máquina de aprendizaje, se pueden reconocer varios tipos de algoritmos. Entre los más conocidos se encuentran el Aprendizaje Supervisado, el Aprendizaje Semisupervisado y el aprendizaje No Supervisado; con respecto al tipo específico de técnicas utilizadas para implementarlos para labores de clasificación o regresión sobre un conjunto de datos a partir del cual se quiere realizar un proceso de aprendizaje, estos enfoques se encuentran agrupados en 7 principales tipos a saber: Árboles de Decisiones, Reglas de Asociación, Algoritmos Genéticos, Redes Neuronales Artificiales, Máquinas de Vectores de Soporte, Algoritmos de Agrupamiento y Redes Bayesianas.

El desarrollo de las redes neuronales artificiales se inspiró en el funcionamiento del cerebro humano para procesar información y aprender a partir de

su entorno. El descubrimiento de este mecanismo de funcionamiento del cerebro fue realizado en 1888 por Ramón y Cajal [4], quien descubrió la estructura celular del sistema nervioso. Inspirado en este mecanismo, en 1986 Rumelhart, Hinton y McClelland [5] definieron una neurona artificial como un dispositivo capaz de generar una única salida a partir de un set de entradas. En la Figura 1 se puede observar una red neuronal artificial con una única salida.

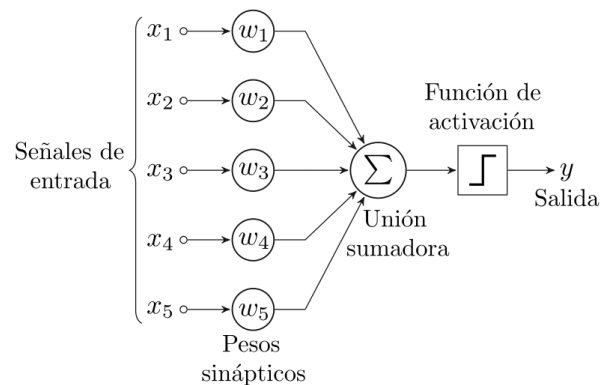


Figura 1: Red Neuronal Artificial

Las señales de entrada son denotadas como x , cada una de ellas es modificada por un factor denominado peso sináptico, el cual es diferente para cada entrada. La suma de todas las entradas ponderadas por su respectivo peso sináptico es aplicada a una función de activación, que es la función encargada de romper la linealidad del modelo para generar una salida y . Esta red neuronal puede ser utilizada para clasificación binaria o regresión. En la Figura 2 se puede observar una red neuronal artificial con múltiples salidas.

A fin de que el algoritmo de aprendizaje pueda ser efectivo en la tarea de clasificación o regresión la función de salida obtenida a partir de los datos de entrenamiento (en el caso del aprendizaje supervisado) se debe comparar con la etiqueta que cada ejemplo posee. A esta función de comparación se le llama función de costo, que depende de las variables w y b (pesos y bias de la red entrenada). El ajuste óptimo de los pesos se basa en la minimización de la función de costo, lo cual se traduce en un alto nivel de predicción. Para ello una de las técnicas más usadas es la del

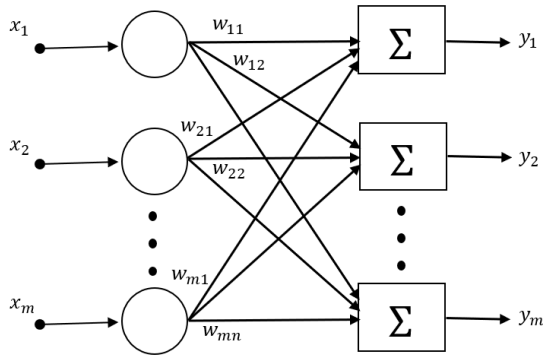


Figura 2: Red Neuronal Artificial Multi-salida

gradiente descendiente, la cual consiste en calcular la dirección del gradiente de la función de coste en el punto establecido por cada evaluación del conjunto de entrenamiento, para así ajustar los pesos y bias en la dirección donde dicho gradiente descienda.

1.2. Aprendizaje Profundo (Deep Learning)

Las redes neuronales artificiales pueden ser utilizadas en arquitecturas de múltiples capas, donde cada capa sea capaz de procesar y extraer determinadas características de la información a analizar para lograr un mayor nivel de abstracción, del mismo modo que lo hacen las capas de neuronas en el cerebro humano. A pesar de haberse diseñado modelos de redes neuronales desde los años 80 del siglo XX, la capacidad computacional de los procesadores existentes en esa época era insuficiente para aplicaciones con muchas capas de profundidad. No es sino hasta bien entrada la segunda década del presente siglo cuando la capacidad de cálculo de los modernos procesadores y el desarrollo de potentes tarjetas de video (GPU) han potenciado el desarrollo de aplicaciones de redes neuronales de múltiples capas. Esto es lo que se conoce como aprendizaje profundo (referido como Deep Learning en idioma inglés).

El uso de múltiples capas para la extracción de características complejas en los datos de entrenamiento ha permitido el análisis automático de imágenes, el procesamiento natural de lenguaje y la mejora en la predicción de sistemas complejos (clima, análisis de variables en bolsa

de valores entre otras). Junto con la evolución en capacidad de cálculo se ha requerido el desarrollo de investigaciones matemáticas, cálculo numérico y optimización de algoritmos para hacer frente a problemas que se presentan en dichas redes neuronales, entre los cuales se cuentan el desvanecimiento del gradiente, la explosión del gradiente o la sensibilidad a la traslación o rotación en el reconocimiento de imágenes.

1.2.1. Redes Neuronales Convolucionales

Las redes neuronales convolucionales están basadas en los trabajos realizados por Hubel y Wisel en 1959 [6], que dieron luz sobre la comprensión del mecanismo de funcionamiento de la corteza visual del cerebro. Su origen está en la arquitectura de red neuronal presentada por Kunihiko Fukushima en 1980 [7], a la cual le dio el nombre de Neocognitron; no obstante no es sino hasta el año 2011 [8] en que se logra refinar esta técnica con uso de backpropagation e implementarla en un GPU para hacer clasificación de imágenes.

En las redes neuronales convolucionales se programan neuronas con campos receptivos de creciente complejidad, aplicando técnicas de filtrado al campo visual de cada neurona a fin de generar, entre otras cosas, insensibilidad con respecto a la posición de un objeto dentro del campo visual completo de la imagen, proveer un mecanismo de reducción de complejidad computacional y extracción de características de las imagen. La aplicación de la técnica de convolución a un conjunto de píxeles de una imagen o a la salida de una capa de neuronas se realiza a través de la ecuación (1).

$$[Y_j = [b_j + \sum_i K_{ij} Y_i]] \tag{1}$$

En donde Y_j es la salida de la capa de convolución agregada a la salida de una capa de neuronas o los píxeles de una imagen Y_i , K_{ij} es la matriz o núcleo de convolución correspondiente y que depende de la característica que se desea extraer y b_j es un factor de desplazamiento o bias. Esta operación que es similar en cierto sentido a la operación de convolución matemática en el tiempo que es ampliamente usada en los

sistemas de comunicaciones, tiene el efecto de lograr un filtrado de características de la capa de neuronas anterior a través de los parámetros del núcleo aplicado. Aplicando capas de convolución a imágenes, combinadas con funciones de activación y otras técnicas de regularización y reducción de dimensionalidad, es posible el análisis de imágenes complejas para extraer de ellas características correspondientes a objetos presentes en ellas e incluso la relación entre dichos objetos (encima, debajo, delante, detrás, etc.).

1.2.2. Redes Neuronales Recurrentes

El uso de Redes Neuronales secuenciales o convolucionales en aplicaciones para datos de tipo secuencial tales como vídeo, música o análisis de lenguaje son limitados debido a que no poseen capacidad de tomar en cuenta las decisiones anteriores como base para realizar una predicción; por ejemplo, para poder hacer la traducción de un texto de un lenguaje a otro, es importante hacerlo en el contexto de las palabras ya traducidas a fin de que la respuesta obtenida tenga sentido. Por ello se crearon las redes neuronales recurrentes, las cuales poseen lazos de transmisión de información que permiten que la información de decisiones anteriores persista.

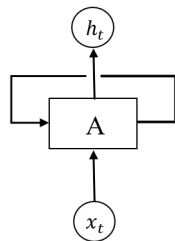


Figura 3: Red Neuronal Recurrente Enrollada

En la Figura 3 se puede observar un diagrama simplificado donde además de los estados de entrada y salida existe un estado interno oculto que se transmite de forma cíclica a través del lazo dibujado, el cual representa un estado de memoria que permite recordar las decisiones anteriores y es utilizado en la celda A para decidir la salida h_t . Este diagrama se puede desenrollar, de manera de ver con mayor claridad que estos estados internos

básicamente transmiten información del pasado a través del tiempo (Figura 4).

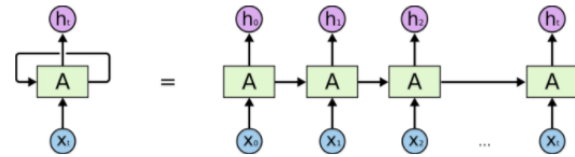


Figura 4: Red Neuronal Recurrente Desenrollada

Las primeras redes neuronales recurrentes a pesar de poder transmitir información de estados pasados para poder tomar mejores decisiones presentes, tenían limitaciones en cuanto a la cantidad de información que pueden efectivamente almacenar. Si los estados temporales transmitidos incluyen memorización a largo plazo, los algoritmos de optimización utilizados tienden a ser poco efectivos para esta arquitectura de aprendizaje [9]. Para contrarrestar esto se desarrollaron las Redes Neuronal Recurrente LSTM, que son una arquitectura de aprendizaje desarrollada para aumentar la efectividad del entrenamiento con dependencias temporales largas [10].

LSTM es el acrónimo de Long Short Term Memory. Con este esquema, se controla la información de los estados internos permitiendo olvidar o mantener información conforme sea requerido (de la misma forma en que trabaja la mente humana); este diseño está específicamente desarrollado para evitar el problema de larga dependencia temporal, razón por la cual actualmente es ampliamente utilizado en la resolución de una gran variedad de problemas de índole secuencial.

1.3. Antecedentes

Tanto el reconocimiento automático de imágenes como el procesamiento natural de lenguaje han comenzado a converger en un nuevo nicho de investigación relacionado con la extracción de características de imágenes para su correcta descripción. Uno de los trabajos que se ha convertido en una referencia en este campo es el paper titulado “Show and Tell: A Neural Image Caption Generator” [11] desarrollado por integrantes del equipo investigador de Google, Inc. Este trabajo fue una de las primeras aproximaciones

basada en la combinación de un modelo de red neuronal convolucional que alimenta a una red neuronal recurrente para la obtención de una descripción de la imagen procesada.

En “Long-term Recurrent Convolutional Networks for Visual Recognition and Description” [12] se utiliza una variante denominada “Long-term Recurrent Convolutional Network” (LRCN) a fin de procesar secuencias de imágenes de video con la finalidad de generar descripciones coherentes de las acciones que ocurren en el mismo. Para ello hicieron pruebas con varios modelos pre-entrenados, con ajustes de varios hiperparámetros, con diversos tipos y números de capas en redes neuronales recurrentes a fin de lograr descripciones con la mayor robustez posible.

Dado lo novedoso de este campo de aplicaciones, aún se está lejos de alcanzar niveles de descripción comparables al logrado por el ser humano. Es por ello que han surgido propuestas orientadas a la mejora en la precisión y naturalidad de las descripciones obtenidas, bien sea para imágenes o secuencias de imágenes. “Boosting Image Captioning with Attributes” [13] y “Exploring Models and Data for Remote Sensing Image Caption Generation” [14] son dos de los estudios más recientes enfocados a la optimización de modelos y parámetros orientados a la generación de descripciones basadas en imágenes, lo cual es un área de estudio en constante desarrollo para lograr aplicaciones cada vez más precisas en tiempo real.

Un área del aprendizaje supervisado que actualmente se encuentra lo suficientemente madura como para el desarrollo de aplicaciones robustas es el reconocimiento automático de texto; es en este sentido que se han desarrollado multitud de propuestas orientadas a la aplicabilidad de esta tecnología para la asistencia a personas con discapacidad visual. Tanto en la investigación “A Smart Reader for Visually Impaired People Using Raspberry Pi” [15] como en “Design of An Electronic Narrator on Assistant Robot for Blind People” [16] se han utilizado computadores de placa reducida Raspberry Pi como base para la construcción de sistemas narrativos que puedan, a través de una cámara instalada en dicho dispositivo, capturar imágenes de un texto y procesar el

mismo para convertir dicha información en audio que pueda ser escuchado por una persona con impedimento visual.

Este mismo enfoque puede ser utilizado para describir el ambiente circundante a personas que tengan limitaciones para poder observarlo. Precisamente esta fue la idea presentada en el estudio “Automated Neural Image Caption Generator for Visually Impaired People” [17], que planteó la posibilidad de utilizar las descripciones provenientes de un sistema entrenado con la utilización de una red neuronal convolucional como entrada a una red neuronal recurrente de forma tal de poder explicar de algún modo el ambiente circundante y servir como referencia a personas que posean diversos tipos de discapacidad visual.

2. Metodología

La presente investigación se circunscribe a la aplicación de un sistema de aprendizaje supervisado para la identificación de los principales elementos presentes en una imagen capturada en tiempo real y la interrelación existente entre ellos, a fin de crear una frase coherente que permita describir la escena presentada de una forma concisa. Esta aplicación se implementó en un computador de placa reducida, en aras de sentar este precedente como base a la futura implementación de este tipo de desarrollos en sistemas de bajo costo accesible para países y personas de bajos ingresos. Igualmente es importante el hecho de que el prototipo propuesto puede funcionar sin necesidad de conexión a internet.

El computador escogido para la implementación de este sistema ha sido el Raspberry Pi 3, debido a su versatilidad, bajo costo y software libre para desarrollo. La frase generada para describir la escena capturada se convierte a una señal de voz, a fin de que la persona con impedimento visual tenga la oportunidad de escuchar la descripción de la imagen a través de un sistema de audífonos instalados en el computador de placa reducida. Igualmente se acopla una cámara capaz de hacer captura de la imagen deseada y un sistema de

accionamiento que permite al usuario decidir el momento en que quiere tener una descripción de su entorno.

El software de programación para la implementación del presente trabajo de investigación fue Python, en su versión 2.7, debido a su compatibilidad con las librerías TensorFlow, Keras y OpenCV en el sistema operativo Raspbian utilizado por el Raspberry Pi 3 (las tres librerías mencionadas fueron necesarias para las tareas de entrenamiento del sistema de aprendizaje supervisado, para la construcción del modelo de predicción y para la captura de imágenes desde la cámara instalada en el dispositivo); Raspbian (actualmente Raspberry Pi OS) es una distribución del sistema operativo GNU/Linux basado en Debian, y es de manera oficial el sistema operativo primario de la familia de placas Raspberry Pi.

Es importante recalcar que los objetos identificables en las imágenes para la generación de las descripciones están basados en el set de entrenamiento utilizado para entrenar el sistema de aprendizaje profundo, por lo que objetos fuera de ese conjunto pueden crear también descripciones menos precisas.

El set de datos utilizado para el entrenamiento de la red neuronal recurrente dispone de descripciones en idioma inglés, razón por la cual este fue el idioma de las descripciones generadas tanto en el set de pruebas como en las imágenes capturadas por la cámara instalada en el Raspberry Pi.

2.1. Modelo Utilizado

El modelo elegido para ser implementado en el Raspberri Pi se basó en el NIC (Neural Image Caption Generator) publicado por el equipo de Google Inc. en el año 2015 [11]. Este modelo permitió generar descripciones nuevas en imágenes nunca vistas, tomando ventaja de la técnica de extracción de características de modelos previamente entrenados y que resultaron exitosos en tareas de clasificación de imágenes. Se escogió el mismo debido a su mecanismo de validación bien documentado (lo cual es un punto fuerte de este trabajo y que no pudo ser apreciado en muchos de los trabajos revisados), además de que la potencia de cálculo requerida sería lo suficiente

para ser manejado por una laptop de potencia media (por el uso de la mencionada extracción de características).

El modelo seleccionado consta de dos sistemas de redes neuronales a saber: una red neuronal convolucional que se encarga de procesar la imagen objetivo (etapa de Visión), y una red neuronal recurrente que genera una frase coherente que describa con mayor probabilidad a la imagen (etapa de Generación de Lenguaje).

La etapa de Visión se compone de una red neuronal profunda convolucional, cuya entrada se compone únicamente de la imagen objetivo, y que se debe encargar de extraer las características de dicha imagen. Debido a las exigencias computacionales de entrenar una red de este tipo desde cero, y dado que ya existen modelos pre entrenados y públicos para descarga libre en la red suficientemente probados, se decidió utilizar uno de estos modelos para la extracción de las características necesarias. La etapa de Generación de Lenguaje se implementó con una red neuronal LSTM, cuyo estado inicial se alimenta con las características que se extraen de la imagen a fin de escoger dentro del vocabulario definido la próxima palabra con mayor probabilidad de ocurrencia, repitiendo esta operación hasta completar una frase completa que describa a la imagen. En la Figura 5 se muestra el diagrama de bloques del modelo.

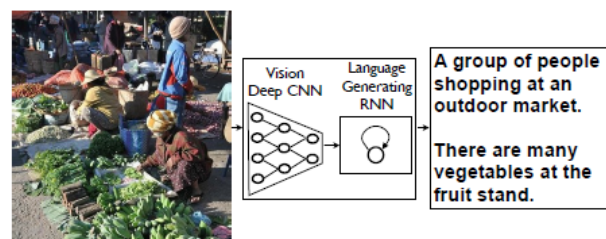


Figura 5: Modelo NIC [11]

Para la programación de la etapa de Visión se compararon varios modelos pre entrenados que están incluidos dentro de la librería Keras. Debido a la limitación en la capacidad disponible de cálculo para el desarrollo del presente trabajo, donde se contó sólo con una laptop con un procesador Intel i5 con dos núcleos, de 8 GB de memoria RAM y

sin GPU instalado, la escogencia de un modelo que minimizase la cantidad de operaciones requeridas para extraer las características de las imágenes en la fase de entrenamiento era primordial; lo mismo aplica a la limitada capacidad en procesador y memoria del Raspberry Pi para que el tiempo de respuesta por cada imagen no sea excesivo. Por ello se decidió utilizar el modelo InceptionV3 debido a su tamaño reducido (92 MB en total) y elevada precisión (78,8 % en el Top-1 Accuracy de Imagenet y 94,4 % en el Top-5 Accuracy del mismo set de imágenes).

Para elegir el set de datos óptimo para entrenamientos, se verificó en Vinyals [11] que el set de datos COCO mostró el mejor desempeño a nivel de evaluación de las frases generadas. Es por ello que se decidió utilizar este mismo set de datos, en específico el correspondiente al año 2014. El mismo está compuesto por imágenes de 91 clases, además de poseer un archivo de datos de referencia donde, por cada imagen, se tienen 5 descripciones que la representan. A partir de las imágenes se obtienen las características que se extraen por medio del modelo inceptionV3 pre entrenado sobre Imagenet, y a partir de las descripciones se obtiene el vocabulario que sirve como parámetro de evaluación para la etapa de generación de lenguaje; la forma de obtener las características de la imagen (y no la probabilidad de pertenencia a cada una de las clases sobre la que fue entrenado) es eliminando la última capa de decisión (SoftMax); con esto se obtiene el conjunto de características extraídas propias de cada imagen que van a servir como parámetros iniciales para la red neuronal recurrente LSTM. El set de datos usado estuvo compuesto por 82.783 imágenes de entrenamiento (con un tamaño aproximado de 13 GB) y por 40.504 imágenes de validación (con un tamaño aproximado de 6,5 GB). De este último conjunto de imágenes, se utilizaron 104 imágenes para pruebas, quedando el conjunto de validación compuesto por 40.400 imágenes en total. En la Figura 6 se observa el diagrama de bloques del modelo entrenado.

A fin de optimizar el entrenamiento de la red neuronal recurrente (la cual es la única que requiere ser entrenada, puesto que para la etapa de visión se

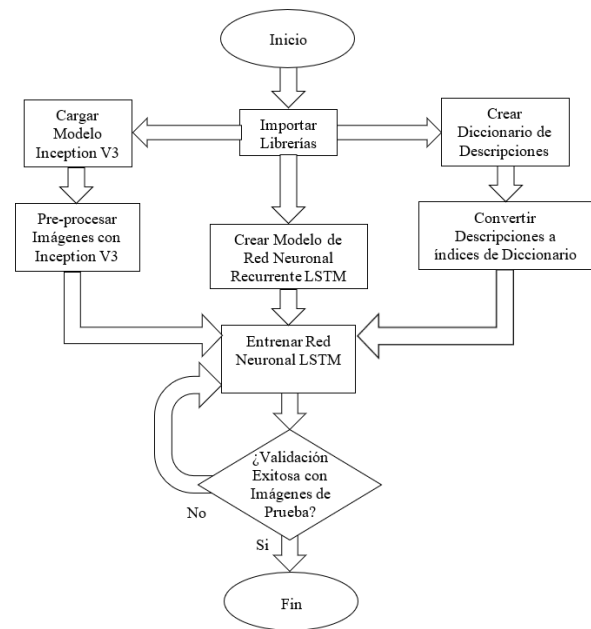


Figura 6: Diagrama de Bloques del Modelo Entrenado

está utilizando una red pre entrenada), se hizo una extracción previa de las características de cada una de las imágenes de entrenamiento y validación. Con el computador utilizado, se requirieron alrededor de 15 horas de cómputo para completar esta tarea.

Para implementar la etapa de Generación de Lenguaje es necesario el establecimiento de un vocabulario a partir del cual se van a escoger las palabras para la formación de las oraciones, este se construyó a partir de las descripciones que acompañan a las imágenes del conjunto de entrenamiento y constó de 8765 palabras, agregándose además 4 palabras adicionales: una para inicio de cada frase, una para fin de cada frase, una para identificar palabras desconocidas del vocabulario y una para especificar la ausencia de palabras. En total, el diccionario generado estuvo compuesto por 8769 palabras. El vocabulario generado se almacenó en formato *.pickle para su posterior utilización.

Las características obtenidas de cada imagen por medio del modelo InceptionV3 sirven como estado inicial a la celda LSTM. Para simplificación del modelo, se escogió una red LSTM de 756 unidades, la cual alimenta dos capas de redes neuronales, la última de las cuales tiene el tamaño

del vocabulario utilizado. A esta última se le aplica una máscara para el cálculo de probabilidad solo sobre las palabras que son conocidas, y luego una capa de cálculo SoftMax para selección de la palabra con mayor probabilidad de ser utilizada para construcción de la frase. La palabra escogida, así como las características de la imagen y el vocabulario es nuevamente alimentado a otra celda LSTM de la misma forma anterior para la escogencia de la siguiente palabra; el proceso se repite hasta que la palabra escogida es la palabra de fin de oración o hasta que se alcanza la longitud máxima de oración establecida, la condición que ocurra primero. Finalmente, la función de pérdidas usada para evaluar el aprendizaje fue la de entropía cruzada.

El entrenamiento se efectuó por batches de imágenes. Cada batch se definió de 64 imágenes, y cada época de entrenamiento se estipuló en 1.290. Para validación se tomaron 300 batches. Con estos parámetros se pudo observar un decrecimiento continuo de la función de pérdidas asociadas, y con 18 épocas se pudo obtener resultados de predicción precisos en el set de pruebas. El modelo completo se almacenó por cada época de entrenamiento para referencia en un archivo en el disco duro. En la Figura 7 se observa la curva de pérdidas por cada época de entrenamiento obtenida para ambos sets.

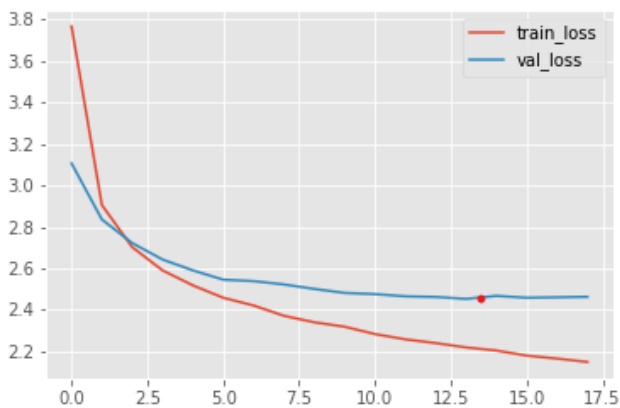


Figura 7: Curva de Precisión del Entrenamiento

2.2. Implementación en Raspberry Pi

En la Figura 8 se puede observar el diagrama de bloques del prototipo implementado.

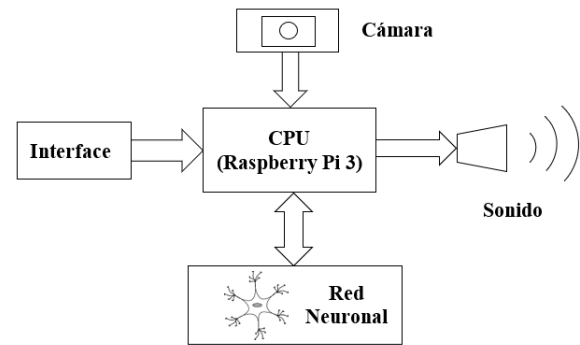


Figura 8: Diagrama de Bloques del Prototipo Implementado

Al computador de placa reducida que se observa en el centro de la Figura 8 se agregó una cámara para poder tomar la imagen del entorno. La interface está compuesta por dos pulsadores, uno para iniciar el proceso de toma de imagen-generación de descripción y otro para repetir la última descripción procesada, un sistema de sonido que está conformada por un set de auriculares, y un modelo de red neuronal previamente entrenado cuya misión es procesar la imagen capturada y generar la descripción de la misma. De la etapa de entrenamiento se obtuvieron dos archivos: el modelo generado con un peso de alrededor de 105 MB y el vocabulario cuyo tamaño fue de unos 167 kB.

La cámara utilizada fue una Kuman modelo SC15, con bus de comunicación para ser instalada en el puerto de cámara del computador de placa reducida. La instalación de la cámara requirió la habilitación de la misma en el entorno de configuración del Raspberry Pi, accesible a través del comando `raspi-config` en la consola de comando del mismo. Para el módulo de interface se agregó un pequeño protoboard, en el cual se incluyeron dos mini pulsadores; estas señales fueron conectados a las entradas del Raspberry Pi en las GPIO 23 y 24; estos pines debieron ser configuradas en el mismo código del programa implementado para utilizar su resistor de pull-up interno y ser configurados como entradas. Finalmente se añadió un set de audífonos, a través de los cuales se pueden escuchar las descripciones generadas del entorno. A fin de

proveer portabilidad, se le incorporó una batería portátil de equipos celulares a través del puerto micro USB. El prototipo final implementado, con todos sus componentes funcionales operativos instalados se puede observar con detalle en la Figura 9.



Figura 9: Prototipo Implementado

Por su carácter de licencia pública de uso y facilidad de instalación se utilizó la aplicación de conversión de voz a texto eSpeak, a fin de poder generar las indicaciones audibles que van a servir de guía a las personas con deficiencia visual que requieran usar el prototipo. Finalmente se hizo prueba funcional del código y se configuró el Raspberry Pi para ejecutar este programa cada vez que se encendiese el mismo de forma automática con la intención de poder hacer uso de la aplicación sin necesidad de conectar el equipo a ninguna pantalla ni ordenador. En la Figura 10 se muestra el diagrama de flujo simplificado del código implementado en el computador de placa reducida.

3. Análisis y discusión de resultados

A fin de hacer la medición cuantitativa de las descripciones generadas por el modelo de predicción entrenado, se escogieron 10 imágenes de forma aleatoria del conjunto de pruebas; 5 de ellas se muestran en la Figura 11 como referencia. Es importante indicar que estas imágenes deben pertenecer a una muestra estadística similar a las imágenes usadas para entrenamiento a fin de lograr un equilibrio Bias-Variance que minimizase el factor reducible del error esperado [18]. Esta es

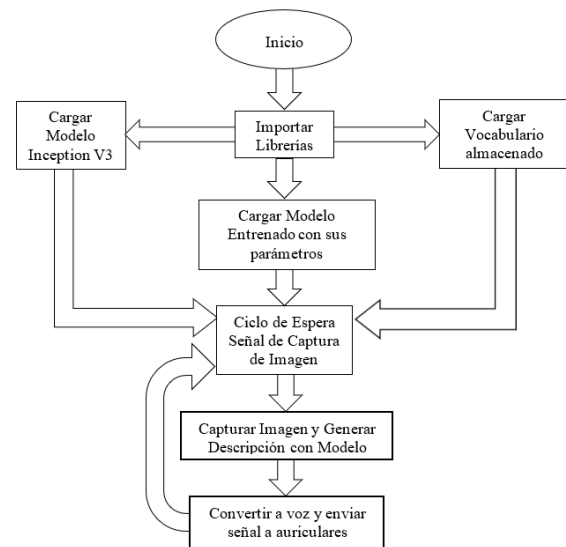


Figura 10: Diagrama de Flujo del Código Implementado en Raspberry Pi

la manera de garantizar la pertinencia del modelo; para ello este set no debe ser utilizado en ninguna de las fases de aprendizaje y debe solo ser reservado para la fase de pruebas.

A nivel de sistemas de traducción automáticos basados en sistemas de inteligencia artificial, e incluso en algunos trabajos de descripción de imágenes se han utilizado diversas técnicas de análisis automático que aumentan la rapidez de las pruebas. Sin embargo, partiendo de las premisas generadas por Vinyals [11], la métrica utilizada para cuantificar la naturalidad y correspondencia de las descripciones generadas para poder medir las sutilezas de lenguaje que un ser humano es capaz de detectar debe ser realizada por una evaluación humana de las mismas. Debido a la indisponibilidad de individuos cuya lengua materna fuese el idioma inglés, la población utilizada para evaluación estuvo compuesta por 5 personas bilingües cuyo idioma materno era el español. La evaluación de las descripciones se hizo a través de una única pregunta, que daba un ranking del 1 al 4 a la calidad de la descripción generada para cada una de las diez imágenes a evaluar.

En la Figura 12 se puede observar la gráfica de distribución normal inversa acumulada de los resultados, tanto para las valoraciones dadas a las descripciones de referencia (las que acompañan



Figura 11: Algunas Imágenes para Validación del Modelo Entrenado

al set de datos) como para valoraciones dadas a las descripciones generadas por el modelo desarrollado. Allí se puede observar que las puntuaciones dadas en general para las descripciones que acompañaban a cada una de las imágenes fue superior a las generadas por el modelo. Sin embargo, este comportamiento era de esperarse si se compara con los resultados mostrados en la Figura 13, y que se corresponden con la publicación original del modelo que sirvió de base para el presente trabajo de investigación.

El promedio de puntuación dado para las imágenes que acompañaban a cada imagen fue de 3.54, en comparación con el valor de 3.89 en el estudio de referencia [11]. En el mismo orden de ideas, el promedio de puntuación dado

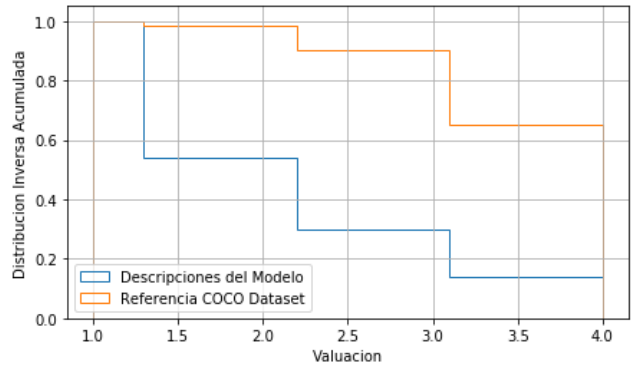


Figura 12: Gráfico de Análisis de Resultados

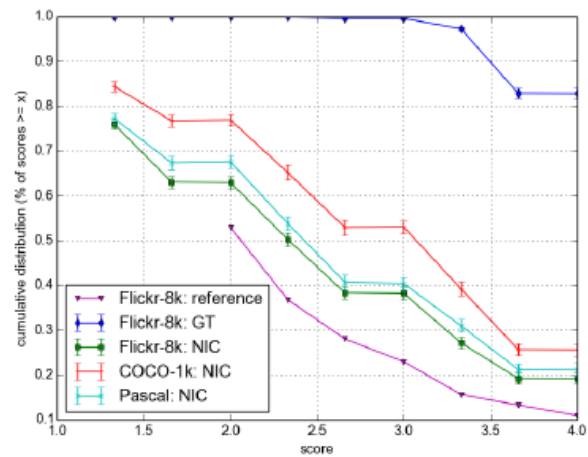


Figura 13: Gráfico de Resultados de Referencia [11]

para las imágenes generadas por el modelo que se desarrolló en el presente trabajo fue de 1,98; en comparación con valores que van desde 2,08 en el estudio referencial. Igualmente se puede observar que los comportamientos de ambas Figura 12 y Figura 13 es similar en cuanto a la forma de la distribución inversa acumulada de probabilidad.

Para validar el funcionamiento del sistema construido con el Raspberry Pi se procedió a la recolección de un pequeño conjunto de imágenes con sus respectivas descripciones en tiempo real, en diversos entornos de tipo interior y exterior. En las Figuras 14 y 15 se expone una muestra de 6 imágenes con sus respectivas descripciones, las cuales fueron obtenidas directamente del prototipo implementado. Queda para trabajos futuros la evaluación de estas descripciones.

Se pudo observar que las descripciones generadas por las imágenes recolectadas por el prototipo funcional de ajustaron de forma bastante aproximada a los elementos capturados y la relación entre ellos. Cabe destacar la capacidad de generalización para describir imágenes con un vocabulario y cantidad de ejemplos relativamente limitados.



Figura 14: Imágenes obtenidas del Prototipo 1-3



Figura 15: Imágenes obtenidas del Prototipo 4-6

4. Conclusión

El presente trabajo demostró la factibilidad de utilizar modelos basados en redes neuronales para la obtención de descripciones del entorno en tiempo real, que sean capaces de ayudar a personas con impedimento visual a mejorar su calidad de vida. En particular, se pudo evidenciar que es posible la obtención de modelos cuyo tamaño y complejidad computacional una vez entrenado sean adecuados para su implementación en computadores de placa reducida, los cuales son una alternativa actual de bajo costo para el desarrollo de prototipos orientados al mercado de personas con bajos recursos económicos.

El nivel de precisión obtenido en las descripciones predichas por un modelo de parámetros reducidos (a fin de generar tiempos de entrenamiento razonables en un computador con un procesador de gama media y sin GPU incorporado) demostró ser similar al obtenido por el estudio referencial [11]. Los tiempos de respuesta entre la captura de la imagen y la obtención de la respuesta en el prototipo fue inferior a los 10 segundos, lo cual resultan ser tiempos poco manejables para aplicaciones que generan descripciones en tiempo real, pero que con implementaciones más orientadas al uso en usuarios finales pueden ser mejorados.

El aspecto más relevante del presente trabajo es el precedente que se genera en cuanto a la posibilidad de desarrollar sistemas de bajo costo, sin necesidad de conexión dedicada a internet y con software de licenciamiento público para la solución de un problema que puede mejorar la vida de más de 75 millones de personas a nivel mundial de forma directa, y de muchos más en forma indirecta (cuando se incluyen los familiares y amigos que sirven de apoyo continuo para soporte vital de las personas con diversos grados de impedimento visual).

El desarrollo de nuevos modelos que mejoren la precisión en la generación de descripciones a partir de una imagen dada, el incremento del tamaño de los datos disponibles tanto para entrenamiento como para validación del entrenamiento del modelo implementado, el aumento de la capacidad

de cálculo en los sistemas computacionales actuales, y la disminución de costos asociado con el desarrollo de nuevos sistemas de implementación de placa reducida, son todos factores que apuntan a la intensificación de los esfuerzos en aras de conseguir soluciones prácticas y factibles para ayudar a todas las personas de bajos recursos en países de bajos ingresos a ser incorporados de manera paulatina al mercado laboral global.

Es la intención que los resultados obtenidos en el presente trabajo, en cuanto a la factibilidad de implementación y la utilización de modelos de máquinas de aprendizaje para la obtención de descripciones del entorno, sirvan como aporte para la mejora futura de los millones de personas que pueden servirse de la técnica y el desarrollo tecnológico para mejorar su calidad de vida y tener la oportunidad de valerse por sí mismos por encima de las limitaciones que alguna discapacidad visual pueda intentar poner en su horizonte.

5. Referencias

- [1] Organización Mundial de la Salud, *Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud*. Décima revisión. Organización Panamericana de la Salud, 2008.
- [2] N. Chenthamil, N. Rekha, and P. Poovizhi, “Portable Camera Based Identification System for Visually Impaired People,” *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 19, no. 3, pp. 19 141–19 146, 2016.
- [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, USA: The MIT Press, 2016.
- [4] S. Ramón y Cajal, “Estructura de los Centros Nerviosos de las Aves,” *Revista trimestral de Histología Normal y Patológica*, vol. 1, no. 1, pp. 314–318, 1888.
- [5] D. Rumelhart, G. Hinton, and J. McClelland, “A General Framework for Parallel Distributed Processing,” *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, 1986.
- [6] D. Hubel and T. Wiesel, “Receptive fields of single neurones in the cat’s striate cortex,” *The Journal of physiology*, vol. 148, no. 3, pp. 574–591, 1959.
- [7] K. Fukushima, “Neocognitron: A hierarchical neural network capable of visual pattern recognition,” *Neural Networks*, vol. 1, no. 2, pp. 119–130, 1988.
- [8] D. Ciresan, M. Ueli, J. Masci, L. Gambardella, and J. Schmidhuber, “Flexible, High Performance Convolutional Neural Networks for Image Classification,” in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, vol. 2, 2011, pp. 1237–1242.
- [9] Y. Bengio, P. Simard, and P. Frasconi, “Learning Long-Term Dependencies with Gradient Descent is Difficult,” *IEEE Transactions On Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [10] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [11] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, “Show and Tell: A Neural Image Caption Generator,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3156–3164.
- [12] J. Donahue, L. Hendrick, M. Rohrbach, S. Venugopalan, S. Guadarrama, K. Saenko, and T. Darrell, “Long-term Recurrent Convolutional Networks for Visual Recognition and Description,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 677–691, 2017.
- [13] T. Yao, P. Yingwei, Y. Li, Z. Qiu, and T. Mei, “Boosting Image Captioning With Attributes,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 4904–4912.
- [14] X. Lu, B. Wang, X. Zheng, and X. Li, “Exploring Models and Data for Remote Sensing Image Caption Generation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 4, pp. 2183–2195, 2018.
- [15] D. Velmurugan, M. Sonam, S. Umamaheswari, S. Parthasarathy, S. Guadarrama, K. Saenko, and T. Darrell, “A Smart Reader for Visually Impaired People Using Raspberry PI,” *International Journal of Engineering Science and Computing*, vol. 6, no. 3, pp. 2997–3001, 2016.
- [16] R. Ardiansyah, “Design of An Electronic Narrator on Assistant Robot for Blind People,” in *MATEC Web of Conferences*, vol. 42, no. 03013, 2016, pp. 03 013p.1–03 013p.5.
- [17] C. Elamri and T. de Planque, “Automated Neural Image Caption Generator for Visually Impaired People,” *IOSR Journal of Engineering (IOSRJEN)*, vol. 10, pp. 28–33, 2018.
- [18] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. New York, USA: Springer, 2015.

Detection of incipient faults of bearings in a Generator Synchronous using Support Vector Machines (SVMs)

Aniello Sparano , Jesús Ramírez , Ledy Guerra* , Rubén Terán 

Departamento de Potencia, Escuela de Ingeniería Eléctrica, Universidad de Carabobo, Valencia, Venezuela

La selección de este artículo fue realizada en el marco de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica "Prof. César Rodolfo Ruiz" Octubre 2020, siendo la evaluación, arbitraje, aceptación y edición a cargo de Revista Ingeniería UC.

Abstract.- Since the synchronous machine in its operation as a generator constitutes a vital element in electrical power systems, programs development and predictive maintenance techniques, in order to identify and fix problems in the machine before irreversible damage is done, is priority. A methodology is established below to detect incipient failures in a tested synchronous generator bearings using Support Vector Machines (SVM). Phase currents analysis is presented, using various methods for data reduction and matching algorithms learning based on support vector machines as data classifier element, in order to obtain a model that is capable of discriminate the condition of a machine, as far as mechanical failures are concerned. The training of support vector machines to obtain the classification criteria was executed using MatLab's SVM commands. The functions provided by this software allowed: training, optimization and subsequent classifier validation. The results obtained showed that the method applied based on SVM has a good fault recognition ability presented in the different components of the bearing, for different operating conditions.

Keywords: synchronous generator; failures; bearings; armature current; support vector machines (SVM); predictive maintenance.

DetECCIÓN DE FALLAS INCIPIENTES EN RODAMIENTOS DE GENERADORES SINCRÓNICOS UTILIZANDO MÁQUINAS DE VECTORES DE SOPORTE

Resumen.- Dado que la máquina sincrónica en su funcionamiento como generador constituye un elemento vital en los sistemas eléctricos de potencia, el desarrollo de programas y técnicas de mantenimiento predictivo, a fin de identificar y solucionar problemas en la máquina antes de que se produzcan daños irreversibles en la misma es prioridad. A continuación se establece una metodología para detectar fallas incipientes en rodamientos de un generador sincrónico de prueba a través de máquinas de vectores de soporte (SVM por sus siglas en inglés). Se presenta el análisis de las corrientes de fase, utilizando métodos variados para la reducción de los datos y algoritmos de aprendizaje automático basados en las máquinas de vectores de soporte como elemento clasificador de datos; a fin de obtener un modelo que sea capaz de discriminar la condición de una máquina, en cuanto a fallas mecánicas se refiere. El entrenamiento de las máquinas de vectores de soporte para obtener los criterios de clasificación se ejecutó utilizando los comandos para SVM de MatLab, las funciones que brinda este software permitieron: el entrenamiento, la optimización y la posterior validación de los clasificadores. Los resultados obtenidos demostraron que el método aplicado basado en la SVM posee una buena capacidad de reconocimiento de fallas presentadas en los diferentes componentes del rodamiento, para distintas condiciones de operación.

Palabras clave: generador sincrónico; fallas; rodamientos; corriente de armadura; máquinas de vectores de soporte; mantenimiento predictivo.

Recibido: 22 de octubre, 2020.

Aceptado: 23 de enero, 2021.

1. Introducción

Los generadores sincrónicos son esenciales en el proceso de conversión de energía eléctrica, por lo que su uso hoy en día se ha vuelto indispensable para el desarrollo y funcionamiento de la sociedad tal como la conocemos. Este tipo de máquinas

* Autor para correspondencia:
Correo-e: llguerra@uc.edu.ve (L. Guerra)

está propensa a sufrir tanto fallas externas como internas, siendo que estas últimas pueden ocurrir tanto en el rotor como en el estator de la máquina. Las principales fallas que pueden presentarse en estas máquinas son: [1, 2, 3]

- Fallas en rodamientos. (40–50 %)
- Fallas en aislamiento del estator. (30–40 %)
- Fallas en el devanado de amortiguamiento. (5–10 %)
- Excentricidad y otras.

Los rodamientos son elementos sobre los que se apoyan los dos extremos del eje que sustenta a la parte móvil de toda máquina eléctrica rotativa. Por consiguiente cualquier desperfecto que pueda presentar la máquina se traduce en esfuerzos mecánicos sobre estos dispositivos, razón por la que son los componentes con el porcentaje de falla más elevado [1].

El resultado de estas fallas se traduce en un daño o deterioro de la máquina, el cual puede impedir su óptimo desempeño y acarrear como consecuencias reducción del rendimiento, degradación en la calidad de energía generada así como disminución de la producción, desbalances en el sistema eléctrico, originar fallas más severas, entre otras; razón por la cual resulta necesario determinar la existencia de las fallas en estado incipiente.

El desarrollo de técnicas para supervisión de las condiciones de operación de las máquinas eléctricas rotativas, permite monitorear las variables de estado de la máquina a fin de detectar problemas incipientes antes de que se presenten fallas catastróficas. Se han desarrollado estudios y técnicas enfocadas en la detección, predicción y diagnóstico de fallas a través de sistemas de monitoreo continuo mientras la máquina esta en operación, mediante la medición de distintas variables y parámetros que representen síntomas de mal funcionamiento, con la finalidad de evitar fallas peligrosas, paradas de emergencia, preservar la vida útil del equipo y reducir costos en la producción y el mantenimiento, [4, 5]. En este sentido, los resultados obtenidos en esta investigación muestran que con las SVM se puede discriminar el estado de los rodamientos de una máquina sincrónica en su estado incipiente, lo que

contribuye a establecer planes de mantenimiento predictivo asociados a dichas máquinas.

2. Campos magnéticos en máquinas sincrónicas

Las máquinas sincrónicas son máquinas eléctricas donde la velocidad mecánica de rotación η se encuentra entrelazada o sincronizada con la frecuencia de la red eléctrica f en la cual están operando. La relación entre la velocidad de rotación de los campos magnéticos de la máquina y la frecuencia eléctrica del estator se expresa mediante la ecuación (1).

$$\eta = \frac{120}{p} f, \quad (1)$$

donde:

η : Velocidad de rotación, rpm.

f : Frecuencia de la red, Hz.

p : número de polos de la máquina, adim.

El campo magnético en el entrehierro de una máquina eléctrica es el resultado de las Fuerzas Magnetomotriz (Fmm) combinadas de los devanados inductor (normalmente en el rotor de la máquina sincrónica) e inducido normalmente en el estator de la máquina sincrónica), que actúan en esa región. En principio, el devanado inductor es el que produce el campo magnético en el entrehierro (campo del rotor), induciendo Fuerzas Electromotrices (Fem) en el devanado del inducido, que dan lugar a corrientes cuando se cierra el circuito. Al circular estas corrientes se crea una Fmm de reacción de inducido o de armadura que produce un campo magnético de armadura (campo del estator) que al combinarse con la F.m.m. del inductor, produce una Fmm resultante que produce un campo magnético resultante en el entrehierro de la máquina. En las máquinas eléctricas de corriente alterna se puede aproximar a que la distribución espacial de los campos magnéticos a lo largo del entrehierro, es sinusoidal, tal como se muestra en la Figura 1.

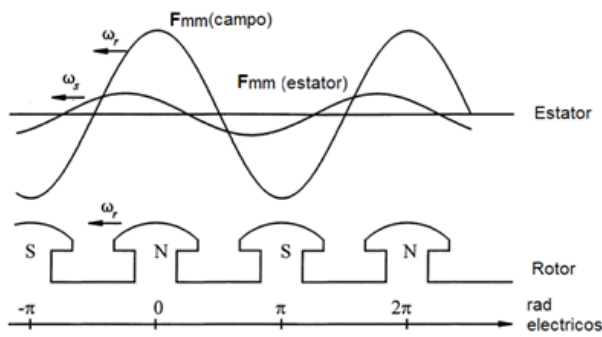


Figura 1: Distribución espacial de las Fuerzas Magnetomotrices en el entrehierro de la Máquina Síncrona [6]

3. Fallas en rodamientos

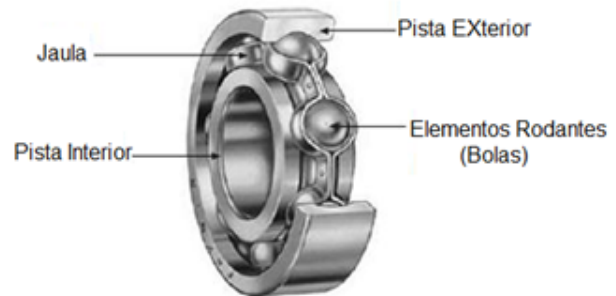
El rodamiento es una pieza que realiza una conexión móvil entre dos elementos de un mecanismo permitiendo la rotación de un elemento respecto al otro, con precisión y mínima fricción. En las máquinas eléctricas, los rodamientos deben soportar y mantener el eje del rotor, garantizando que el entrehierro de la máquina sea de un tamaño y distribución, adecuado y uniforme, de forma tal que se transfieran cargas mecánicas del eje a la carcasa de la máquina. Estos elementos deben operar a altas y bajas velocidades, minimizando la fricción y reduciendo las pérdidas de energía. En máquinas pequeñas y horizontales, los rodamientos que comúnmente se emplean son los rodamientos rígidos de bolas [7, 5].

Todos los rodamientos están compuestos básicamente por: pista exterior (PE), pista interior (PI), elementos rodantes (ER) y jaulas. En la Figura 2 se indican las partes de un rodamiento y se muestra que los elementos rodantes (bolas y rodillos), se encuentran alojados entre los aros. La jaula asegura la rodadura de los elementos.

- Pistas (Exterior e Interior): Los aros o pistas de los rodamientos soportan cargas que transmiten a los asientos de los ejes o estructuras de soporte tales como los alojamientos. Estos aros presentan un camino de rodadura endurecida donde se localizan y en el que rotan los elementos rodantes.
- Elementos Rodantes: Los elementos rodantes se clasifican en dos grupos: bolas y rodillos.

Los rodillos, a su vez, pueden ser: cilíndricos, de agujas, cónicos y esféricos.

- Jaula: Las jaulas se utilizan para contener a las bolas o rodillos y mantener la distancia entre los mismos en su dirección circunferencial. Estas se fabrican por prensado o forjado con aceros más blandos, poliamida, resina fenólica, latón o bronce.



Fuente: SKF (2004) Rolling Bearings in Electric Motors and Generators

Figura 2: Partes de un Rodamiento

Como toda pieza mecánica, los rodamientos tienen un tiempo de vida útil. Este tiempo está asociado a la velocidad de operación así como a la carga dinámica y estática que pueden soportar. Asimismo, el tiempo de vida depende de las condiciones de operación presentes, por lo que alteraciones en las mismas pueden producir fallas en los rodamientos, comprometiendo el desempeño del proceso de la máquina. Entre las principales causas de fallas en rodamientos se tiene [8, 7]:

- Carga excesiva.
- Desgaste por lubricación deficiente o excesiva.
 - Desgaste ocasionado por partículas abrasivas.
 - Desgaste ocasionado por lubricación inadecuada.
- Exposición a vibración mientras el rodamiento no se encuentra rodando.
- Contaminación externa.
- Instalación Incorrecta.
- Fatiga.
- Cambios bruscos de velocidad.

4. Máquinas de Vectores de Soporte (SVM por sus siglas en inglés)

Las máquinas de vectores de soporte (SVM, del inglés Support Vector Machines) fueron concebidas como un algoritmo para resolver problemas de clasificación binaria, hoy en día se han desarrollado como una técnica robusta para clasificación y regresión aplicadas a grandes conjuntos de datos complejos, apoyándose en criterios de generalización y optimización matemática. La SVM consiste en un modelo de clasificación que sirve como herramienta para distinguir objetos pertenecientes a distintas categorías o clases. Para obtener el modelo de clasificación es necesario cumplir dos fases: entrenamiento y prueba [9, 10, 11].

En la fase de entrenamiento, en base a las propiedades del conjunto inicial de datos, se seleccionan los parámetros que serán utilizados para definir la separación de clases. En esta fase, el algoritmo de aprendizaje define, en función a los ejemplos de entrenamiento, los elementos cuya clase es conocida, un hiperplano de separación, el cual establece el límite para distinguir entre los elementos pertenecientes a cada clase. Este plano se define a través de vectores construidos por los puntos más cercanos de cada clase al límite de separación, los cuales son llamados vectores de soporte. Posteriormente, en la fase de prueba, el modelo de separación entre clases, es aplicado a un conjunto de muestras (conjunto de pruebas) cuyas etiquetas de clase se desconocen, con el objetivo de determinar su clase. A continuación se describen los modelos de clasificación considerando que los datos son linealmente separables, que los datos son cuasi separables linealmente y que los datos no son separables linealmente [1, 12].

El modelo más simple de las SVM, conocido como clasificador de margen máximo, funciona sólo para datos linealmente separables en el espacio de características, es decir, se tiene un conjunto de datos en el que cada punto de entrenamiento $x_i \in R^n$ se corresponde con alguna de dos posibles clases, las cuales se designan con un valor $y_i \in \{-1, 1\}$ para $i = 1, \dots, l$.

Inicialmente, para un conjunto de entradas

separables linealmente, denotado como $S = (x_1; y_1); \dots; (x_n; y_n)$, se puede definir un hiperplano de separación como una función lineal de la forma indicada en la ecuación (2).

$$D(x) = \langle w \cdot x \rangle + b. \quad (2)$$

Donde w representa el vector normal del plano y b es llamado el intercepto del plano. El vector normal define la orientación del plano en el espacio, mientras que la relación entre la norma del vector w y b , define la distancia entre el hiperplano y el origen del espacio. Ambos parámetros son coeficientes reales, además el hiperplano de separación cumplirá con las restricciones indicadas en la ecuación (3) para todo x_i del conjunto de ejemplos:

$$\begin{aligned} \langle w, x_i \rangle + b &\geq 0 & \text{si } y_i = +1 \\ \langle w, x_i \rangle + b &\leq 0 & \text{si } y_i = -1, i = 1, \dots, n. \end{aligned} \quad (3)$$

En la Figura 3 se muestran los posibles hiperplanos que permiten la separación de los datos. Como se observa, el plano de separación no es único. Para ello se hace necesario fijar un criterio de optimización, entonces se define el margen (τ) de un hiperplano, el cual representa la distancia mínima entre el plano y los ejemplos de cada clase más cercanos al mismo; por consiguiente un hiperplano es óptimo si su margen es máximo, en otras palabras, el hiperplano de separación a seleccionar, será aquel que entre el conjunto de infinitos planos, guarde la máxima distancia entre él y los ejemplos más cercanos de ambas clases.

La distancia entre el plano y un punto de ejemplo x' viene dada por la ecuación (4).

$$\frac{D(x')}{\|w\|}. \quad (4)$$

Tomando en cuenta las restricciones de la ecuación (2), todos los ejemplos de entrenamiento cumplirán con la ecuación (5).

$$\frac{y_i D(x_i)}{\|w\|} \geq \tau, \quad \text{con } i = 1, \dots, n. \quad (5)$$

Por consiguiente, encontrar el hiperplano óptimo equivale a encontrar el valor de w que maximice

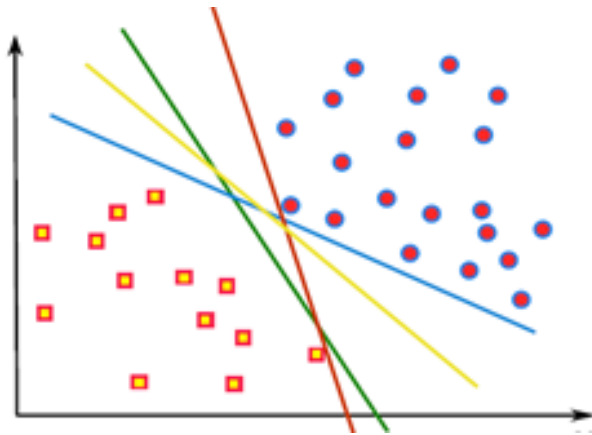


Figura 3: N posibles hiperplanos de separación en un espacio bidimensional [11]

el margen. Comprendiendo que existen infinitas soluciones que difieren solo en la escala de w , se reduce el número de posibles soluciones a una sola, fijando arbitrariamente el valor del producto de τ y la norma w , igual a la unidad, quedando como la ecuación (6).

$$\tau \|w\| = 1. \tag{6}$$

De esta manera, aumentar el margen equivale a disminuir el valor de la norma del vector w . La maximización del margen estará directamente relacionada con la capacidad de generalización del hiperplano de separación, de forma tal que a mayor margen, existirá mayor distancia de separación entre las clases. Por lo tanto, un hiperplano óptimo de separación, de acuerdo con su definición, es aquel que posee un margen máximo, en consecuencia, un mínimo valor de $\|w\|$, sujeto la restricción indicada en la ecuación (7).

$$y_i(\langle w, x_i \rangle + b) \geq 1, i = 1, \dots, n; y_i \in (-1, 1). \tag{7}$$

La restricción presentada reflejará una relación de igualdad para los ejemplos que están situados a ambos lados del hiperplano óptimo y que definen el margen; estos reciben el nombre de vectores soporte. Puesto que estos ejemplos son los más cercanos al hiperplano de separación, serán los más difíciles de clasificar y, por tanto, deberían ser los únicos ejemplos a considerar al momento de construir ese hiperplano. Estos puntos se observan en la Figura 4.

En este sentido, para encontrar el hiperplano equidistante a las dos clases linealmente separables

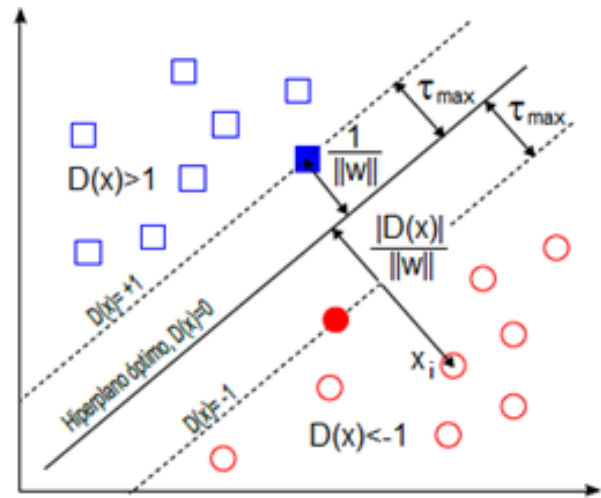


Figura 4: Margen máximo de separación e hiperplano óptimo de ejemplos linealmente separables [10]

en el espacio de entrada, es necesario encontrar solución al problema de optimización de la ecuación (8).

$$\begin{aligned} \text{Max} &\rightarrow \frac{1}{\|w\|} \\ \text{sujeto a } &y_i(\langle w, x_i \rangle + b) \geq 1, i = 1, \dots, n. \end{aligned} \tag{8}$$

Que se puede escribir como la ecuación (9).

$$\begin{aligned} \text{Min} &\rightarrow \frac{1}{2} \langle w, w \rangle \\ \text{sujeto a } &y_i(\langle w, x_i \rangle + b) \geq 1, i = 1, \dots, n. \end{aligned} \tag{9}$$

De acuerdo a la teoría de optimización, el problema primal tiene una forma dual si la función objetivo y las restricciones son funciones que no tienen máximos o mínimos locales. De esta manera, resolver el problema dual implica obtener la solución del problema primal.

Inicialmente se construye un problema de optimización sin restricciones, utilizando el Lagrangiano del problema primario, como la ecuación (10).

$$\begin{aligned} \mathcal{L}(w, b, \alpha) &= \frac{1}{2} \langle w, w \rangle \\ &- \sum_{(i=1)}^n \alpha_i (y_i(\langle w, x_i \rangle + b) - 1). \end{aligned} \tag{10}$$

Donde $\alpha_i \geq 0$ representa los multiplicadores de Lagrange, los cuales indican la dificultad intrínseca en el cumplimiento de cada una de las restricciones de la función objetivo, mientras mayor es el valor del multiplicador, más difícil es cumplir con la restricción asociada a $(y_i(\langle w, x_i \rangle + b) - 1)$.

El siguiente paso consiste en aplicar las condiciones de Karush – Kuhn – Tucker, indicadas en las ecuaciones (11), (12) y (13).

$$\frac{\partial \mathcal{L}(w^*, b^*, \alpha)}{\partial w} = w^* - \sum_{i=1}^n \alpha_i x_i y_i = 0, \quad (11)$$

$$\frac{\partial \mathcal{L}(w^*, b^*, \alpha)}{\partial b} = \sum_{i=1}^n \alpha_i y_i = 0, \quad (12)$$

$$\alpha_i (1 - y_i(\langle w^*, x_i \rangle + b^*)) = 0. \quad (13)$$

Obteniéndose las ecuaciones (14) y (15).

$$w^* = \sum_{i=1}^n \alpha_i x_i y_i, \quad (14)$$

$$\sum_{i=1}^n \alpha_i y_i = 0. \quad (15)$$

Sustituyendo en la ecuación (9) se obtiene la ecuación (16).

$$\mathcal{L} = \sum_{(i=1)}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j \langle x_i \cdot x_j \rangle. \quad (16)$$

Se ha transformado el problema primal en su correspondiente dual que consiste en maximizar la ecuación (16) sujeto a las restricciones de la ecuación (15) y a que $\alpha_i \geq 0, 1, \dots, n$.

De esta manera el vector de peso w^* queda en función de α^* como se muestra en la ecuación (17).

$$w^* = \sum_{i=1}^n \alpha_i y_i x_i. \quad (17)$$

El valor de b no aparece en el problema dual por lo que b^* se obtendrá a partir de la ecuación (13), en la que se cumple la condición de $\alpha_i > 0$, como se indica en la ecuación (18).

$$y_i(\langle w^*, x_i \rangle + b^*) = 1. \quad (18)$$

El ejemplo (x_i, y_i) satisface la condición del problema primal de la ecuación (9). Adicionalmente, los ejemplos del entrenamiento que satisfacen la condición de igualdad, es decir aquellos que tienen multiplicadores de Lagrange distintos de cero, serán un vector de soporte. Despejando b^* de la ecuación (18) se tiene la ecuación (19).

$$b^* = y_{v_s} - \langle w^*, x_{v_s} \rangle. \quad (19)$$

Donde (x_{v_s}, y_{v_s}) representa al vector soporte con su respectiva clase. Sin embargo en la práctica se obtiene el valor de b^* a partir del promedio de todos los vectores de soporte, tal como se indica en la ecuación (20).

$$b^* = \frac{1}{N_{v_s}} \sum_{i=1}^{N_{v_s}} (y_{v_s} - \langle w^*, x_{v_s} \rangle). \quad (20)$$

Conocidos w^* y b^* el hiperplano óptimo se puede expresar de acuerdo a la ecuación (21).

$$D(x) = \sum_{i=1}^n \alpha_i^* y_i \langle x, x_i \rangle + b^*. \quad (21)$$

En algunos casos los datos no son linealmente separables en el espacio de entrada, por lo que la estrategia a seguir es permitir la clasificación errónea de algunos ejemplos del conjunto de datos de entrenamiento. Se considera que un ejemplo es no separable si no cumple la condición en (7), lo que equivale a dos posibles opciones; la primera donde el ejemplo cae dentro del margen asociado a la clase correcta, de acuerdo con la frontera de decisión establecida por el hiperplano de separación seleccionado; y la segunda donde el ejemplo cae en el lado equivocado de la frontera de decisión.

Para abordar este problema, se añade a la condición del hiperplano de separación, un conjunto de variables reales y positivas, denominadas variables de holgura, $\varepsilon_i, = 1, \dots, n$; estas permitirán cuantificar el número de ejemplos no separables que se está dispuesto a admitir, quedando la restricción del hiperplano de acuerdo a la ecuación (22).

$$y(\langle w, x_i \rangle + b) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0, \quad (22)$$

$$i = 1, \dots, n; y_i \in (-1, 1).$$

Para un ejemplo (x_i, y_i) , ε_i representa la desviación en caso de ser un ejemplo no separable, por lo que una variable de error de valor cero corresponde a un ejemplo separable, si es mayor que cero corresponde a un ejemplo no separable y mayor que uno corresponde a un ejemplo no separable y mal clasificado. La sumatoria de todas las variables de error permite tener una idea del precio asociado al número de ejemplos que son no separables; donde mientras mayor sea esta sumatoria, menor es la precisión de la clasificación. En la Figura 5 se muestra margen máximo de separación con la adición de la variable de error.

En la ecuación (23) y (24) se muestra la función a optimizar, la cual cuenta con C , un parámetro contante elegido por el usuario para controlar el grado de influencia del término de error de ejemplos no separables (ε_i) que permite regular la proporción entre el sobreajuste del clasificador final y el número de ejemplos no separables. En tal sentido, si C tiene un valor muy alto, permitirá tener valores de ε_i pequeños, donde si el límite de C se tender a infinito, todos los ε_i serán iguales a cero; por lo que se estaría considerando que el total del conjunto de datos es separable lo que acarrea un sobreajuste del clasificador. Si por el contrario el valor de C es muy pequeño, en el caso límite tendiendo a cero, se estaría permitiendo que los ejemplos tengan asociados valores de ε_i muy grandes, por lo que los ejemplos son mal clasificados. Se considera que $C \sum_{i=1}^n \varepsilon_i$ es un término del coste asociado a la clasificación de los elementos.

$$f(w, \varepsilon) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \varepsilon_i, \quad (23)$$

$$\text{Min} \rightarrow \frac{1}{2} \langle w, w \rangle + C \sum_{i=1}^n \varepsilon_i. \quad (24)$$

Con las restricciones indicadas en la ecuación (25).

$$\begin{aligned} y_i(\langle w, x_i \rangle + b + \varepsilon_i - 1) &\geq 0; \\ \varepsilon_i &\geq 0, i = 1, \dots, n. \end{aligned} \quad (25)$$

Al igual que para el problema de los casos linealmente separables se procede con la metodología

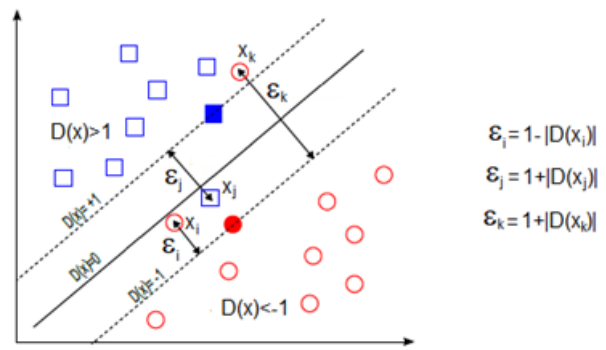


Figura 5: Margen máximo de separación de ejemplos cuasi-separables, con la adición de la variable de holgura [10]

de dualidad mostrada anteriormente obteniendo la ecuación (26).

$$\text{Max} \rightarrow \sum_{(i=1)}^n \alpha_i - \frac{1}{2} \sum_{(i,j=1)}^n \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle. \quad (26)$$

Con las restricciones de la ecuación (27).

$$\sum_{i=1}^n \alpha_i y_i = 0; 0 \leq \alpha_i \leq C; i = 1, \dots, n. \quad (27)$$

La solución a este problema permite expresar la frontera de decisión en términos de los coeficientes α_i^* como se muestra en la ecuación (28).

$$D(x) = \sum_{i=1}^n \alpha_i^* y_i \langle x, x_i \rangle + b^*, \quad (28)$$

donde b^* viene dado por la ecuación (29).

$$b^* = y_i \sum_{j=1}^n \alpha_j^* y_j \langle x_j, x_i \rangle; \quad (29)$$

$$\forall \alpha_i \text{ tal que } 0 < \alpha_i < C.$$

Adicionalmente existen conjuntos de datos que no son linealmente separables, por lo que resulta necesario hacer una transformación del espacio de entrada a un espacio de mayor dimensionalidad utilizando funciones de transformación llamadas Kernel para buscar el hiperplano óptimo en ese nuevo espacio [9, 10, 11].

Una función Kernel es una función $K : X \times X \rightarrow R$ que asigna a cada par de

elementos del espacio de entrada X , un valor real correspondiente al producto escalar interno de las imágenes de los elementos del espacio de entrada, en un nuevo espacio llamado espacio de características. La función Kernel se expresa como se indica en la ecuación (30).

$$\begin{aligned} \phi : X &\rightarrow F \\ K(x \cdot x') &= \langle \phi(x), \phi(x') \rangle = \langle \phi_1(x), \phi_1(x') \rangle \\ &+ \dots + \langle \phi_m(x), \phi_m(x') \rangle. \end{aligned} \quad (30)$$

En la Figura 6 se muestra la transformación del espacio de entrada al espacio de características, donde la función ϕ permite realizar dicha transformación de un espacio donde los datos no son linealmente separables a un espacio donde sí lo son.

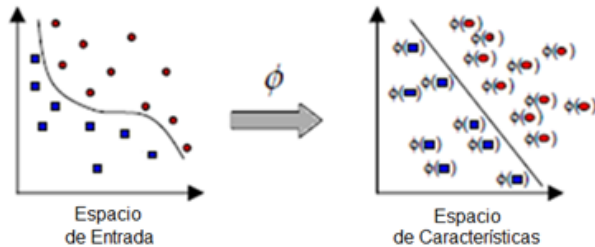


Figura 6: Transformación de un espacio de entrada a un espacio de características [9]

En la Figura 7 se evidencia que los datos que no son linealmente separables en el espacio de entrada de dos dimensiones, pueden separarse linealmente en el espacio tridimensional de características.

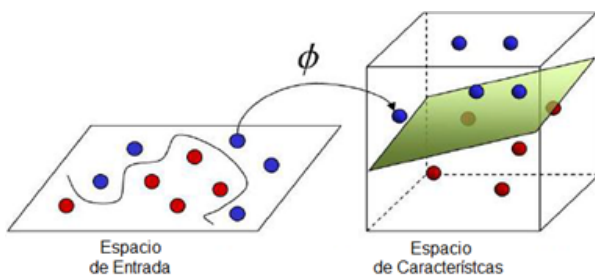


Figura 7: Separación de clases mediante una función Kernel [9]

Como se explicó anteriormente el problema se reduce a la optimización de la ecuación (31).

$$\text{Max} \rightarrow \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j). \quad (31)$$

Con la restricción indicada en la ecuación (32).

$$\sum_{i=1}^n \alpha_i y_i = 0, 0 \leq \alpha_i \leq C, 1 = 1, \dots, n. \quad (32)$$

En la Tabla 1 se muestra las funciones Kernel más utilizadas.

Varios estudios realizados sobre la eficacia del entrenamiento de máquinas de vectores de soporte coinciden en que es recomendable empezar el entrenamiento a través del Kernel RBF, ya que esta función permite realizar mapeos no lineales de ejemplos a una mayor dimensión, a diferencia del Kernel lineal, el RBF puede manejar casos en los que la relación entre las etiquetas y los atributos (datos de entrenamiento) no es lineal. Adicionalmente el Kernel RBF permite facilidades en la selección de parámetros, ya que solo es necesario que el usuario escoja los parámetros C y γ , a diferencia de otras funciones las cuales requieren que se fijen parámetros adicionales.

Tabla 1: Funciones Kernel más comunes [10]

Kernel Lineal	$K(x, x') = \langle x, x' \rangle$
Kernel polinómico de grado p	$K_p(x, x') = (\gamma \langle x, x' \rangle + \tau)^p$
Kernel gaussiano o Función de Base Radial (RBF):	$K_p(x, x') = e^{-\gamma \ x - x'\ ^2}, \gamma > 0$
Kernel Sigmoid	$K(x, x') = \tanh(\gamma \langle x, x' \rangle + \tau)$

Sin embargo, existen algunas situaciones en las cuales el Kernel RBF no es confiable, en particular cuando el número de características es muy grande, se puede usar el Kernel lineal. Si el número de características es grande, tal vez no se necesita mapear la data a un espacio de mayor dimensión. Esto se debe a que el mapeo no lineal no mejora el desarrollo. Usar el kernel lineal es suficiente, y solo se necesita ajustar el parámetro C .

En principio se desconoce cuáles serán los mejores valores de C y γ para un problema determinado, en consecuencia, se debe llevar a cabo un procedimiento para la selección de los mismos. El objetivo de la selección de parámetros es identificar valores (C, γ) de manera que el clasificador realice eficientemente la predicción de datos de prueba.

La forma más común de llevar a cabo el procedimiento es separar los datos en tres partes, donde algunos autores recomiendan que se debe escoger el 50 % de ellos para entrenamiento, 25 % como conjunto de prueba y el 25 % restante como conjunto de validación. Asimismo, otros autores consideran una división que comprende 70 % para entrenamiento y el 30 % restante para verificación. Entonces la eficiencia de la predicción sobre el conjunto desconocido puede reflejar de forma precisa la eficacia del modelo para clasificar en conjunto de datos desconocidos [9, 10, 11, 1, 12].

Una versión desarrollada del proceso de selección es la validación cruzada donde se divide el conjunto de entrenamiento en z subconjuntos de igual tamaño. Seguidamente un subconjunto es probado usando el clasificador entrenado con los $z - 1$ subconjuntos restantes. Por lo tanto, cada instancia del conjunto de entrenamiento completo es predicha una vez, y con ello la eficacia de la validación cruzada representaría el porcentaje de data que ha sido clasificada correctamente.

Es recomendable realizar el ajuste de C y γ usando validación cruzada, donde son probados pares de valores de los parámetros y aquel par con la mayor eficiencia en validación cruzada se selecciona. En este sentido se tiene que probar con secuencias exponenciales crecientes de C y γ , es un método práctico para identificar buenos valores de estos parámetros.

5. Pruebas realizadas

En el Laboratorio de Máquinas Eléctricas de la Universidad de Carabobo se cuenta con un prototipo de máquina sincrónica para el estudio de fallas incipientes. En la Tabla 2 se muestra las características de dicha máquina, en la Figura 8 se muestra el diagrama del montaje y en la Tabla 3 se muestra las especificaciones de los equipos utilizados para las pruebas.

Adicionalmente se utilizaron tres (3) amperímetros DC, tres (3) amperímetros AC, un (1) voltímetro DC, un (1) voltímetro AC y un (1) tacómetro digital a fin de monitorear las condiciones de funcionamiento del circuito. Por otro lado, no fue necesario realizar el

Tabla 2: Características de la Máquina Sincrónica Utilizada [13]

Característica	Unidades
Ranuras	42
Número de polos	14
Devanado	Imbricado doble capa
Paso de bobina	2 ranuras
Paso polar	3 ranuras
Número de espiras	3
Número de bobinas por grupo	1
Número de bobinas por fase	14
Conductor de cobre esmaltado	N°17
Ángulos de ranuras	8.57 °mec, 60 °EE

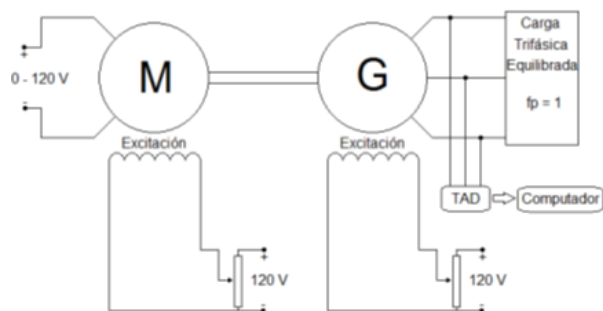


Figura 8: Diagrama del Montaje

condicionamiento de las señales debido a que se obtienen niveles de tensión que se pueden manejar directamente con la Tarjeta de adquisición de Datos (TAD).

Asimismo, para simular las fallas en rodamientos se emplearon 3 cojinetes, que inicialmente se encontraban en perfecto estado, con el fin de simular las fallas en Pista Interna, Pista Externa y Elemento Rodante. Para lograrlo, al primer cojinete se le realizó un agujero en la pista interna, al segundo se le hizo un agujero en la pista externa y para el tercero se realizó la extracción de una de las bolas internas del cojinete [13]. En la Figura 9 se pueden observar las fallas antes indicadas.



Figura 9: Fallas en los cojinetes [13]

Por otro lado, la máquina objeto de estudio se trabajó en dos tipos de conexión (*Y* y *YY*) con tres niveles de carga de tipo resistiva (Vacío, 50%, 75% y 100%) y en cada condición de operación se adquirió las tres corrientes de línea [13]. Seguidamente los datos se organizaron como indica la Figura 10.

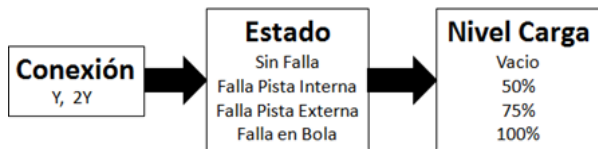


Figura 10: Organización de los datos

Seguidamente se realizó un estudio de los datos para establecer los parámetros característicos a utilizar en el entrenamiento de las SVM prevaleciendo cuatro (04) características a saber: media aritmética, desviación estándar, curtosis y asimetría.

En el entrenamiento de la SVM se proponen dos modelos de clasificación: el primero corresponde a una clasificación binaria (SVM binaria), la cual, al comparar la máquina en condición “sin falla” contra condiciones de falla, haga posible la detección de fallas en rodamientos. El segundo modelo corresponde a una clasificación multiclase (SVM multiclase), la cual apunta a descubrir una falla incipiente, así como identificar la naturaleza de la misma, es decir, detectar y clasificar fallas en los rodamientos del generador. Adicionalmente, para los dos modelos propuestos se utilizó el 70% de los datos totales para el entrenamiento y el 30% restante para validación y prueba. Por otro lado para la clasificación multiclase se utilizaron los siguientes algoritmos:

- 1 VS Resto: para este modelo se generan varias SVM binarias, donde para una de ellas, una de las clases es asignada como positiva y el resto son negativas. El algoritmo consiste en realizar todas las combinaciones posibles y aquella con la función de salida más alta asigna la clase. Aquí, el número de clasificadores es *N*, siendo *N* igual al número de clases.

Tabla 3: Especificaciones de los equipos [13]

Motor DC			
Fabricante	Electrónica Veneta	Modelo	M-12/EV
Tipo	EXC. SEP.	Potencia (W)	450
Voltaje (V)	220	Corriente (A)	2,1
Voltaje de excitación (V)	220	Corriente de excitación (A)	0,33
Velocidad (RPM)	3600	Torque (Nm)	1,5

Fuente de Alimentación Marca Terco		
Selector	Tensión (V)	Corriente (A)
DC fijo	220	3,5
DC variable	0 – 220	16
AC fijo	220	10
AC variable	0 – 220	10

Reóstatos para la Carga Trifásica	
Voltaje (V)	220
Corriente máxima (A)	16
Resistencia (Ω)	1,2

Reóstatos Shunt	
Marca	Terco
Voltaje (V)	220
Corriente Máxima (A)	2
Resistencia (Ω)	440

Tarjeta de Adquisición de Datos (TAD)	
Modelo	USB - 1208FS
Voltaje de alimentación (V)	120
Frecuencia de alimentación (Hz)	60
Nº de canales externos	8

- 1 VS 1: en este caso, se generan múltiples SVM donde, para cada una de ellas, una clase es positiva, otra negativa y el resto de las clases son ignoradas. La estrategia aquí consiste en acumular puntos de aciertos. Al introducir un nuevo ejemplo, cuya clase es desconocida, en cada combinación la SVM predice la clase del ejemplo, y cada vez que una clase es seleccionada, se le suma un punto. Al final, aquella con mayor cantidad de puntos se escoge como la clase ganadora. El número de clasificadores requeridos viene dado por la ecuación (33).

$$\text{Número de SVM} = \frac{N(N - 1)}{2} \quad (33)$$

Siendo *N* el número de clases. Para el diseño del programa de entrenamiento multiclase, se tiene que el número de clases es igual a 4 y haciendo uso de

Tabla 4: Matriz de decisión para el entrenamiento Multiclase

Clase	SVM					
	SVM 1	SVM 2	SVM 3	SVM 4	SVM 5	SVM 6
Sin falla	1	1	1	0	0	0
Falla Pista Interna	-1	0	0	1	1	0
Falla Pista Externa	0	-1	0	-1	0	1
Falla en Bola	0	0	-1	0	-1	-1

Tabla 5: Resultados del entrenamiento en conexión Y

SVM KERNEL	Precisión de la SVM (%)			Precisión SVM Multiclase (%)	
	Clasificación máquina sin falla VS:			Algoritmo de clasificación	
	Falla Pista Interna	Falla Pista Externa	Falla en Bola	1 Vs 1	1 Vs Resto
Linear	57,1	57,1	54,6	30,3	32,1
Quadratic	81	80	67,6	45,8	40,7
Cubic	86,7	87,3	81,3	63,4	60,9
Fine RBF	95,2	93,7	96,2	83,1	81,4
Medium RBF	85,1	89,2	80,5	62,7	61
Coarse RBF	65,1	67,3	57,6	34	41

Tabla 6: Resultados del entrenamiento en conexión YY

SVM KERNEL	Precisión de la SVM (%)			Precisión SVM Multiclase (%)	
	Clasificación máquina sin falla VS:			Algoritmo de clasificación	
	Falla Pista Interna	Falla Pista Externa	Falla en Bola	1 Vs 1	1 Vs Resto
Linear	67,9	68,8	63,2	45,1	34,4
Quadratic	73	87,9	77,8	59	49,7
Cubic	83,2	90,8	80	70,6	61,5
Fine RBF	95,2	96,2	94	89	89,7
Medium RBF	81,1	90,2	82,2	72	70,3
Coarse RBF	63,2	75,6	74,6	48,5	47,4

la ecuación (31), se obtuvo que para el algoritmo 1 VS 1 se requirieron 6 clasificadores individuales. En este sentido la matriz de decisión, en la cual las filas son las clases de estudio y las columnas los clasificadores por paso, queda como se indica en la Tabla 4.

Es posible mejorar el desempeño de clasificación de las SVM al modificar los parámetros de penalización del error (C) y la escala de la función Kernel (γ) utilizada. El parámetro C controla el equilibrio entre generar un margen amplio contra el error en la clasificación de los ejemplos cercanos al mismo, mientras que γ permite realizar ajustes sobre la curvatura de los datos. Se recomienda aplicar una secuencia exponencial geométrica de al menos 10 valores para el parámetro C , por ejemplo $C = 2^{-5}, 2^{-3}, \dots, 2^3, 2^5$. Por el contrario, para el ajuste de γ los software utilizan un procedimiento

heurístico para seleccionar el valor más apropiado para el modelo.

6. Resultados

En las Tablas 5 y 6 se muestran los resultados de precisión de la SVM, al comparar el generador sincrónico tanto para conexión en Y como para conexión YY, en condición sin falla contra las condiciones de falla en pista interna, falla en pista externa y falla en bola, tanto para clasificación simple como para clasificación multiclase, adicionalmente se refleja la variación de la función aplicada de la SVM.

Se observa indudablemente, que la función RBF es la más apta para optimizar el entrenamiento, donde para clasificación binaria se obtuvieron

Tabla 7: Clasificación Binaria para Conexión Y. Función Kernel = RBF

C	z = 5			z = 10		
	Falla Pista Interna	Falla Pista Externa	Falla Bola	Falla Pista Interna	Falla Pista Externa	Falla Bola
0,00193	57,1	57,1	57,1	57,1	57,1	57,1
0,00781	57,1	57,1	57,1	57,1	57,1	57,1
0,03125	57,1	57,3	57,1	57,2	57,3	57,2
0,125	90,7	92,4	89,4	90,6	92,2	90,7
0,5	94,8	93,8	94,4	95,9	93,6	94,4
2	96	95,9	97	96,6	95,5	96,2
8	96,5	96,4	98	97,1	96,7	96,9
32	96,2	95,9	98,2	97	97,0	96,7
128	96,5	95,3	98	96,5	96,1	96,8
512	95,9	94,7	97,9	96,1	95,8	96,6
1000	95,7	95	97,6	96,3	95,6	96,3

Tabla 8: Clasificación Multiclase para Conexión Y. Función Kernel = RBF

C	z = 5		z = 10	
	1 VS 1	1 VS RESTO	1 VS 1	1 VS RESTO
0,00193	30,8	78,3	30,8	79,8
0,00781	30,8	78,8	30,8	80
0,03125	30,8	79,8	30,8	79,7
0,125	73,6	78,9	75,4	80,3
0,5	83	82,8	83,2	82,9
2	86,4	85,6	85,9	85,8
8	87,4	85,3	87	85,8
32	87,2	85,1	87,3	85,7
128	86,5	84,9	86,4	85,1
512	85,5	84,8	86,3	84,9
1000	85,6	84,5	86,1	85

Tabla 9: Clasificación Binaria para Conexión YY. Función Kernel = RBF

C	z = 5			z = 10		
	Falla Pista Interna	Falla Pista Externa	Falla Bola	Falla Pista Interna	Falla Pista Externa	Falla Bola
0,00193	57,1	57,1	57,1	57,1	57,1	57,1
0,00781	57,1	57,1	57,1	57,1	57,1	57,1
0,03125	60,9	79,4	58,7	64,6	75,4	57,3
0,125	81,8	87,8	90,6	86,7	87,7	88,7
0,5	94,2	94,3	92,9	94,6	94,9	94,1
2	96,1	95,9	94,1	95,2	96,7	96,7
8	96	96,5	94,4	96,5	96,5	96,8
32	96	97	95,7	96,1	96,5	96,1
128	95,8	97	95	95,7	95,5	95,3
512	95,5	96,8	94,1	95,6	96,5	95,3
1000	95,5	97	93,1	95,6	96,2	95,3

porcentajes de acierto superiores a 93 % y para clasificación multiclase superior a 83 %.

Adicionalmente, para fijar los parámetros que optimicen el proceso de clasificación, se evalúan los parámetros que controlan la validación cruzada (z) y la minimización del error (C). Este

procedimiento se aplicó para la clasificación binaria y multiclase tanto en conexión Y como en YY, lo que se muestra en las Tablas 7, 8, 9 y 10.

En las Tablas 7 y 8 se observa que para valores de C muy bajos, la clasificación es baja, del orden del 57 % de aciertos para clasificación binaria y en

Tabla 10: Clasificación Multiclase para Conexión YY. Función Kernel = RBF

C	z = 5		z = 10	
	1 VS 1	1 VS RESTO	1 VS 1	1 VS RESTO
0,00193	30,8	83,6	30,8	84,3
0,00781	30,8	83,9	30,8	84,5
0,03125	41,1	83,9	43,6	84,4
0,125	78,3	85,1	80,7	85,9
0,5	89,5	88,8	90,3	89,8
2	91,8	91,8	92,8	92,4
8	92,7	92	93,2	93,1
32	92,5	91,3	92,7	92,6
128	91,8	90,6	92,5	91,5
512	91,5	90,3	92,3	91,6
1000	91,5	90,3	92,3	91,5

Tabla 11: Clasificación de fallas en función del nivel de carga. Función Kernel = RBF; C = 8; z = 10

Conexión	% Carga	% Precisión SVM al comparar máquina “sin Falla” vs:			% Precisión SVM Multiclase	
		Falla Pista Interna	Falla Pista Externa	Falla en Bola	1 vs 1	1 vs Resto
Y	50	99,7	100	100	96,3	96,2
	75	99,7	96	97,7	91	91,7
	100	100	100	97,3	91,5	88,8
YY	50	94,7	99	99	95,2	96,7
	75	99,3	97,7	99,7	97,8	96,8
	100	98,7	99,7	98,7	98,7	98,5

el orden de 31 % para 1 vs 1. Para valores mayores de C, el porcentaje de aciertos se incrementó notoriamente, siendo la media del 96,4 % para clasificación binaria y de 85,8 % para clasificación multiclase. La precisión fue aproximadamente igual para todas las fallas individuales, al igual que para todos los valores de z probados. Los mejores resultados se obtuvieron para valores de C en el rango $2 \leq C \leq 32$, sin embargo, para valores de $C > 1$, la variación no es significativa.

Al igual que para conexión Y, en conexión YY cuyos resultados se muestran en las Tablas 9 y 10, se observa que para valores de C muy bajos, la clasificación es baja, del orden del 57 % de aciertos para clasificación binaria y en el orden de 31 % para 1 vs 1. Para valores mayores de C, el porcentaje de aciertos se incrementó notoriamente, siendo la media del 95,8 % para clasificación binaria y de 91,9 % para clasificación multiclase.

La precisión fue aproximadamente igual para todas las fallas individuales, al igual que para todos los valores de z probados. Los mejores resultados se obtuvieron para valores de C en el rango $2 \leq C \leq 1000$ para clasificación binaria y $C = 8$ para

clasificación multiclase; sin embargo, para valores de $C > 1$, la variación no es significativa.

Por otro lado, se consideraron datos de entrenamiento para los niveles específicos de carga de 50 %, 75 % y 100 % obteniéndose los porcentajes de acierto que se muestran en la Tabla 11 observándose que, al entrenar las máquinas de vectores de soporte para un nivel de carga específico, los niveles de acierto se encuentran en el orden del 98 %. Adicionalmente, se observó que al ser comparados contra el entrenamiento sin atender al nivel de carga, los resultados de precisión se encuentran en el mismo rango de valores, lo que significa que la SVM es capaz de discriminar fallas de manera apropiada independientemente del nivel de carga.

7. Conclusiones

Los estudios asociados a fallas representan uno de los aspectos fundamentales a considerar actualmente para la adecuada operación en máquinas sincrónicas, así como para el desarrollo de nuevas tecnologías que permitan perfeccionar

su funcionamiento. Debido a la importancia que adquieren estas máquinas dentro de los sistemas eléctricos, el desarrollo de nuevas técnicas y modelos de mantenimiento predictivo, basadas en métodos de aprendizaje supervisado, tal como la clasificación con SVM, permite establecer mejoras en el proceso de operación y mantenimiento de la máquina, así como la calidad del servicio que con ellas se suministra. En relación a lo anterior, para el entrenamiento de la SVM, se comprobó que organizar adecuadamente la data permite establecer con mayor claridad el problema de clasificación principal y cómo abordarlo, a su vez que permite aprovechar una de las principales ventajas de las SVM, la cual radica en la facilidad para el entrenamiento.

Asimismo, se demostró que, una vez definido el problema de clasificación, así como estructurados apropiadamente los casos de estudio, es posible diseñar y realizar el entrenamiento de una máquina de vectores de soporte para la detección y clasificación de fallas en rodamientos en el generador sincrónico, a partir del análisis de la corriente de armadura como variable de estudio. Debido a las propiedades de este clasificador, se comprobó que es posible su aplicación para el estudio del comportamiento en cuanto condiciones de fallas mecánicas, en lo que a máquinas eléctricas rotativas se refiere, en este caso específicamente en la máquina sincrónica operando como generador.

Finalmente, basado en el análisis de los resultados de los entrenamientos realizados, se concluye que la máquina de vectores de soporte es capaz de reconocer efectivamente la presencia de una falla en los rodamientos de la máquina sincrónica, a través de la observación de la corriente estática, independientemente del nivel de carga y de la conexión del devanado de armadura que se utilice. Adicionalmente es capaz de discriminar si la falla del cojinete se encuentra en la pista interna, pista externa o en bola.

Asimismo, los tres tipos de fallas de cojinete estudiados se discriminan en igual proporción debido a que para los tres tipos de fallas (en la pista interna, en la pista externa y en bola) los porcentajes de acierto se mantuvieron, a lo largo del análisis, en el mismo rango de valores, esto

permite concluir que la SVM discrimina todas las fallas en cojinetes por igual. En líneas generales, los resultados obtenidos son superiores a 85 % para clasificación multiclase y superiores a 96 % para clasificación binaria, por lo que el modelo aplicado y el desempeño demostrado por el clasificador son ciertamente satisfactorios.

8. Referencias

- [1] M. Fernández-Cabamas, M. García-Melero, G. Alfonso-Orcajo, J. Cano-Rodríguez, and J. Solares-Sariego, *Técnicas para el mantenimiento y diagnóstico de máquinas eléctricas rotativas*. ABB Service S.A.-Marcombo Boixareu Editores, 2000.
- [2] I. Tabatabaei, J. Faiz, H. Lesani, and M. Nabavi-Razavi, "Modeling and Simulation of a salient-pole synchronous Generator with dynamic eccentricity using modified winding function theory," *IEEE Transactions on Magnetics*, vol. 40, no. 3, pp. 1550–1555, 2004.
- [3] N. Prabhakar, "Stator fault analysis of Synchronous Machines," Phd. Thesis, Victoria University, 2007.
- [4] P. Tavner, "Review of condition monitoring of rotating electrical machines," *IET Electric Power Applications*, vol. 2, no. 4, pp. 215–247, 2008.
- [5] H. Toliyat, S. Nandi, S. Choi, and H. Meshgim-Kelk, *ELECTRIC MACHINES. Modeling, Condition Monitoring, and Fault Diagnosis*. Taylor & Francis Group, LLC, 2013.
- [6] P. Delgado-Arredondo, D. Morínigo-Sotelo, R. Osornio-Rios, J. Avina-Cervantes, H. Rostro-Gonzalez, and R. Romero-Troncoso, "Methodology for fault detection in induction motors via sound and vibration signals," *Mechanical Systems and Signal Processing*, vol. 83, no. 15, pp. 568–589, 2017.
- [7] Z. Medrano, C. Perez, J. Gomez, and M. Vera, "Nueva metodología de diagnóstico de fallas en rodamientos en una máquina síncrona mediante el procesamiento de señales vibro-acústicas empleando análisis de densidad de potencia," *Ingeniería Investigación y Tecnología*, vol. 17, no. 1, pp. 73–85, 2016.
- [8] M. Castillo and F. Garzón, "Tipología de Fallas en Generadores Síncronos Hidráulicos de la Empresa de Energía del Pacífico," Tesis para optar al título de Ingeniero Electricista, Universidad del Valle, 2016.
- [9] G. Betancourt, "Las Máquinas de Soporte Vectorial (SVMs)," *Scientia Et Technica*, vol. 1, no. 27, pp. 67–72, 2005.
- [10] E. Carmona, *Tutorial sobre Máquinas de Vectores de Soporte (SVM)*. Universidad Nacional de Educación a Distancia (UNED), 2014.
- [11] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.

- [12] L. Cao and W. Chong, “Feature extraction in support vector machine: a comparison of PCA, XPCA and ICA,” in *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02*, vol. 2, 2002, pp. 1001–1005.
- [13] J. Amaya and M. Rangel, “Estudio Preliminar de Fallas en Máquinas Síncronas para la Aplicación de Mantenimiento Predictivo a partir del Análisis del Espectro de Señales Eléctricas.” Trabajo Especial de Grado para optar al título de Ingeiero Electricista, Universidad de Carabobo, 2012.

EEG Signal Clustering With Learned Features Using Deep Autoencoder

Sergio Villazana ^{*a} , César Seijas ^a , Guillermo Montilla ^b , Egilda Pérez ^a 

^aUniversidad de Carabobo, Facultad de Ingeniería, Centro de Procesamiento de Imágenes, Valencia, Venezuela

^bYttrium-Technology Corp., Panamá, Panamá.

La selección de este artículo fue realizada en el marco de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica “Prof. César Rodolfo Ruiz” Octubre 2020, siendo la evaluación, arbitraje, aceptación y edición a cargo de Revista Ingeniería UC.

Abstract.- This work proposes a convolutional autoencoder based non-supervised feature extractor, to find clusters of electroencephalographic signal (EEG) to supporting to the physicists to diagnose the epilepsy condition. Three autoencoders were designed with input dimensions of 4096×1 , 2048×2 and 768×6 , to analyze the impact of the signal length on latent representation generated by autoencoders. Latent representation was used as input to the clustering algorithms K -means and support vector clustering. Latent representation was mapped onto a two-dimensional space of mean and standard deviation to visualize it, and to apply the clustering algorithms. Results showed a good latent representation of the three autoencoders, with a maximum reconstruction error of 1,47 % for the worst case. Clustering algorithms got visually consistent clusters compared with the ground-truth distribution onto the two-dimension latent space. The best performance was achieved with the K -means algorithm and the best latent representation of the input signal. Resultant clusters were impacted by the length of the input segment, where K -means with an input length of 4096 samples had the best performance.

Keywords: Epileptic EEG signals; autoencoder; K -means; SVC.

Agrupamiento de Señales EEG con Rasgos Aprendidos Usando Autoencoder Profundo

Resumen.- Este trabajo propone un algoritmo basado en autoencoders convolucionales como extractor de rasgos no supervisado, para hallar grupos o clusters de señales electroencefalográficas (EEG), como apoyo para el especialista médico para facilitar el diagnóstico de la condición de epilepsia. Se diseñaron tres autoencoders con señales de entrada de 4096×1 , 2048×2 y 768×6 , para analizar el efecto de la longitud de la señal sobre la representación latente generada por los autoencoders. La representación latente se utilizó como entrada a los algoritmos de agrupamiento K -means y basado en vectores de soporte. La representación latente se llevó a un espacio bidimensional donde se obtuvo la media y la desviación estándar para visualizarla, y operar sobre ellas los algoritmos de agrupamiento. Los resultados demostraron una buena representación latente de los tres autoencoders, con un error máximo de reconstrucción de las señales de entrada de 1,47 % para el peor caso. Los algoritmos de agrupamiento lograron obtener unos grupos visualmente consistentes con la distribución de los puntos de referencia en el espacio bidimensional latente. La mejor medida de desempeño se logró con el algoritmo K -means con la mejor representación latente de las señales de entrada. Los grupos resultantes fueron influenciados por la longitud del segmento de entrada, donde el algoritmo K -means con una longitud de entrada de 4096 muestras tuvo la mejor medida de desempeño.

Palabras clave: Señales EEG epilépticas; autoencoders; K -means; agrupamiento basado en vectores de soporte.

Recibido: 22 de octubre, 2020.

Aceptado: 27 de noviembre, 2020.

1. Introducción

Uno de los desórdenes neurológicos más importante que afecta a la actividad cerebral

es la epilepsia. La epilepsia es una condición que padecen 50 millones de personas alrededor del mundo [1], que produce crisis convulsivas que afecta la calidad de vida del paciente [2]. Para la detección de la condición de epilepsia el especialista tiene que analizar e interpretar un conjunto de señales electroencefalográficas (EEG) muy extensas, lo que hace que la detección sea un proceso muy tedioso, y además dependiente

* Autor para correspondencia:

Correo-e:svillaza@gmail.com (S. Villazana)

del criterio del especialista que realiza dicho análisis. Es imperativo automatizar la detección de señales epilépticas mediante el análisis de las señales EEG para coadyuvar al especialista en su tarea de analizar estas señales. Actualmente, existen técnicas para el análisis de las señales EEG, basados en algoritmos de inteligencia artificial, destacando en los últimos años las redes neuronales profundas (entre ellas las redes convolucionales y las redes recurrentes). Las diferentes clases de señales epilépticas poseen rasgos subyacentes particulares que caracterizan a cada clase en consideración, pero los mismos no son visibles en las señales EEG temporales. Estos rasgos subyacentes permiten discriminar entre diferentes señales EEG normales y patológicas, y por tanto, asignarles una etiqueta; este proceso de análisis se denomina *clustering* (agrupamiento) [3]. El *clustering* es una técnica de análisis de datos que permite asignarles etiqueta a los datos, lo que se puede entender como una técnica de clasificación no supervisada [3], ya que no se cuenta con etiquetas previas con las cuales compararlas.

Una característica a destacar de las redes neuronales profundas es la capacidad de extraer la información subyacente (rasgos) de los datos de entrada, por medio de sus numerosas capas intermedias, lo que elimina la necesidad utilizar enfoques para extraer rasgos a mano, que utilizan una vasta cantidad de técnicas lineales y no lineales, basadas en el tiempo o en la frecuencia o combinación de ellas. La efectividad de los rasgos extraídos manualmente de la señal para modelar exitosamente los rasgos subyacentes, y detectar los grupos (*clusters*) “naturales” de los datos es dependiente de la naturaleza de los datos bajo análisis, y por tanto, exige un trabajo arduo del investigador para la selección de los mismos durante la fase de experimentación.

Los algoritmos de aprendizaje profundo para extraer los rasgos subyacentes de las señales EEG, cuando no se cuenta con una discriminación a priori (etiqueta o clase) de las señales, son los *autoencoders* (AE) [4]. Los AE son redes neuronales convolucionales profundas que, fundamentalmente, tratan de emular la función identidad, es decir, la salida es una reproducción de

su entrada [5]. Los AE se componen de dos bloques en cascada denominados codificador (*encoder*) y decodificador (*decoder*). La salida del primero es una representación latente (rasgo subyacente) de la entrada, que es la entrada del decodificador, que finalmente reproduce la entrada [5]. En la medida que el error de reconstrucción (diferencia de aproximación entre la entrada y la salida) sea más pequeño, la representación latente contiene rasgos más relevantes de la entrada [5].

Los métodos para agrupar los datos, en este trabajo, con apoyo en la representación latente (rasgos subyacentes) obtenidos con el AE son el agrupamiento basado en *K*-means [6] y agrupamiento basado en vectores de soporte (*Support Vector Clustering*, SVC de sus siglas en inglés) [7, 8]. En el agrupamiento *K*-means [6] conocido también como agrupación dura (*hard clustering*), los datos se dividen en grupos distintos, donde cada elemento de dato pertenece exactamente a un grupo (*cluster*). El agrupamiento basado en las SVC es un algoritmo que puede detectar formas arbitrarias de grupos de un espacio de datos [7]. EL SVC proyecta los puntos de datos a un espacio de rasgos hiperdimensional usando una función de núcleo (*kernel*) Gaussiana [8].

Este estudio propone un proceso de extracción no supervisada de rasgos señales EEG usando autoencoders convolucionales, para agrupar las señales EEG en grupos (*clusters*) utilizando dos métodos de agrupamiento basado en *K*-means y SVC.

2. Fundamentación teórica

2.1. Autoencoders

Los autoencoders son redes neuronales profundas para obtener una representación de los datos en un nuevo espacio conocido como espacio latente. El autoencoder está compuesto de dos bloques fundamentales: a) El encoder que transforma los datos de entrada, x , a un espacio, de menor dimensión que el espacio original de los datos, mientras que mantiene los rasgos más relevantes $z = f_{\phi}(x)$ [5], donde f y ϕ son la función de proyección del espacio de entrada al espacio latente, y los parámetros de la función de

proyección f del encoder, respectivamente; b) El decoder que reconstruye los datos originales a partir de la representación latente, $r = g_\theta(z) \cong x$ [5], donde g y θ son la función de proyección del espacio latente (z) al espacio de entrada, y los parámetros de la función de proyección g del decoder, respectivamente. El AE sintetiza una composición de funciones $r = g_\theta(f_\phi(x)) \cong x$, donde para obtener una buena representación latente (z) de la entrada (x), se requiere que la reconstrucción r sea lo más aproximado posible a la entrada x . La función objetivo a minimizar, para obtener la mejor reconstrucción de la entrada, es el error medio cuadrático. Dado el conjunto de entrada x_1, x_2, \dots, x_n , se requiere minimizar la función objetivo $\min_{\phi, \theta} \frac{1}{n} \sum_{i=1}^n \|x_i - g_\theta(f_\phi(x_i))\|^2$, donde ϕ y θ son los parámetros del encoder y el decoder, respectivamente [5].

2.2. Agrupamiento basado en las Máquinas de Vectores de Soporte

El agrupamiento basado en las Máquinas de Vectores de Soporte (SVC siglas en inglés de *Support Vector Clustering*) se basa sobre el clasificador de una clase [9, 10] y un enfoque de etiquetamiento de los grupos [8]. El objetivo principal de esta máquina de aprendizaje es hallar al hiperplano más alejado del origen, en el espacio de rasgos, que separa a todos los datos hacia el lado (del hiperplano) opuesto al origen. Esto se logra al minimizar la función objetivo $\min \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{\nu n} \sum_{i=1}^n \xi_i - \rho$, sujeto a las restricciones $\mathbf{w}^T \Phi(x_i) \geq \rho - \xi_i$, $\xi_i \geq 0$, $\rho \geq 0$, donde \mathbf{w} es la normal del hiperplano, ρ es la distancia del hiperplano al origen, y ν controla la influencia de los datos atípicos (*outliers*). Φ es una función no lineal (función kernel) que proyectan los puntos desde el espacio de datos a un espacio de rasgos hiperdimensional. La función Lagrangiano de la función objetivo es $L = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{\nu n} \sum_{i=1}^n \xi_i - \rho - \sum_{i=1}^n \alpha_i (\mathbf{w}^T \Phi(x_i) + \xi_i - \rho) - \sum_{i=1}^n \beta_i \xi_i$, sujeta a $\alpha_i, \beta_i \geq 0$. Tomando las derivadas del Lagrangiano L con respecto a \mathbf{w} , ξ_i , ρ e igualando a cero conduce a $\mathbf{w} = \sum_{i=1}^n \alpha_i \Phi(x_i)$, $\alpha_i + \beta_i = 1/\nu n$,

$\sum_{i=1}^n \alpha_i = 1$. El producto interno entre todos los vectores $\Phi(x_i)$ y $\Phi(x_j)$ en el espacio de rasgos es la matriz kernel, y un elemento de dicha matriz es $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$. La función kernel a utilizar es la función Gaussiana, $K(x_i, x_j) = e^{(-\|x_i - x_j\|^2 / (2\sigma^2))}$, como lo propuso [8], donde el parámetro σ es el ancho de la Gaussiana, algunos autores utilizan la siguiente igualdad $\gamma = 1/2\sigma^2$. Sustituyendo las relaciones obtenidas del Lagrangiano y sustituyéndolas en la función objetivo primaria conduce a la forma dual de esta $\min \frac{1}{2} \sum_{i=1}^n \alpha_i \alpha_j \langle \Phi(x_i), \Phi(x_j) \rangle = \min \frac{1}{2} \sum_{i=1}^n \alpha_i \alpha_j K(x_i, x_j)$. Las condiciones KKT [11] establecen que $\alpha_i (\mathbf{w}^T \Phi(x_i) + \xi_i - \rho) = 0$ y $\beta_i \xi_i = (1/\nu n - \alpha_i) = 0$. Los vectores de soporte (VS) son aquellos datos x_i para los cuales los multiplicadores de Lagrange α_i son distintos de cero con $\xi_i \geq 0$, y los datos x_i ordinarios que no son VS su correspondiente α_i son iguales a cero.

El enfoque de etiquetamiento o de asignación de grupos se basa en una técnica muy rápida propuesta en [8], este enfoque de etiquetamiento de grupo está basado en conos [8]. Debido a que la función kernel es la Gaussiana, se puede demostrar, que los datos en el espacio de entrada se proyectan a una hiperesfera de radio unitario, de allí que los autores definen un cono entre dos vectores de soporte. Si dos o más conos se interceptan los vectores de soporte que definen a cada cono pertenecen al mismo grupo. En este caso, el máximo número de grupos corresponde al número de vectores de soporte. Cada cono se define por el ángulo que forman dos vectores de soporte y el origen. En [8] los autores demostraron que el ángulo entre las imágenes de dos vectores $\Phi(x_i)$ y $\Phi(x_j)$ en el espacio de rasgos, cuando la función kernel es una Gaussiana, está definida por $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle = \cos \omega$, donde ω es el ángulo entre los vectores $\Phi(x_i), \Phi(x_j)$. Se puede demostrar que las distancias en el espacio de datos corresponden a ángulos en el espacio de rasgos y viceversa. Esta afirmación es importante debido a que el proceso de agrupamiento es hecho en el espacio de datos usando la distancia, afectada por un factor arbitrario a criterio del investigador para

ajustar el número de grupos.

2.3. Agrupamiento basado en *K-means*

Este algoritmo divide a los datos en K grupos o particiones, C_1, C_2, \dots, C_K , representados por sus centros. El centro de cada grupo se calcula como la media de todos datos que pertenecen a dicho grupo. Los K centros iniciales se seleccionan aleatoriamente. En cada iteración se reasignan los datos al grupo más cercano utilizando la distancia Euclidiana entre el k -ésimo centro y el dato bajo consideración. Luego los centros de los grupos se recalculan con los nuevos miembros incorporados y/o desincorporados. El k -ésimo centro se calcula como $\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q$ [12], donde N_k es el número de datos que pertenecen al k -ésimo grupo y μ_k es la media del grupo k . El criterio de parada del algoritmo es observar el error de asignación al grupo, si este no se reduce con la reubicación de los centros, o alcanzar un número máximo de iteraciones, lo que suceda primero.

2.4. Medidas de evaluación de desempeño

Para evaluar el desempeño de los modelos de agrupamiento existen dos medidas ampliamente usadas en la literatura. La primera medida es la exactitud basada en el algoritmo Húngaro y la segunda es la información mutua normalizada.

2.4.1. Exactitud basada en el algoritmo Húngaro:

Es una medida directa que refleja la efectividad del algoritmo de agrupamiento. La exactitud viene dada por $ExacHung = \frac{1}{n} \sum_{i=1}^n \delta(c_i, map(\hat{c}_i))$, donde c_i es la etiqueta verdadera de los datos, y \hat{c}_i es la etiqueta asignada por el algoritmo de agrupamiento; $\delta(i, j) = 1$ si $i = j$, $\delta(i, j) = 0$ si $i \neq j$; map es una función de proyección optimizada que proyecta cada cluster hallado a la clase real [13], y n es el número de datos. Esta medida halla la mejor correspondencia entre los grupos hallados y las clases de referencia [13].

2.4.2. Información mutua normalizada:

La información mutua normalizada (*IMN*) [14] se calcula como $IMN = 2 \times I(C, \hat{C}) / (H(C) +$

$H(\hat{C}))$, donde C son las clases de referencia, \hat{C} son las etiquetas asignadas por el algoritmo de agrupamiento, I es la información mutua, y H es la entropía.

3. Metodología

3.1. Base de datos

La base de datos utilizada en este trabajo corresponde a las señales EEG de la Universidad de Bonn descrita en [15], está conformada por cinco conjuntos A, B, C, D y E, cada uno con 100 señales electroencefalográficas (EEG) monocanal de 23,6 segundos de duración cada una. Estas señales EEG están libres de artefactos debido a la actividad muscular o movimientos de los ojos. Los conjuntos A y B consisten en segmentos tomados de los registros EEG superficiales obtenidos de cinco voluntarios sanos usando un esquema de colocación de los electrodos estandarizada, conocida como sistema 10-20. Los conjuntos A y B corresponden a voluntarios despiertos, relajados y con los ojos abiertos (A) y los ojos cerrados (B). Los conjuntos C, D y E, de pacientes diagnosticados con epilepsia, corresponden a EEG profundos o intracraneales. Las señales en el conjunto C fueron obtenidas de la formación hipocampal del cerebro. El conjunto D se obtuvo dentro de la zona epileptogénica. Los conjuntos C y D contienen registros de la actividad cerebral medida durante los intervalos sin crisis epilépticas (interictal). El conjunto E contiene registros durante la actividad convulsiva, o periodo ictal.

Todos estos segmentos EEG se registraron con un amplificador de 128 canales, un convertidor analógico-digital de 12 bits a una frecuencia de muestreo de 173,61 Hz, y se les aplicó un filtro pasabanda con ajustes de 0,53 Hz y 40 Hz [15]. El número de registros en total es 500, con 4097 muestras cada uno (23,6 segundos de registro por EEG). La Figura 1 presenta una muestra de cada una de las cinco señales por cada conjunto, la unidad de los ejes verticales está en microvoltios (μV), y la unidad del eje horizontal es muestras (adimensional).

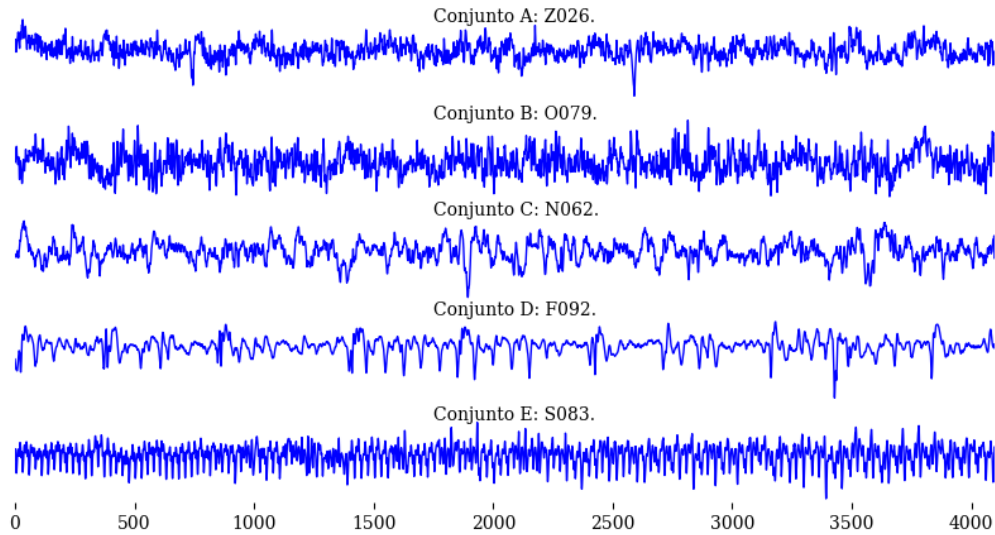


Figura 1: Muestras de señales de los conjuntos A, B, C, D y E

3.2. Entrenamiento de los autoencoders

La Figura 2 muestra un diagrama de bloques del proceso de extracción de rasgos (z) y de clustering. El proceso de entrenamiento está dividido en dos fases: a) Entrenamiento del autoencoder para obtener la representación latente de la señal (z) y, b) Entrenamiento del algoritmo de agrupamiento para determinar los grupos o clusters.

Para agrupar los datos se realizaron tres esquemas para la extracción de los rasgos con el autoencoder y analizar su efecto sobre los grupos generados. El primer esquema fue utilizar las señales unidimensionales completas (4096 muestras) como entrada al autoencoder, la segunda estrategia fue dividir cada señal en dos segmentos de 2048 muestras (descartando la última muestra) y formando con ellos una nueva señal de 2048×2 , la tercera y última estrategia fue dividir cada señal en segmentos de 768 muestras. En esta última estrategia el número de dimensión de la nueva señal sería de 768×5 , quedando descartadas 257 muestras, por lo que se decidió incluir un segmento adicional, tomando las últimas 511 muestras del quinto segmento, por tanto la dimensión resultante fue de 768×6 .

La arquitectura del autoencoder para las señales de dimensión 4096×1 se muestra en la Tabla 1.

Todas las capas de convolución (Conv1D, Conv1DTranspose) se implementaron con un filtro de longitud 7 y el número de filtros (Profundidad) depende de la ubicación de la capa en la arquitectura de la red. La capa Conv1DTranspose (llamada también como capa de des-convolución), realiza la operación de convolución, pero, aumenta a su salida la dimensión de la señal de entrada, a diferencia de la capa Conv1D que reduce a su salida la dimensión de la señal de entrada. Todas las capas de Dropout eliminan aleatoriamente el 20% de las neuronas de la capa precedente, estas se incluyeron para agregar un factor de regularización para mejorar la capacidad de generalización de la red. Todas las capas de convolución (Conv1D, Conv1DTranspose), excepto la última, le sigue una capa con la función de activación LeakyReLU con su parámetro por defecto ($\alpha = 0,3$). El encoder lo conforman las capas de convolución, su entrada es la señal a codificar (4096×1) y su salida es de dimensión 1024×16 , como se observa, se redujo su longitud de 4096 a 1024, pero se incrementó la dimensión (de 1 a 16). El decoder lo conforman las capas de des-convolución, con una entrada de 1024×16 y salida de 4096×1 . El número de parámetros del autoencoder con señal de entrada de 4096×1 es 9505.

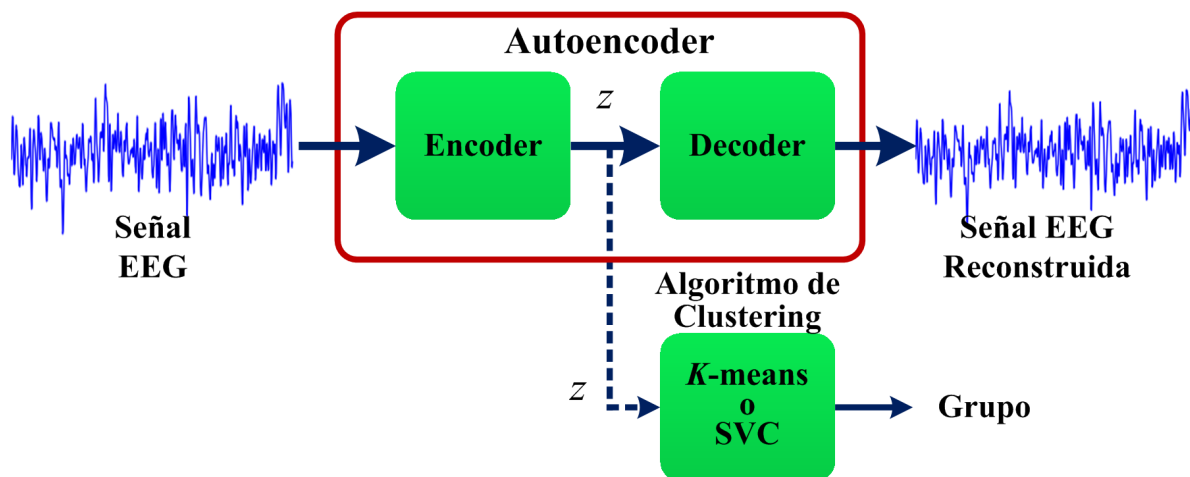


Figura 2: Diagrama de bloques del proceso de agrupamiento

Tabla 1: Arquitectura del autoencoder

	Tipo de capa	Filtro			Salida
		Dimensión	Profundidad	Paso	Dimensión
Encoder	Entrada	4096 × 1	-	-	-
	Conv1D	7 × 1	32	2	2048 × 32
	Dropout	-	-	-	2048 × 32
	Conv1D	7 × 1	16	2	1024 × 16
Decoder	Conv1DTranspose	7 × 1	16	2	2048 × 16
	Dropout	-	-	-	2048 × 16
	Conv1DTranspose	7 × 1	32	2	4096 × 32
	Conv1DTranspose	7 × 1	1	2	4096 × 1

La arquitectura de los autoencoders se mantiene para las otras configuraciones de entrada, variando solamente las dimensiones de entrada, la salida del autoencoder y las variables latentes (salida del encoder). El número de parámetros de los autoencoders para las entradas de 2048 × 2 y 768 × 6 es de 9954 y 11750, respectivamente, y las dimensiones de la salida del encoder son 512 × 16 y 192 × 16, para las señales de entrada de 2048 × 2 y 768 × 6, respectivamente.

Los ajustes para entrenar fueron: El optimizador elegido fue *Adam* con sus parámetros por defecto, la función objetivo fue *mse* (error medio cuadrático). El número de épocas fue 100, y el tamaño del lote (*batch size*) fue de 81.

La división de los datos, para todos los ensayos, fue de 90 % (450 señales) para la fase de entrenamiento y 10 % (50 señales) para prueba. Del conjunto de entrenamiento se tomó el 90 % (405 señales) para entrenar y 10 % (45 señales) para validación en cada época de entrenamiento, y supervisar si se está sobreentrenando (*overfitting*)

o subentrenando (*underfitting*). Todas las señales de entrenamiento fueron escaladas por columnas entre 0 y 1, ambos valores inclusive. Las señales de prueba se escalaron entre 0 y 1 utilizando como referencia los valores máximos y mínimos de cada columna o dimensión de las señales de entrenamiento.

La Tabla 2 muestra los errores de entrenamiento y prueba (o generalización) de los autoencoders. Se observa un aumento de los errores en la medida que el número de muestras de la señal de entrada disminuye, ya que el número de parámetros se incrementa, lo que significa que la complejidad aumenta, y a mayor complejidad mayor error.

Tabla 2: Errores de entrenamiento y validación

Autoencoder		Errores (%)	
Entrada	Nro. de Parámetros	Entrenamiento	Prueba
4096 × 1	9505	0,033	0,220
2048 × 2	9954	0,117	0,604
768 × 6	11750	0,314	1,468

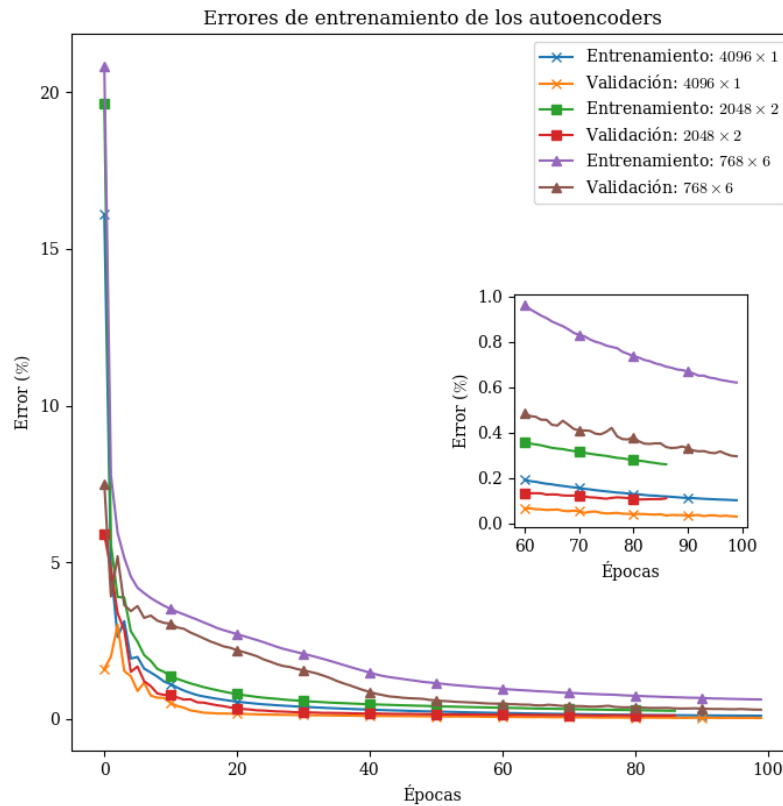


Figura 3: Errores de entrenamiento y validación durante el entrenamiento

La Figura 3 muestra los errores de entrenamiento y de validación en función de las épocas de entrenamiento de los tres autoencoders.

Se observa que los errores de validación siempre son menores que los errores de entrenamiento durante todo el entrenamiento, lo que significa que no hubo sobreentrenamiento de los autoencoders.

Las Figuras 4, 5 y 6 muestran las señales de entrada y salida (reconstrucción) de los encoders con entradas 4096×1 , 2048×2 y 768×6 , respectivamente.

Las Figuras 4, 5 y 6 muestran una excelente reconstrucción de las señales, lo que es un indicativo de que la representación latente (salida del encoder) de las entradas es bastante representativa de cada señal.

3.3. Entrenamiento de los algoritmos de clustering

El procedimiento para entrenar los algoritmos de agrupamiento fue el siguiente: Se tomaron las representaciones latentes de las señales de prueba y se redimensionaron como vectores unidimensionales, ello se hizo porque las dimensiones de la representación latente fueron de 1024×16 , 512×16 y 192×16 para los autoencoders con entradas 4096×1 , 2048×2 y 768×6 , respectivamente. El número de vectores unidimensionales fue 50 con longitudes 16384, 8192 y 3072 para los autoencoders con entradas 4096×1 , 2048×2 y 768×6 , respectivamente. Luego por cada vector se obtuvo su media y su desviación estándar para obtener 50 vectores de dos dimensiones, $z = (z_0, z_1)$, donde z_0 es la media del vector unidimensional latente, y z_1 es la desviación estándar del vector unidimensional latente, para cada autoencoder. Las Figuras 7, 8 y 9 muestran la

Tabla 3: Parámetros del algoritmo SVC

Encoder	Parámetros del clasificador de una clase		Parámetro de etiquetamiento	Nro. de vectores de soporte	Número de grupos
	ν	σ			
Entrada	ν	σ	Factor		
4096×1	0,38	0,0707	0,32	19	4
2048×2	0,30	0,0707	0,50	16	4
768×6	0,25	0,0707	0,45	15	5

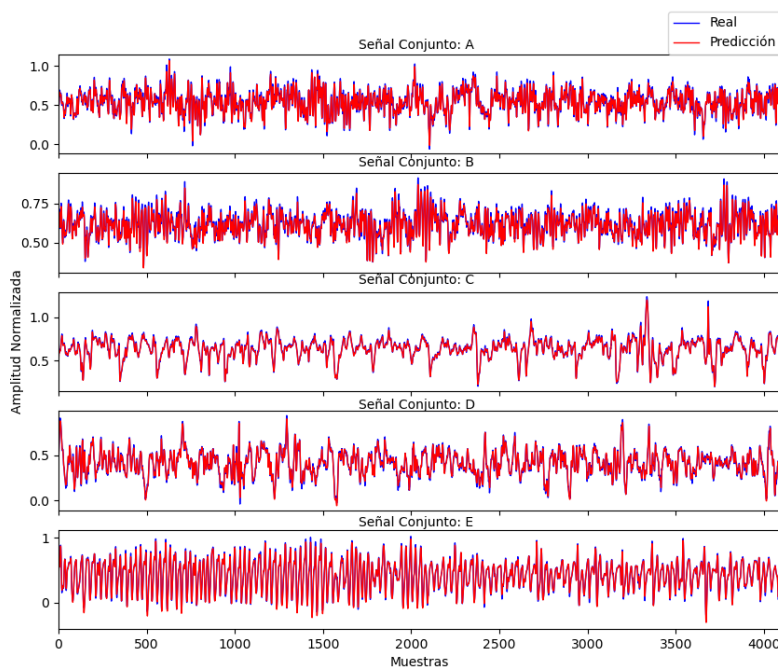


Figura 4: Señales de entrada de prueba y sus reconstrucciones. Encoder 4096×1.

distribución de los nuevos vectores latentes, donde también se indica el conjunto de origen de las señales EEG. Las distribuciones mostradas en las Figuras 7, 8 y 9 son las referencias (*ground-truth*) para evaluar el desempeño de los algoritmos de agrupamiento para todos los autoencoders.

Si se detallan cuidadosamente los rangos de los ejes de las Figuras 7, 8 y 9, se observa que las medias (eje horizontal) y las desviaciones estándar (eje vertical) de las variables latentes, independientemente de las clases, disminuyen con la longitud de la señal de entrada del encoder. También se ve que la posición relativa de los puntos, representando cada señal del conjunto de prueba, se mantienen aproximadamente invariable, independientemente

del encoder. También es notable la superposición de los puntos pertenecientes a las señales de los conjuntos A y B (ambos de EEG normales). En general, el conjunto D es el que está bien diferenciado del resto de los conjuntos, algunos puntos del conjunto C se superponen con puntos del conjunto A, y el conjunto E (señales ictales), posee algunas señales que se solapan con los puntos de los conjuntos A y B, y otras señales que definitivamente se alejan muy claramente del resto.

3.3.1. Agrupamiento con el método K-means

Se fijó el número de clusters (grupos) igual al número de conjuntos de la base de datos, $K = 5$. Los centroides de los grupos se

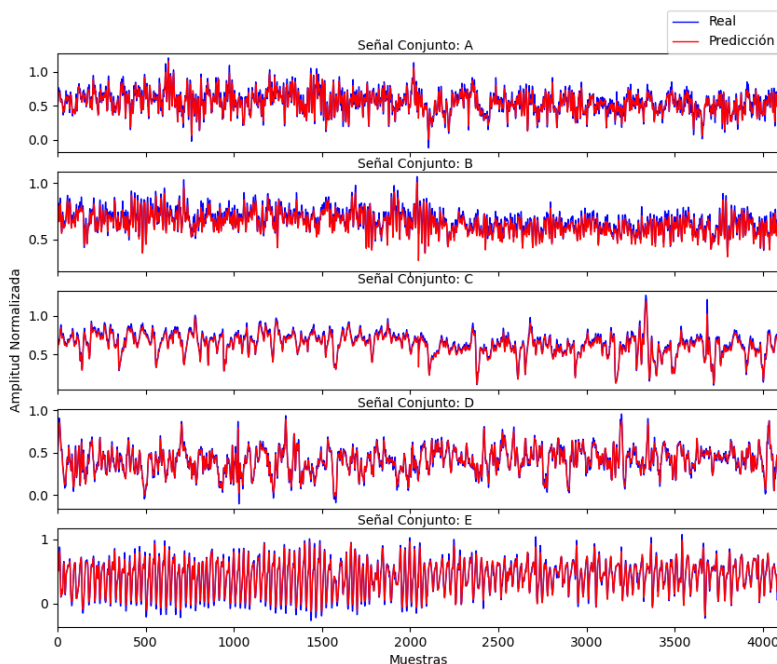


Figura 5: Señales de entrada de prueba y sus reconstrucciones. Encoder 2048×2.

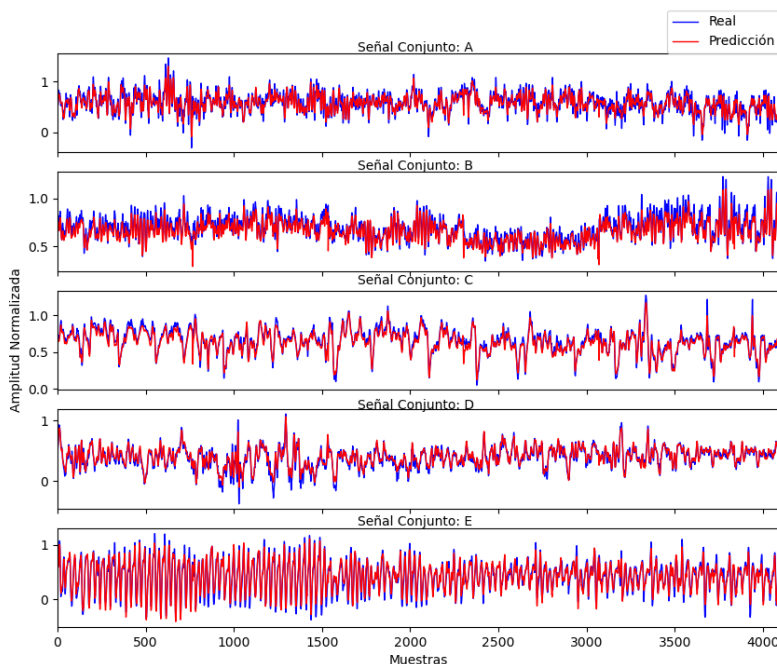


Figura 6: Señales de entrada de prueba y sus reconstrucciones. Encoder 768×6.

inicializaron aleatoriamente. El resultado para los tres autoencoders se muestran en la Figuras 10, 11

y 12. Se observa que la posición de los clusters con la misma referencia numérica no necesariamente

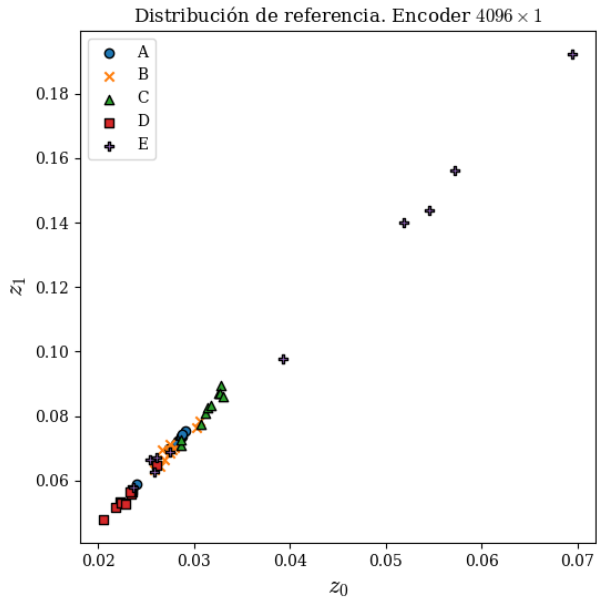


Figura 7: Distribución de las señales de los conjuntos A, B, C, D y E para el Encoder 4096x1.

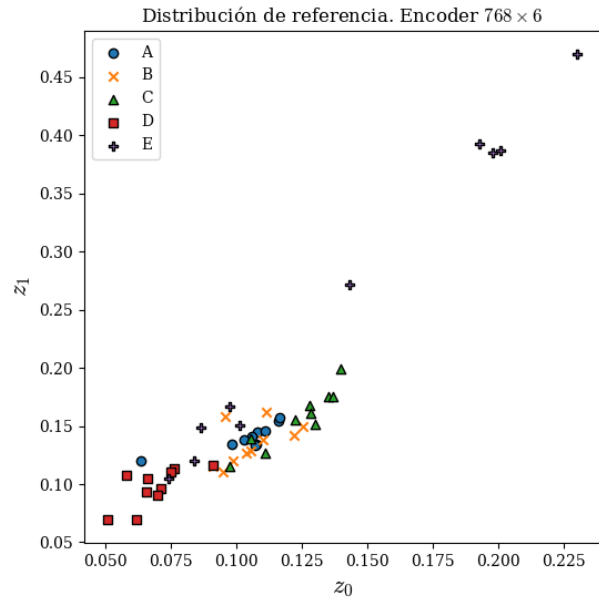


Figura 9: Distribución de las señales de los conjuntos A, B, C, D y E para el Encoder 768x6.

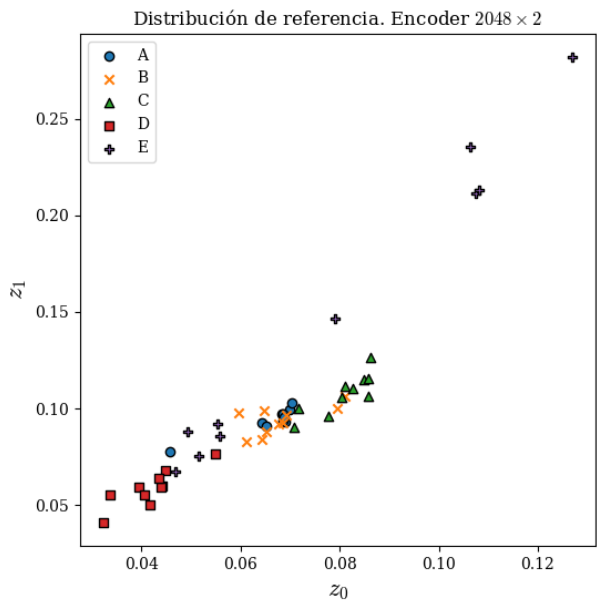


Figura 8: Distribución de las señales de los conjuntos A, B, C, D y E para el Encoder 2048x2.

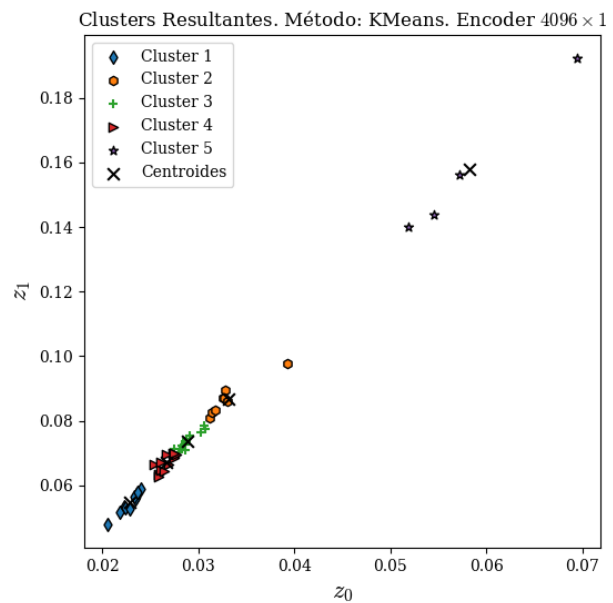


Figura 10: Grupos obtenidos con el método K-means para el Encoder 4096x1.

están ubicados en la misma región relativa del espacio $z_0 - z_1$.

3.3.2. Agrupamiento con el método SVC

Los ajustes para el entrenamiento de las tres clasificadores de una sola clase y del etiquetamiento basado en conos se muestran

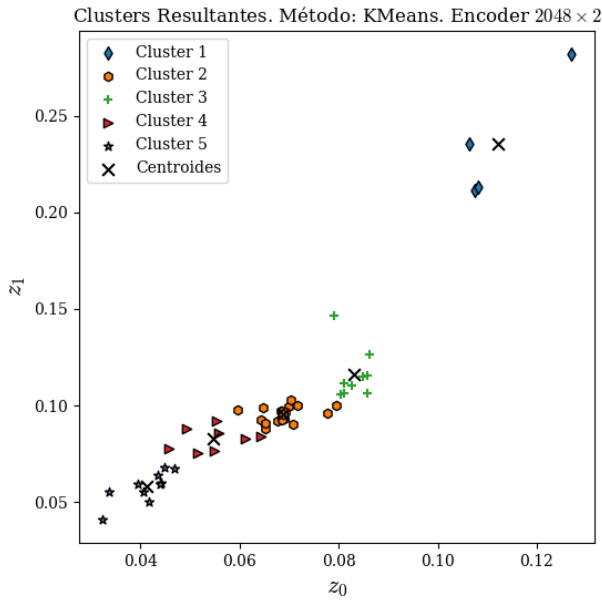


Figura 11: Grupos obtenidos con el método K -means para el Encoder 2048×2 .

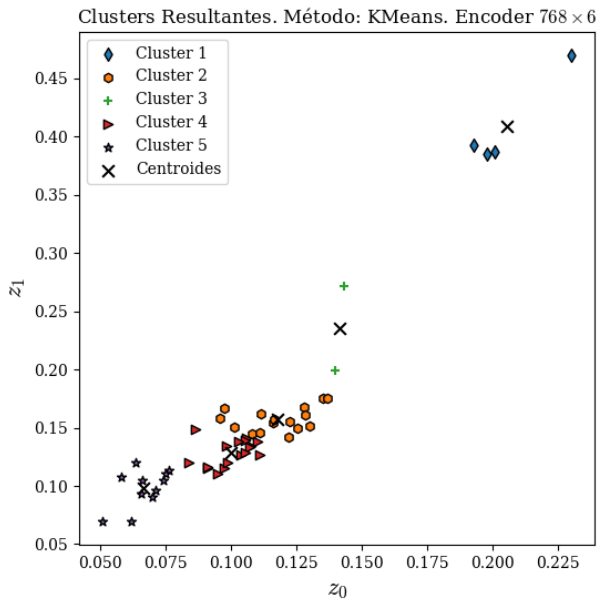


Figura 12: Grupos obtenidos con el método K -means para el Encoder 768×6 .

en la Tabla 3. Los valores se ajustaron para tener un máximo de cinco grupos. El factor de etiquetamiento lo que hace es modificar el rango de alcance (distancia) de influencia de los clusters definidos por los vectores de soporte. La Tabla 3

también muestra el número de vectores de soporte y el número de grupos obtenidos con el método SVC para cada una de las representaciones latentes de los autoencoders. Las Figuras 13, 14 y 15 muestran los clusters definidos para cada autoencoder.

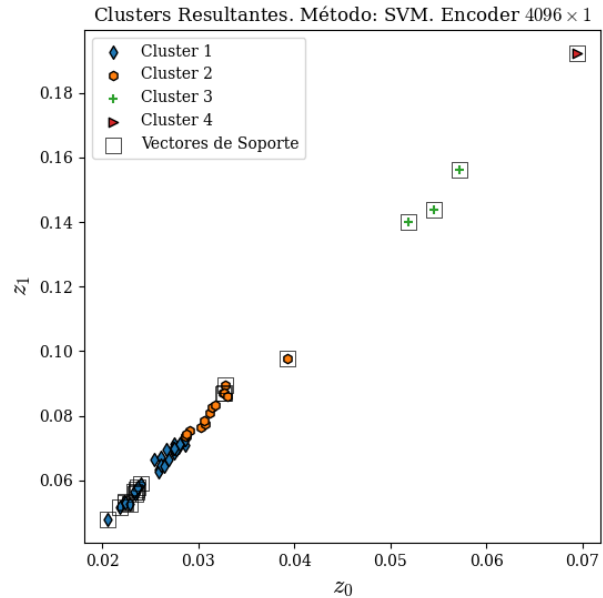


Figura 13: Grupos obtenidos con el método SVC para el Encoder 4096×1 .

La Tabla 4 muestra el desempeño de los métodos de agrupamiento basados en K -means y SVC, aplicando la formulación de la exactitud basada en el algoritmo húngaro (*ExacHung*) y la información mutua normalizada (*IMN*). La medida *ExacHung* no está disponible para el método SVC, para los autoencoders 4096×1 y 2048×2 , porque el número de grupos hallados fue 4.

Tabla 4: Desempeño de los algoritmos K -means y SVC

Encoder	Medidas de desempeño			
	<i>ExacHung</i>		<i>IMN</i>	
	K -means	SVC	K -means	SVC
4096×1	0,66	-	0,52	0,28
2048×2	0,62	-	0,54	0,42
768×6	0,50	0,44	0,36	0,36

Se puede observar en la Tabla 4 que el mejor agrupamiento lo obtuvo el método K -means con

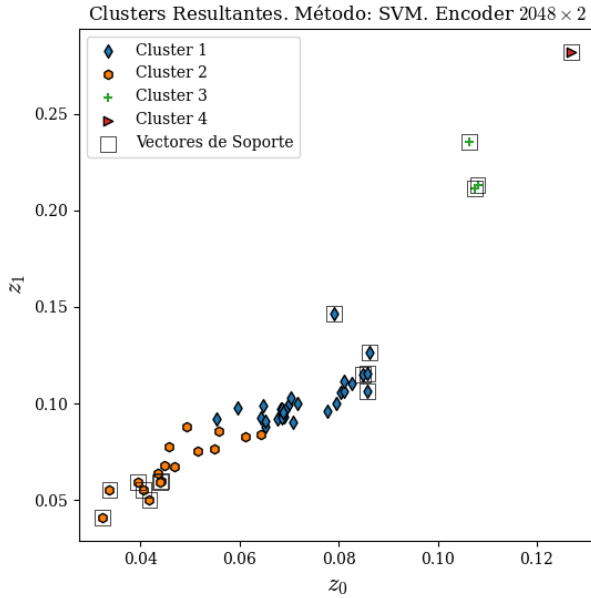


Figura 14: Grupos obtenidos con el método SVC para el Encoder 2048x2.

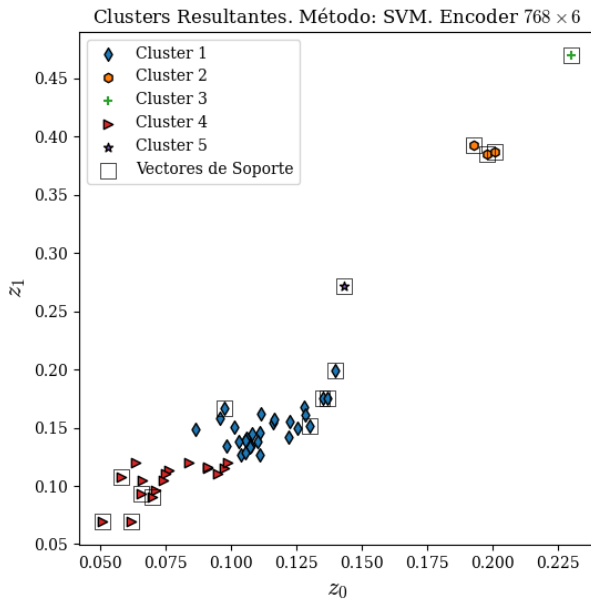


Figura 15: Grupos obtenidos con el método SVC para el Encoder 768x6.

las entradas latentes del autoencoder 4096x1, lo que confirma que la mejor representación latente fue aquella con el menor error de reconstrucción (0,22 %). No obstante, el agrupamiento obtenido con el autoencoder 2048x2 también tuvo excelente

desempeño.

4. Análisis y discusión

Se puede verificar, de los resultados obtenidos, que el error de reconstrucción de los autoencoders disminuye cuando se entrena con la señal completa, en comparación a cuando se entrena con la señal segmentada, esto puede estar influenciado por el proceso de normalización de las señales, ya que cada columna se normaliza como si fuera una señal independiente (cierto solo para las señales sin segmentar), lo que no es cierto para las señales segmentadas. A pesar de lo anterior, el error del peor caso (señal segmentada en 6 segmentos) es 1,47 % evaluado sobre el conjunto de pruebas, lo que indica una buena capacidad de reconstrucción de las señales por los tres autoencoders.

Para poder tener una representación geométrica visualizable de las variables latentes (salida del autoencoder) se hallaron las medias (z_0) y las desviaciones estándar (z_1) de dichas variables latentes, para obtener un vector bidimensional y poder visualizar su distribución sobre el plano $z_0 - z_1$. Se observa, casi una invariabilidad en la posición relativa de los puntos o señales pertenecientes a los conjuntos A, B, C, D y E, lo que puede deberse a la buena representatividad de la variables latentes a las señales originales, obtenidas por cada autoencoder. El efecto observable es una disminución de la media (eje z_0) y la desviación estándar (eje z_1) con el aumento de la longitud de la señal y/o la disminución de su dimensión, lo cual se comprueba por simple inspección al observar los rangos de los ejes z_0 y z_1 de las Figuras 7, 8 y 9.

El proceso de agrupamiento obtenido por cada método (K-means y SVC) mostraron resultados con buen desempeño, lo que se confirma visualmente. Pero de acuerdo a las medidas de desempeño, los mejores resultados de agrupamiento se lograron con el método K-means y los autoencoders 4096x1 y 2048x2. Se puede ver que el efecto de la longitud del segmento de señal influencia la detección de los grupos, ya que el aumento o disminución de la distancia relativa entre los puntos en el plano $z_0 - z_1$, facilita o dificulta el proceso de agrupamiento. El método SVC pudo determinar 5 grupos solo para

el caso de la señal de longitud de 768 muestras, y solo 4 grupos para las señales de longitudes de 4096 y 2048 muestras, pero con grupos de solo de 1 miembro, correspondiente a los puntos muy alejados del resto.

5. Conclusiones

Este trabajo presentó la propuesta de un esquema de agrupamiento de señales EEG basados en variables latentes obtenidas de un autoencoder convolucional. Se pudo verificar que, independientemente de la longitud de la señal de entrada al autoencoder, el error de reconstrucción obtenido fue bastante bajo, menos del 1,5 %, aún para el caso más exigente (señal de 768×6).

Se verificó la fortaleza de los autoencoders como extractor de rasgos no supervisado, al obtener una buena representación latente de las señales de entrada.

Los algoritmos de agrupamientos obtuvieron unos grupos con miembros (puntos) consistentemente homogéneos, y con buenas medidas de desempeño para el método K -means. El mejor desempeño de agrupamiento se obtuvo con el método K -means con la representación latente de las señales obtenidas con el autoencoder 4096×1 .

6. Referencias

- [1] A. Neligan and J. W. Sander, *Epidemiology of Seizures and Epilepsy*. John Wiley & Sons, Ltd, 2014, ch. 4, pp. 28–32.
- [2] D. Shorvon, Simon, *Handbook of epilepsy treatment. Forms, causes and therapy in children and adults*, 2nd ed. Blackwell Publishing, 2005.
- [3] S. Miyamoto, H. Ichihashi, and K. Honda, *Algorithms for Fuzzy Clustering: Methods in c-Means Clustering with Applications*, 1st ed. Springer Publishing Company, 2008.
- [4] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [5] E. Min, X. Guo, Q. Liu, G. Zhang, J. Cui, and J. Long, “A survey of clustering with deep learning: From the perspective of network architecture,” *IEEE Access*, vol. 6, pp. 39 501–39 514, August 2018.
- [6] J. B. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, L. M. L. Cam and J. Neyman, Eds., vol. 1. University of California Press, 1967, pp. 281–297.
- [7] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, “Support vector clustering,” *Journal of Machine Learning Research*, vol. 2, pp. 125–137, march 2001.
- [8] S.-H. Lee and K. M. Daniels, “Gaussian kernel widths selection and fast cluster labeling for support vector clustering,” University of Massachusetts Lowell, Tech. Rep., 2005.
- [9] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the support of a high-dimensional distribution,” *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
- [10] B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, J. C. Platt *et al.*, “Support vector method for novelty detection.” in *NIPS*, vol. 12. MIT Press, 1999, pp. 582–588.
- [11] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. Springer, 2000, vol. 1.
- [12] L. Rokach, “A survey of clustering algorithms,” in *Data Mining and Knowledge Discovery Handbook*, 2nd ed., O. Maimon and L. Rokach, Eds. Springer, 2016, ch. 14, pp. 269–297.
- [13] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval research logistics quarterly*, vol. 2, no. 1–2, pp. 83–97, 1955.
- [14] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, “Normalized mutual information feature selection,” *IEEE Transactions on Neural Networks*, vol. 20, no. 2, pp. 189–201, February 2009.
- [15] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. Elger, “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,” *Physical Review E*, vol. 64, no. 061907, pp. 061 907–1–061 907–8, 2001.

Model identification and comparative evaluation of PID controller tuning methods for a catalytic regeneration process in Scilab-XCOS

Gerardo Lanza ^{*,a} , Francisco Arteaga-Bravo ^b 

^aInnovent, C.A. Automatización y Controles Industriales, Valencia, Venezuela.

^bCentro de Investigación y Tecnología en Automatización, Electrónica y Control (CITAEC), Departamento de Sistemas y Automática, Escuela de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de Carabobo, Valencia, Venezuela.

La selección de este artículo fue realizada en el marco de la Jornada de Investigación de la Escuela de Ingeniería Eléctrica "Prof. César Rodolfo Ruiz" Octubre 2020, siendo la evaluación, arbitraje, aceptación y edición a cargo de Revista Ingeniería UC.

Abstract.- In this paper, model identification and feedback control are presented for a catalytic regeneration process, having as controlled variable the temperature of the catalyst layer. To represent the system, an approximate first-order model, obtained by identification methods is used. Two methods are used for PID controller tuning, a direct synthesis method and an organized approximate trial and error method. The first one requires the precise determination of the dominant characteristics of the process based on identification techniques, while the second one is an improvement-extension of the original Ziegler-Nichols method, taking as an essential condition that to have a well-tuned loop there must be a predictable relationship between the period of oscillation and the integral time. These procedures are applicable to control loops with self-regulated processes using proportional integral controller. A comparative analysis of these methods is developed and the simulations are performed in Scilab-XCOS. Conclusions are established on the method that is more appropriate for PID Controller tuning based on the simulations performed.

Keywords: model identification; PID tuning; feedback control; catalytic regeneration; Scilab-XCOS.

Identificación de modelo y evaluación comparativa de métodos de sintonización de controlador PID para proceso de regeneración catalítica en Scilab-XCOS

Resumen.- En este trabajo se presenta la identificación de modelo y el control por realimentación de un proceso de regeneración catalítica, teniendo como variable controlada la temperatura de la capa de catalizador. Para representar el sistema se utiliza un modelo aproximado de primer orden, obtenido por métodos de identificación. Para la sintonización del controlador PID se utilizan dos métodos, un método de síntesis directa y un método organizado aproximado de ensayo y error. El primero requiere la determinación precisa de las características dominantes del proceso en base a técnicas de identificación, mientras que el segundo es una mejora-extensión del método original de Ziegler-Nichols, tomando como condición esencial que para tener un lazo bien sintonizado debe existir una relación predecible entre el periodo de oscilación y el tiempo integral. Estos procedimientos son aplicables a lazos de control con procesos autorregulados usando controlador proporcional integral. Se desarrolla un análisis comparativo de estos métodos y las simulaciones se efectúan en Scilab-XCOS. Se establecen conclusiones sobre el método que resulta más apropiado para la sintonización del controlador PID en base a las simulaciones realizadas.

Palabras clave: identificación de modelo; sintonización de PID; control por realimentación; regeneración catalítica; Scilab-XCOS.

Recibido: 22 de octubre, 2020.

Aceptado: 12 de diciembre, 2020.

1. Introducción

En esta investigación se aborda el problema de la identificación de modelo y una evaluación comparativa de dos métodos de sintonización de controlador PID para el proceso de regeneración

* Autor para correspondencia:

Correo-e: glanza@innovent.com.ve (G. Lanza)

catalítica en el ambiente del software libre Scilab-XCOS. Se aplica el método de control PID feedback simple para controlar la temperatura T_C de la capa del catalizador del regenerador. En este proceso el objetivo que se persigue es la regeneración del catalizador del reactor químico. Una descripción de este proceso se puede encontrar en Smith y Corripio [1]. Este proceso se representa por un modelo aproximado de primer orden con retardo, obtenido a partir de técnicas de identificación. La importancia de este modelo radica en que las características del proceso son utilizadas por el enfoque de sintonización denominado Síntesis Directa, en el cual el resultado del entonamiento o sintonización depende del modelo del proceso, razón por la que también se le llama diseño basado en modelo [2, 3]. El segundo enfoque o método de sintonización a ser aplicado en este trabajo se denomina método organizado aproximado de ensayo y error, propuesto por Harold Wade en 2005. Este método puede ser visualizado como una estrategia organizada de ensayo y error para obtener el entonamiento del controlador PID con el número mínimo de cambios de parámetros y de perturbaciones al proceso que se está controlando [4].

Inicialmente, se debe realizar la identificación del modelo del proceso, usando técnicas de identificación experimental mediante señales especiales. Esta visión o enfoque, es designado por algunos autores como identificación clásica o también identificación empírica de modelo. La señal de prueba empleada con más frecuencia es la señal escalón, y es la señal utilizada en este estudio. Esta señal de prueba tipo escalón se puede aproximar en la práctica por la apertura o cierre súbito de una válvula que tiene incidencia en el proceso. En la aplicación de identificación experimental, el modelo se determina efectuando pequeños cambios de la variable de entrada alrededor de un punto de operación. La respuesta dinámica que resulta se usa entonces para determinar el modelo. Este procedimiento resulta en una linealización experimental del proceso, válida para una región alrededor de las condiciones nominales de operación. La información sobre el comportamiento dinámico del proceso se obtiene

registrando los valores de la entrada aplicada y la correspondiente respuesta del proceso. Existen diversos procedimientos tanto de lazo abierto (controlador en modo “manual”) como de lazo cerrado (controlador en modo “automático”) para obtener un modelo aproximado del proceso. Los métodos de lazo abierto se denominan métodos basados en la curva de reacción. En este trabajo se aplica el método denominado “Bump Test” o Prueba Escalón para aproximar el proceso mediante un modelo de primer orden que es el utilizado en la mayoría de los métodos de sintonización de controladores PID. Este modelo del proceso es entonces empleado en el método de sintonización de Síntesis Directa [5, 6, 7, 8, 9].

En el campo del control automático de procesos el objetivo a lograr es la optimización del proceso, y esto requiere colocar al proceso como el punto de partida para el control. El proceso no puede ser despreciado o subestimado, ya que eso puede implicar que nunca se vaya a lograr la optimización del mismo. Los enlaces o conexiones en la secuencia de pasos para alcanzar la optimización pueden involucrar las operaciones de calibración de transmisores, filtrado de señales, identificación de proceso y calibración de sintonización [2]. El procedimiento general de entonamiento de un controlador PID envuelve la etapa inicial de identificación donde se obtiene información sobre la dinámica aproximada del proceso (ganancia, tiempo muerto y constante de tiempo), y luego la aplicación de las reglas de sintonización para determinar los parámetros del controlador [2, 7]. Los parámetros del controlador PID deben ser ajustados tomando muy en cuenta la dinámica del proceso. Si no se considera realmente el proceso, los parámetros obtenidos para el PID serían deficientes y resultarían en respuestas muy lentas o inestables [10].

El objetivo principal de esta investigación se enfoca en la realización de un análisis y una evaluación comparativa entre dos métodos muy significativos dentro del ámbito de la sintonización de controladores PID. El primer método denominado Síntesis Directa, presentado por Kevin Starr y aplicado por la empresa ABB [2], está basado en el modelo aproximado de

primer orden obtenido mediante la aplicación de una prueba escalón (“Bump Test”) con el controlador en modo “manual”. El segundo método se denomina método aproximado organizado de prueba y error, presentado por Harold Wade en una publicación de la ISA (International Society of Automation) en el año 2005 y está basado en la premisa de que para obtener un lazo bien sintonizado se debe tener una relación predecible entre el periodo de oscilación y el tiempo integral. En este segundo método de Wade se efectúan cambios de Setpoint en el proceso y se registran los cambios estando el sistema en lazo cerrado [4].

2. Identificación de modelo del proceso

La técnica de identificación de proceso llamada “Bump Test” o Prueba de Señal se va a utilizar para encontrar el modelo del proceso de regeneración catalítica. Se inscribe dentro de la clasificación de métodos de identificación experimental o identificación mediante señales especiales, que algunos autores también denominan identificación clásica. En general en el campo de la identificación de procesos se puede establecer la siguiente clasificación: identificación analítica o de principios básicos, identificación experimental mediante señales especiales y también identificación paramétrica [5, 10].

2.1. Clasificación de los métodos de identificación de proceso

En la identificación analítica se considera el desarrollo de un modelo partiendo de los principios físico-químicos fundamentales del proceso en estudio, se plantean ecuaciones diferenciales en base a balances de masa y energía, y normalmente se obtienen modelos complejos y no lineales donde se requieren métodos para simplificar y linealizar tales modelos, siendo la principal desventaja de esta estrategia de identificación la necesidad de conocer muy bien la tecnología del proceso lo cual no siempre es factible. Debido a la complejidad de los sistemas industriales los métodos analíticos de identificación son generalmente difíciles de aplicar [5, 11].

En la identificación experimental o clásica se emplean señales especiales, siendo la señal escalón la más simple y utilizada con mucha frecuencia, además también se utilizan secuencias binarias pseudoaleatorias, y en ocasiones rampas, sinusoidales e impulsos. Estas técnicas experimentales son más directas de aplicar y son las que permiten obtener modelos aproximados a más corto plazo, aunque a veces se encuentra resistencia en personal de operaciones de planta en la introducción de señales de prueba en el proceso. Entre los procedimientos empleados para la identificación experimental se pueden nombrar entre otros los siguientes: método de tanteo (lazo cerrado), método de ganancia límite (lazo cerrado), método de curva de reacción (lazo abierto), método de “Bump Test” (lazo abierto), y en este campo se puede decir que la búsqueda de nuevos métodos es un tema abierto de investigación. También se tiene la identificación paramétrica que está basada en los llamados métodos de minimización del error de predicción, que son producto de la teoría tradicional de mínimos cuadrados [2, 5, 6, 7, 8, 9, 10, 11, 12].

En este trabajo el enfoque se va a plasmar en la identificación experimental basada en el método de prueba denominado “Bump Test” o Prueba Escalón según lo explica Kevin Starr en su libro “Single loop control methods” [2].

2.2. Identificación experimental de proceso mediante “Bump Test”

La obtención del modelo es una de las etapas esenciales en la meta deseada de la optimización de proceso. Como se puede ilustrar en la Figura 1, se tienen que realizar una serie de actividades para alcanzar dicha optimización. Esta serie de acciones incluye calibración de instrumentos, filtrado de señales, obtención de modelo y entonamiento de controlador.

La importancia de obtener un modelo del proceso radica en que se tienen métodos de sintonización del controlador basados en el modelo, tal es el caso del método de síntesis directa que se empleará en este trabajo. El método de síntesis directa se basa en las características del proceso para determinar los ajustes del controlador. Por medio de un método

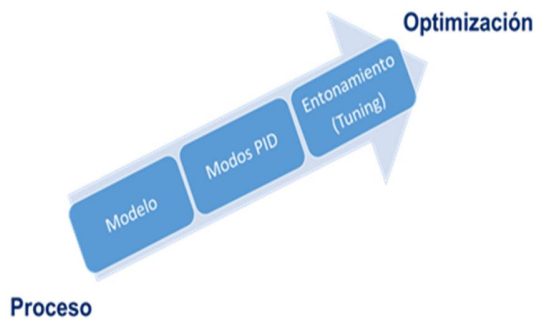


Figura 1: Secuencia de actividades hacia la optimización

de identificación experimental denominado Prueba Escalón o “Bump Test” se puede obtener un modelo aproximado del proceso y se puede predecir cómo responderá el proceso bajo ciertas condiciones. El procedimiento consiste en colocar la salida del controlador en modo manual, y registrar la salida del proceso ante una entrada de tipo escalón tal como se ilustra en la Figura 2. Este es un método de identificación experimental a lazo abierto.

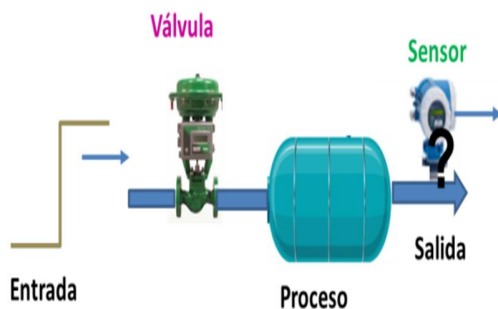


Figura 2: Esquema del procedimiento para el “Bump Test”

Se produce un cambio tipo señal escalón a la salida del controlador en modo manual para manipular la válvula de control y esta señal es la entrada al proceso. En la Figura 3 se supone que tanto el convertidor I/P como la válvula y el transmisor son ideales, y que el proceso es de 1er

orden, por lo que la salida será de primer orden. Se debe hacer un registro de los valores de la señal de salida para cada valor de la señal de entrada y así obtener la curva de respuesta y la dinámica del proceso.

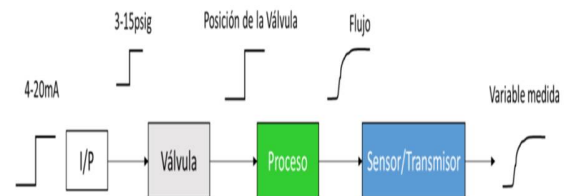


Figura 3: Diagrama de bloques de equipos para realizar el “Bump Test”

2.3. Descripción de la metodología de la Prueba Escalón (“Bump Test”)

Kevin Starr de ABB describe la siguiente metodología para la realización de la Prueba Escalón o “Bump Test”. Antes de hacer la prueba, es altamente recomendable realizar una revisión de la instrumentación de campo, actuador e instrumento de medición del lazo de control. Al iniciar la prueba escalón se debe procurar hacerla en un sentido (abriendo o cerrando la válvula) que mejore o favorezca al proceso. Al comienzo debe hacerse un cambio pequeño en el actuador, por ejemplo un cambio de magnitud de 5 o 10% y trabajando en una porción lineal del elemento final de control. Es importante permitir que el proceso se estabilice antes de generar el próximo escalón. Y también, el escalón se debe generar en direcciones diferentes (abriendo y cerrando la válvula), o sea, usar un método cíclico para la prueba escalón [2].

La prueba escalón o Bump Test se va a aplicar para los tipos de procesos llamados autorregulados. Un proceso autorregulado es aquel que, si todas las entradas son fijas, buscará su propio estado de equilibrio. Es una acción en el proceso que actúa para la restauración del equilibrio. La mayoría

de los procesos en la industria son del tipo llamado autorregulado (más del 90 % de los casos) [2, 13, 14].

2.4. Prueba Escalón “Bump Test” para un Proceso de primer orden

Un proceso de respuesta pura de primer orden no es muy común, pero puede ser usado para modelar casi todos los procesos industriales tales como flujos, presiones, filtros básicos entre otros. Para fines de análisis y diseño de sistemas de control, modelos de primer orden pueden ser perfectamente satisfactorios. Más aún, dichos modelos son el punto lógico de comienzo para el desarrollo y comprensión de modelos más complejos [2, 15].

El proceso de primer orden está definido por los dos parámetros, la ganancia de proceso K_p y la constante de tiempo τ_p . La ganancia del Proceso, K_p representa cuánto cambia en magnitud el proceso y la constante de tiempo, τ_p cuánto tiempo toma en hacerlo. La respuesta ante una entrada escalón del sistema de 1er orden descrito por la ecuación diferencial ordinaria (ecuación (1)):

$$\tau_p \frac{dY(t)}{dt} + Y(t) = K_p U(t), \quad (1)$$

viene dada por la ecuación (2)

$$Y(t) = U(t)K_p(1 - e^{-t/\tau_p}). \quad (2)$$

En la Figura 4 se muestra el esquema del “Bump Test” para el sistema o proceso de primer orden.

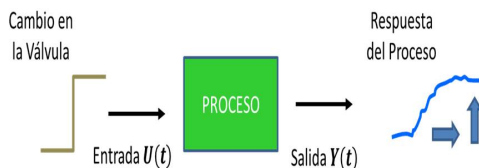


Figura 4: Prueba Escalón “Bump Test” para proceso de 1er orden

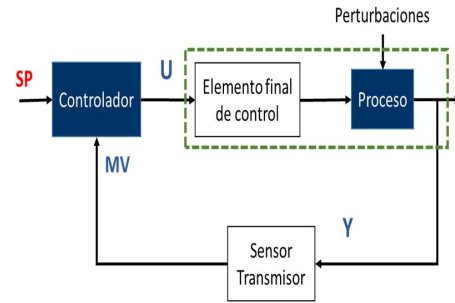


Figura 5: Diagrama del sistema de control en lazo cerrado

A continuación se muestra el esquema del lazo cerrado de control en la Figura 5, indicando la terminología empleada.

Donde SP es el Setpoint o punto de ajuste (valor deseado), PV : Variable de proceso, MV : Variable medida (“measured value” o valor medido), U es la salida del controlador o la entrada al proceso, Y es la entrada al controlador/salida del proceso.

Al efectuar la Prueba Escalón o Bump Test tal como se indica en la Figura 6, se efectúa un cambio del valor en la apertura de la válvula.

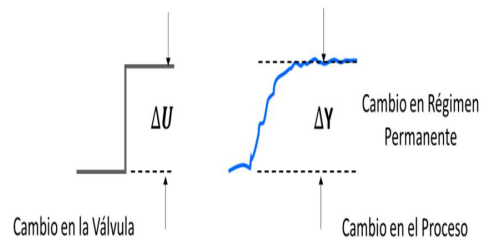


Figura 6: Cambio en la válvula y en el proceso en el “Bump Test”

Entonces la ganancia K_p se obtiene según la ecuación (3)

$$K_p = \frac{\text{Cambio en el Proceso } Y(t)}{\text{Cambio en la Entrada } U(t)} = \frac{\Delta Y(t)}{\Delta U(t)}. \quad (3)$$

En cuanto al valor de la constante de tiempo τ_p del proceso, se obtiene en base al tiempo de

transición o tiempo de ajuste dinámico que es aproximadamente igual a 4 constantes de tiempo ($4\tau_p$) del proceso. En la Figura 7 se ilustra esta aproximación en el proceso.

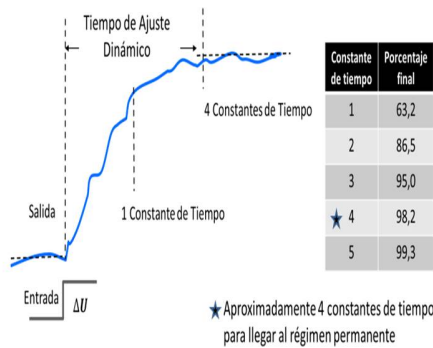


Figura 7: Aproximación de 4 constantes de tiempo del proceso

La constante de tiempo τ_p se define como el tiempo que toma el proceso en alcanzar el 63,2% de su valor en régimen permanente o estado estacionario. Es el tiempo que el proceso tarda en pasar de un valor de régimen permanente a otro para un cambio ΔU de valor del actuador. La constante de tiempo es también referida a veces como “retraso” (lag) debido a que la respuesta del proceso va por detrás del cambio en la entrada del proceso. A medida que la constante de tiempo es más pequeña, más rápido alcanza el proceso el nuevo valor de estado estacionario. La constante de tiempo guarda relación con la velocidad de respuesta del sistema y se puede aproximar usando el tiempo de transición dinámico (o tiempo de estado estacionario o régimen permanente) de acuerdo a la ecuación (4) [1, 2, 3, 16]:

Una constante de tiempo

$$\tau_p \cong \frac{\text{Tiempo de Transición Dinámica}}{4} \quad (4)$$

Es importante tomar en cuenta que las unidades de la constante de tiempo deben coincidir con las unidades de los parámetros del controlador.

A continuación se describe la dinámica del sistema de primer orden con tiempo muerto y la

forma en qué se aplica el método de “Bump Test” para determinar el modelo de este proceso.

2.5. Prueba Escalón “Bump Test” para Proceso de primer orden con retardo

En un proceso de primer orden con tiempo muerto se tienen tres parámetros para describir su dinámica: ganancia de proceso K_p , constante de tiempo τ_p y retardo de proceso (o tiempo muerto) T_{DT} . Este modelo de 3 parámetros es la representación más usada para la dinámica de procesos. El tiempo de retardo o tiempo muerto constituye un aspecto fundamental en los procesos y debe ser considerado en el diseño del sistema de control para evitar inestabilidad en el proceso y lograr una buena sintonización del controlador [17, 18].

La expresión para la salida $Y(t)$ ante un cambio escalón $\Delta U(t)$, viene dada por la ecuación (5)

$$Y(t) = U(t - T_{DT})K_p \left(1 - e^{-(t-T_{DT})/\tau_p}\right) \quad (5)$$

Al realizar el “Bump Test”, la ganancia K_p y la constante de tiempo τ_p se obtienen de la misma forma que para el proceso de primer orden, ecuaciones (3) y (4). El tiempo muerto se puede determinar haciendo el cálculo del tiempo cuando el proceso comienza a responder menos el tiempo cuando el actuador tuvo el cambio, como se muestra en la ecuación (6):

$$T_{DT} = t_{\text{respuesta del proceso}} - t_{\text{acción del actuador}} \quad (6)$$

En la Figura 8, se puede observar la aproximación del modelo de tres parámetros usada al convertir el cambio de la válvula en una respuesta del proceso.

Cuando el retardo no es dominante, o sea, T_{DT} es mucho menor que la constante de tiempo del proceso τ_p , entonces se puede usar un modelo (primer orden puro) de 2 parámetros para aproximar el proceso de primer orden con tiempo muerto. En la Figura 9 se ilustra la aproximación del proceso de primer orden con tiempo muerto usando un modelo de 2 parámetros (K_p, τ_p') [2].

Para obtener el modelo de 2 parámetros aproximado para el proceso de 1er orden con tiempo muerto se necesita:

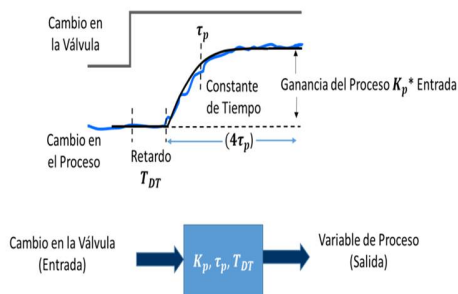


Figura 8: Aproximación del modelo de los tres parámetros K_p, τ_p , y T_{DT}

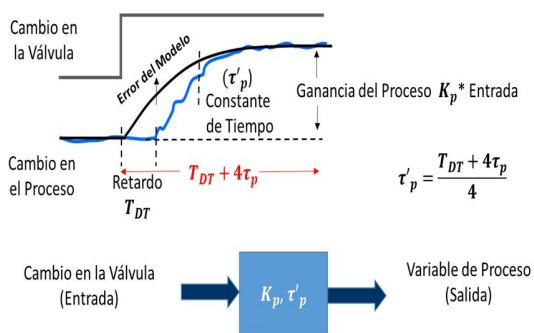


Figura 9: Aproximación del proceso de primer orden con tiempo muerto usando un modelo de 2 parámetros (K_p, τ'_p)

1. Realizar la prueba escalón o “Bump Test”.
2. Registrar la respuesta del proceso.
3. Hacer una aproximación de 1er orden sobre la respuesta observada.
4. Determinar la ganancia del proceso K_p y la constante de tiempo τ'_p del modelo aproximado de 1er orden [2].

3. Proceso de regeneración catalítica

En este trabajo se considera el proceso de regeneración catalítica descrito por Smith, Corripio [1] y que se muestra en la Figura 10.

El objetivo que se quiere lograr en este proceso es la regeneración de un catalizador de un reactor

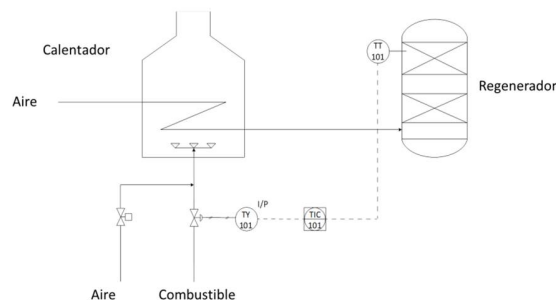
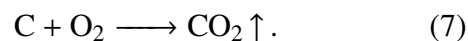


Figura 10: Control Feedback simple para un sistema de regeneración catalítica

químico. Dicho catalizador se emplea en un reactor en el cual se deshidrogena un hidrocarburo. Al transcurrir cierto tiempo, el carbón (C) queda depositado sobre el catalizador, causando su contaminación. Ante este evento, el catalizador pierde su condición y requiere ser regenerado. Para regenerarlo se procede a una etapa donde se quema el carbón depositado, soplando aire caliente en la capa de catalizador, reaccionando el oxígeno del aire con el carbón, formándose CO_2 de acuerdo con la ecuación (7):



Una vez que se ha quemado todo el carbón, el catalizador está preparado para ser aprovechado otra vez. Estamos en presencia de un proceso por lotes donde la etapa de la quema del carbón puede durar unas cuantas horas y el estudio del lazo de control puede ser realizado considerando un proceso continuo en el intervalo de tiempo que tome la regeneración.

La variable controlada en este caso es la temperatura de la capa del catalizador PV: TT101, variable muy importante. Una temperatura alta de la capa podría dañar el catalizador; por el contrario, si es muy baja el tiempo de la quema sería muy largo. La variable manipulada viene a ser el flujo de combustible hacia el calentador ($Out : TY101$). Y se considera la temperatura de aire de entrada como una perturbación (DV). Por fines de simplificación en este esquema de control se ha omitido el control

de relación aire/gas del horno. Del mismo modo, se muestra un solo medidor de temperatura, en la práctica real existen varios medidores y se toma el promedio o la temperatura más alta de la capa. En la Figura 11 se muestra el diagrama de bloques del sistema representado en la Figura 10 para el proceso de regeneración catalítica con control Feedback simple.

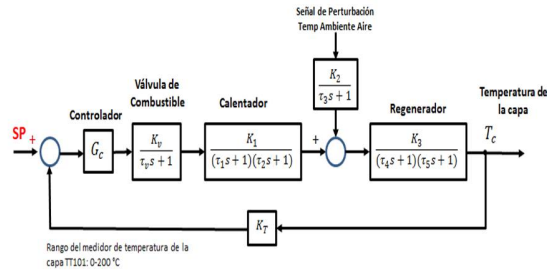


Figura 11: Diagrama de bloques del control Feedback del Regenerador Catalítico

Los valores de los distintos parámetros del sistema vienen dados por [1]:
 $\tau_1 = 3 \text{ min}$ $\tau_2 = 1 \text{ min}$ $\tau_3 = 3 \text{ min}$ $\tau_4 = 4 \text{ min}$
 $\tau_5 = 1 \text{ min}$ $\tau_v = 0,2 \text{ min}$ $K_1 = 1$ $K_2 = 1,5$
 $K_3 = 0,8$ $K_v = 3 \text{ gpm}/\%CO$ $K_T = 0,5 \text{ \%TO}/C$

4. Métodos de sintonización de controlador PID

Antes de explicar los dos métodos que son estudiados, aplicados, analizados y comparados en este trabajo, el método de síntesis directa y el método aproximado organizado de ensayo y error, inicialmente se expone un breve resumen sobre algunas de las formas de clasificación de los distintos métodos disponibles para la sintonización o entonamiento del controlador PID, o sea, la determinación óptima de los parámetros del controlador.

4.1. Clasificación de los métodos de sintonización de controlador PID

Existe una gran variedad de métodos de sintonización de un controlador PID en la bibliografía de sistemas de control, y actualmente sigue siendo un tema continuo

de investigación en el área de control de procesos. A continuación se describen brevemente algunos enfoques de clasificación de estos métodos.

4.1.1. Métodos de sintonización a lazo abierto y métodos a lazo cerrado

Una forma de clasificación está basada en el hecho de si la sintonización (la prueba o experiencia) se realiza a lazo abierto o a lazo cerrado y se le denomina por tanto métodos de sintonización a lazo abierto o métodos de lazo cerrado. Básicamente, la denominación se refiere a la naturaleza de la prueba realizada en el sistema para determinar un modelo del proceso con el cual se pueden generar unas reglas o expresiones para calcular los parámetros involucrados en la sintonización del controlador PID.

4.1.2. Métodos de sintonización con relaciones empíricas y métodos de síntesis directa [3]

Otra forma de clasificación se fundamenta en el tipo de enfoque teórico-práctico y se tienen entonces los métodos con relaciones empíricas y los métodos de sintonización con relaciones de control de modelo interno y síntesis directa (o también llamados diseño basado en modelo).

La sintonización con métodos empíricos, en forma aproximada, envuelve la realización de una prueba experimental a lazo abierto o lazo cerrado y el ajuste de la respuesta a un modelo. Los parámetros del controlador se calculan en base a este modelo y a una tabla de relaciones empíricas. No se puede imponer condiciones en las especificaciones de la respuesta dinámica del sistema. Muchos de estos métodos de relaciones empíricas se basan en datos de lazo abierto y ajustes a un modelo de primer orden con tiempo muerto. Como ejemplos se tienen el método de la curva de reacción de proceso, el método del criterio de mínimo error integral, el método de ciclo final (“ultimate-cycle”) de Ziegler-Nichols, entre otros. En estos métodos se trabaja en el entonamiento hasta que se tenga la respuesta deseada, el punto de partida no es el proceso.

En los métodos de síntesis directa y de control de modelo interno, el diseño del controlador se basa en un modelo del proceso y en una respuesta deseada de lazo cerrado (especificada generalmente para cambios de Setpoint). El resultado de la aplicación del método de síntesis directa depende del modelo del proceso. Con este método se puede obtener una respuesta sin sobreimpulso y con un tiempo de respuesta requerido

por la planta. Al método de síntesis se le llama también diseño basado en modelo, siendo el proceso y su dinámica el punto de partida, o sea, el método empieza con la identificación experimental y la especificación de la respuesta deseada del sistema [3, 19].

4.1.3. Métodos de sintonización basados en el tipo de proceso: autorregulado o integrativo

Harold Wade en su libro “Basic and Advanced Regulatory Control: System Design and Application” [13] presenta una clasificación de los métodos de entonamiento basada en el tipo de proceso, esto da origen a los llamados métodos para entonamiento de procesos autorregulados (sintonización por ensayo y error, sintonización por pruebas de lazo abierto, sintonización mediante pruebas de lazo cerrado, sintonización de mejora del estado encontrado) y métodos para procesos integrativos (sintonización de lazos de control idealizados de nivel de líquidos, y otros enfoques). El método presentado por Harold Wade denominado sintonización organizada por ensayo y error [4, 13] se explica con mucho detalle en la próxima sección para luego ser comparado con el método de síntesis directa presentado por Kevin Starr de la ABB [2].

También se puede hablar de los métodos de tipo manual y los métodos de auto-tuning presentes en el diseño y configuración de muchos controladores industriales fabricados en la actualidad. A continuación se presentan los dos métodos que son el objeto de estudio, análisis y comparación en la presente investigación.

4.2. Método de Síntesis Directa

En este método se define la respuesta de lazo cerrado que se desea obtener y se diseña el controlador de acuerdo con esa respuesta. Con la síntesis directa se seleccionan las ganancias del controlador en base a los parámetros del modelo del proceso por lo que se denomina diseño basado en modelo. A la Síntesis Directa también se le conoce como Método de no sobreimpulso ya que se puede especificar la respuesta deseada de lazo cerrado como una señal aproximada de primer orden que no presente ni oscilaciones ni sobreimpulso [2, 3].

Para la deducción o explicación de este método se va a tomar en cuenta un diagrama de bloques de lazo cerrado del sistema, donde la función de transferencia $G(s)$ representa todas las funciones de transferencia del paso

directo distintas al controlador, $G(s) = G_v(s)G_p(s)$, tal como se ilustra en la Figura 12.

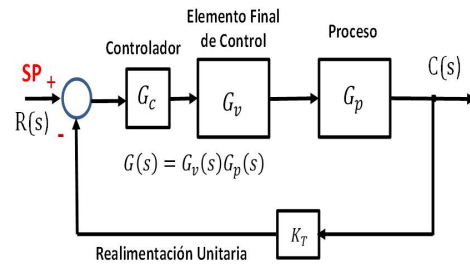


Figura 12: Diagrama de bloques del sistema de control a lazo cerrado

La función de transferencia $G_{lc}(s)$ de lazo cerrado viene dada por la ecuación (8)

$$G_{lc}(s) = \frac{C(s)}{R(s)} = \frac{G_c(s)G(s)}{1 + G_c(s)G(s)}, \quad (8)$$

y a partir de la ecuación (8), la expresión para $G_c(s)$ resulta en la ecuación (9).

$$G_c(s) = \frac{1}{G(s)} \frac{C(s)/R(s)}{1 - C(s)/R(s)}. \quad (9)$$

La ecuación (9) permite obtener una expresión para el controlador $G_c(s)$ en función de la relación salida/entrada $G_{lc}(s) = C(s)/R(s)$ y del proceso $G(s)$. Si se especifica como relación salida/entrada deseada una respuesta aproximada de primer orden dada por la ecuación (10)

$$G_{lc}(s) = \frac{C(s)}{R(s)} = \frac{1}{\tau_{lc}s + 1}, \quad (10)$$

donde τ_{lc} es la constante de tiempo del sistema en lazo cerrado, entonces la expresión para el controlador viene dada por la ecuación (11)

$$G_c(s) = \frac{1}{G(s)} \frac{G_{lc}(s)}{1 - G_{lc}(s)} = \frac{1}{G(s)} \frac{1}{\tau_{lc}s}, \quad (11)$$

y considerando el proceso de primer orden (como resultado del “Bump Test”), se tiene la ecuación (12)

$$G(s) = \frac{K_p}{\tau_p s + 1}. \quad (12)$$

Entonces, la expresión del controlador vendrá dada por la ecuación (13)

$$\begin{aligned}
 G_c(s) &= \frac{1}{G(s)} \frac{1}{\tau_{Ic}s} \\
 &= \frac{\tau_p s + 1}{K_p} \left(\frac{1}{\tau_p s} \right) \\
 &= \frac{\tau_p}{K_p \tau_{Ic}} \left[1 + \frac{1}{\tau_p s} \right].
 \end{aligned}
 \tag{13}$$

Tomando en cuenta el caso del algoritmo PI en la forma Standard, No interactiva, se tiene la siguiente expresión del controlador (ecuación(14)):

Controlador PI en forma Standard $= K_C \left[1 + \frac{1}{T_I s} \right]$, (14)

por tanto, comparando las ecuaciones (13) y (14) se tiene que los valores resultantes de las Reglas de Sintonización por el método de Síntesis Directa de Controlador PI en forma Standard usando modelo aproximado de primer orden para el proceso en lazo abierto resultan en las ecuaciones (15) y (16).

Ganancia del controlador PI $K_C = \frac{\tau_p}{K_p \tau_{Ic}} = \frac{1}{K_p \tau_{ratio}}$, (15)

Tiempo integral $T_I = \tau_p$, (16)

donde $\tau_{ratio} = \left(\frac{\tau_{Ic}}{\tau_p} \right)$ está vinculado con la velocidad de respuesta, K_p es la ganancia de proceso de lazo abierto y τ_p es la constante de tiempo de lazo abierto. Se puede observar que la ganancia proporcional K_C del controlador y la ganancia de proceso K_p a lazo abierto son inversamente proporcionales y el tiempo integral T_I del controlador es igual al valor de la constante de tiempo τ_p del proceso a lazo abierto.

En la aplicación del método Síntesis Directa se pueden establecer los siguientes pasos: 1) Realización de la prueba escalón o “Bump Test” para determinar modelo aproximado del proceso como de primer orden, 2) Obtención de los parámetros del controlador que permitan la respuesta deseada en lazo cerrado (en modo automático), y 3) Consideración del valor del parámetro τ_{ratio} para establecer la velocidad de la respuesta. El

parámetro τ_{ratio} se usa para controlar la velocidad a la cual el proceso responde a una perturbación o a un cambio de Setpoint [2].

El τ_{ratio} es el único parámetro de sintonización ya que K_p y τ_p son parámetros de configuración basados en el modelo del proceso a lazo abierto. Se debe tener en cuenta el manejo de las unidades de ingeniería con las cuales se realiza la identificación del proceso, dichas unidades deben coincidir con las unidades de los parámetros del controlador [2].

El valor de τ_{ratio} debe ser determinado en base a la confiabilidad del modelo aproximado de primer orden (obtenido de la prueba escalón “Bump Test”) y también se debe tomar en cuenta el posible desgaste de la válvula y la condición del proceso aguas arriba. A continuación en la Figura 13 se ilustra la consideración que debe tenerse en la relación entre el τ_{ratio} y la confiabilidad del modelo aproximado de primer orden obtenido para el proceso a lazo abierto obtenido del “Bump Test” [2].

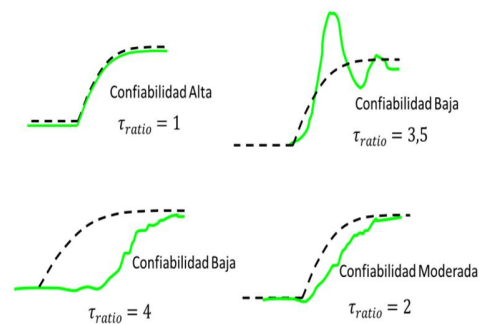


Figura 13: Relación entre el τ_{ratio} y la confiabilidad del modelo de proceso [2]

La confiabilidad del modelo es una forma de observar el error de la aproximación del modelo de 1er orden obtenido de la identificación experimental del proceso (“Bump Test”). Mientras mayor sea el error del modelo se va a requerir un valor mayor de τ_{ratio} . Para un valor alto de τ_{ratio} se obtiene una respuesta lenta del sistema y para un valor τ_{ratio} pequeño, la respuesta es rápida. En la Figura 14 se muestra la relación entre la velocidad de respuesta, el valor de τ_{ratio} , la confiabilidad del modelo y la ganancia del controlador. El tiempo integral se mantiene igual (Modo de controlador PI en forma Standard no interactiva). El valor de τ_{ratio} es una manera de cuantificar o medir el error del modelo.

Confiabilidad del Modelo de 1er Orden	τ_{ratio}	Velocidad de Respuesta	Ganancia del Controlador	Tiempo Integral
Alta	1	Rápida	$K_C = \frac{1}{K_P \cdot 1}$	$T_I = \tau_P$
	2		$K_C = \frac{1}{K_P \cdot 2}$	$T_I = \tau_P$
	3		$K_C = \frac{1}{K_P \cdot 3}$	$T_I = \tau_P$
Baja	4	Lenta	$K_C = \frac{1}{K_P \cdot 4}$	$T_I = \tau_P$

Figura 14: Relación entre τ_{ratio} y la velocidad de respuesta del sistema [2]

Con el método de síntesis directa se puede entonces obtener un proceso estable en modo automático (sistema en lazo cerrado), se puede escoger la velocidad de respuesta en base al valor de τ_{ratio} y es un método basado en el modelo del proceso, i.e., un enfoque que arranca con la determinación del modelo en lazo abierto del proceso. Se debe tener presente que se está trabajando con procesos autorregulados, con una aproximación de primer orden del modelo del proceso y con la forma PID Standard, No interactiva Modo Proporcional-Integrativo PI.

4.3. Método Organizado Aproximado de Ensayo y Error

El método basado en ensayo y error (“Trial and Error”) consiste en la realización del ajuste manual de los parámetros PID del controlador mediante tanteo, es muy utilizado en la práctica industrial cotidiana, en algunas ocasiones es más un arte que una ciencia y depende mucho de la experiencia. Sin embargo, se recomienda realizarlo de forma organizada para lograr resultados satisfactorios, en términos de calidad y tiempo invertido en el entonamiento. En muchas ocasiones, este método se emplea en lazos de control que han estado en funcionamiento (“As found state”) y que requieren ser entonados para mejorar su desempeño por presentar oscilaciones. Dichas oscilaciones pueden ser ocasionadas por cambios en el proceso o luego de algún mantenimiento de la instrumentación de campo.

En esta sección se van a describir tres métodos presentados por la Sociedad Internacional de Automatización (ISA por sus siglas en inglés) llamados: 1) Entonamiento de lazo manual, 2) Entonamiento basado

en la experiencia y 3) Método inteligente de ensayo y error “Intelligent trial-and-error” [4, 13, 20, 21, 22].

4.3.1. Entonamiento de Lazo Manual

Este método presentado por Chuck Cornell en el: “Manual de Referencia Técnica para Ingenieros de Sistemas de Control” (Control Systems Engineer Technical Reference Handbook, ISA 2012) consiste en realizar el entonamiento del lazo de control online con la planta en funcionamiento mediante ensayo y error, aplicando los siguientes pasos para un controlador PID paralelo [20]:

Con el controlador en modo manual, ajustar la salida del controlador hasta que sea posible mantener la PV deseada. Colocar el K_I y K_D en cero. Colocar el controlador en modo automático e ir incrementado K_P hasta que la PV comience a oscilar. Luego disminuir a la mitad el valor de K_P para una respuesta de 1/4 de decaimiento. Luego incrementar K_I hasta corregir el offset dentro del tiempo suficiente de respuesta del proceso (Precaución: mucho K_I podría causar inestabilidad en el proceso). Luego, si es requerido, incrementar K_D hasta lograr obtener una respuesta rápida para alcanzar el setpoint ante cambios de carga o perturbaciones. (Precaución: mucho K_D podría causar una respuesta excesiva u overshoot).

En un lazo de control PID rápido, el entonamiento puede producir ligeros overshoot para alcanzar el setpoint de forma más rápida. Sin embargo, en algunos sistemas es posible que esta respuesta no sea aceptable. En ese caso, se va a requerir un valor significativamente bajo a la mitad de K_P que se había seleccionado.

Este método podría tomar mucho tiempo al requerir varios intentos de entonamiento. También, podría ocasionar situaciones riesgosas durante su ejecución.

4.3.2. Entonamiento Basado en la Experiencia

En este método presentado en el “Manual Condensado de Medición y Control de la ISA” (Condensed Handbook of Measurement and Control, ISA 2007), el entonamiento basado en la experiencia se hace introduciendo en el controlador las constantes P , I y D , partiendo de valores iniciales ya conocidos [21]. Esto se realizaría de forma aproximada y generalmente se requiere más de un intento de entonamiento. Se debe ir realizando ajustes finos de los parámetros hasta conseguir la respuesta del proceso esperada.

4.3.3. Método inteligente de ensayo y error “Intelligent trial-and-error”

Es un método en lazo cerrado que consiste en realizar el entonamiento (“Tuning”) de un controlador modo PI standard para procesos autorregulados (la mayoría de los lazos de control son de este tipo) de forma organizada, por tanteo y error. Este método fue presentado por Harold Wade, publicado en un artículo de la ISA InTech Magazine 2005 llamado: “Trial and Error: An organized procedure” [4]. La idea principal de este método es buscar un comportamiento aceptable en forma rápida, procurando un número mínimo de cambios en los parámetros y evitando causar interferencia al proceso controlado.

Este análisis tiene como punto de partida los parámetros de control encontrados en el sistema (“as-found data”), la ganancia del controlador y el tiempo integral. Se supone que se han descartado otros tipos de inconvenientes, que lo que se tiene es un problema de sintonización y que se quiere obtener un lazo estable (sin oscilaciones), el cual retorne rápido al punto de Setpoint después de algún cambio de carga. Si el lazo se encuentra oscilando, dos parámetros van a caracterizar este comportamiento del proceso, la razón de decaimiento DR y el periodo de oscilación P . El periodo de oscilación y el tiempo integral deben mostrar las mismas unidades, P en minutos y T_I en minutos (por repetición). Los valores existentes de los parámetros (“as-found data”) vienen a ser el punto de comienzo del análisis y procedimiento de sintonización del lazo de control. En el caso de usar algún método de Ziegler-Nichols se estaría comenzando desde cero, sin tomar en cuenta este valioso conocimiento inicial (llamado “as-found data”) [4, 13, 22].

Este método se fundamenta en que un lazo de control está bien entonado cuando se cumple la ecuación (17), lo cual asegura un decaimiento rápido de las oscilaciones:

$$1,5 < \frac{P}{T_I} < 2, \quad (17)$$

donde P es el período de oscilación (min) y T_I es el tiempo integral (min/rep). Las ecuaciones (18) y (19), son expresiones equivalentes que también pueden ser útiles

$$1,5T_I < P < 2T_I, \quad (18)$$

$$0,5P < T_I < 0,67P. \quad (19)$$

Se consideran respuestas aceptables aquellas que presentan un tiempo de decaimiento (“Decay Ratio”) menor a 1/4. Con este criterio se quiere establecer un compromiso o balance entre la respuesta más aceptable (sin sobre impulso) a un cambio de Setpoint y una respuesta satisfactoria a perturbación de carga.

El procedimiento se puede resumir de la siguiente forma (Reglas Si-Entonces) [4, 13, 22]:

- Si el lazo no presenta oscilaciones, entonces se debe aumentar la ganancia del control hasta lograr oscilación amortiguada en el lazo y así poder observar el periodo de oscilación.
- Si el lazo se encuentra oscilando pero el periodo de oscilación no está en el rango 1,5 a 2; entonces se debe reajustar el tiempo integral para que quede en el rango de 1/2 a 2/3 del periodo actual.
- Si el periodo de oscilación se encuentra en el rango 1/2 a 2 veces el tiempo integral T_I , entonces se debe ajustar la ganancia del controlador hasta que se tenga un 1/4 de decaimiento u otra característica deseada de amortiguamiento.

Al hacerse un cambio de un parámetro de control, el efecto de este cambio debe determinarse efectuando un cambio pequeño de Setpoint. Se sugiere cambiar un solo parámetro a la vez. Se aconseja tomar nota de cambios consecutivos de los parámetros de control y respuesta del sistema para observar el avance en cada ensayo. Generalmente, se consigue un buen desempeño después de 3 o 4 pruebas [4, 13, 22].

5. Simulaciones realizadas en Scilab-XCOS

A efectos de realizar la identificación del modelo del proceso de regeneración catalítica y la comparación de los métodos de sintonización estudiados se construyó en el software Scilab-Xcos el siguiente diagrama de bloque mostrado en la Figura 15.

Para simular lo que sería una sesión real de sintonización de un lazo de control, se realizaron ajustes manuales de los parámetros del controlador PID de tal forma de conseguir una respuesta oscilatoria típica, la cual es posible encontrar luego de una jornada de mantenimiento de planta y en específico de los instrumentos de campo y válvulas de control o por cambios en las condiciones del proceso, cuando el operador de planta reporta un desperfecto del lazo de control. A continuación, en la Figura 16 se muestra el comportamiento del lazo de control al incrementar el Setpoint de 50 % a 60 %, con unos parámetros de

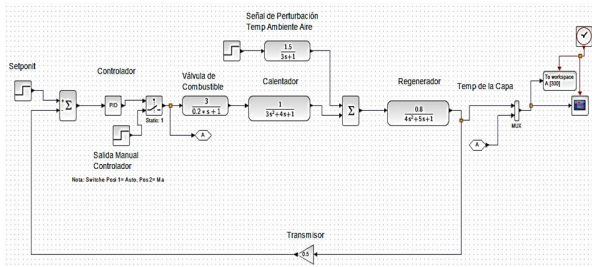


Figura 15: Diagrama de bloque del control del proceso en Scilab-Xcos

entonamiento iniciales de: $K_C = 1,27$; $K_I = 0,27$ rep/min y $T_D = 0$ min. Como información importante se tiene que el rango del instrumento de medición de temperatura de la capa es de 0 a 200 °C y que la forma del controlador en Xcos es PID Paralelo.

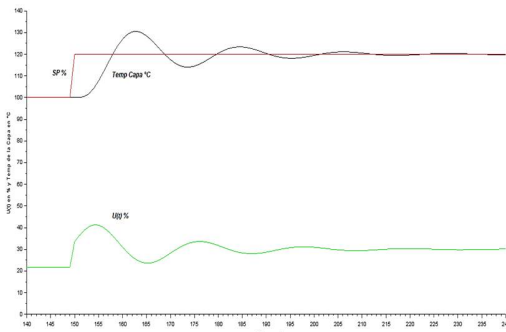


Figura 16: Comportamiento inicial del lazo de control

Como se observa, la respuesta del lazo de control es inadecuada y a partir de la misma se aplicarán los métodos de sintonización estudiados.

5.1. Simulación del lazo de control aplicando el método de síntesis directa

En esta simulación se colocó el lazo de control en modo abierto, es decir, en modo manual, colocando el selector del diagrama de bloque del sistema Xcos en la posición 2 y se realizó la prueba escalón (“Bump

Test”) variando la salida $U(t)$ un 10 % obteniendo los resultados mostrados en la Figura 17.

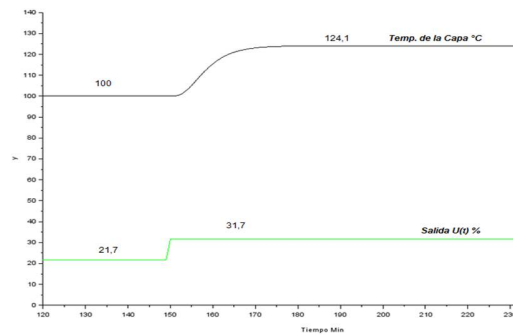


Figura 17: Prueba escalón (“Bump Test”) al proceso

Al procesar mediante Excel los datos obtenidos de la prueba escalón, representados en la Figura 18, es posible obtener los parámetros del proceso como se muestra en las ecuaciones (20) y (21).

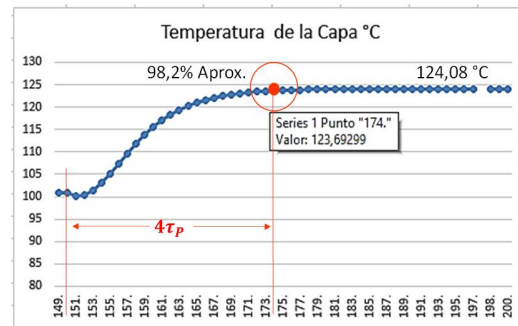


Figura 18: Datos obtenidos de la prueba escalón

$$K_p = \frac{(124,08 - 100) \text{ }^\circ\text{C}}{10 \text{ } \%,} \cdot 100 \text{ } \% = 1,2; \quad (20)$$

$$\tau_p = \frac{(174 - 150) \text{ min}}{4} = 6 \text{ min.} \quad (21)$$

Entonces, es posible aproximar el proceso de acuerdo a la ecuación (22):

$$G_p(s) = \frac{1,2}{6s + 1}. \quad (22)$$

Se considera despreciable el tiempo muerto en la respuesta del proceso.

Al aplicar las reglas de entonamiento del método de síntesis directa se obtienen los siguientes parámetros de entonamiento: Para $\tau_{ratio} = 1$ según la ecuación (23):

$$K_C = \frac{1}{K_p \tau_{ratio}} = \frac{1}{1,2} = 0,83; \quad (23)$$

se obtiene: $K_C = 0,83$ y $T_I = 6$ min.

Para $\tau_{ratio} = 2$ según la ecuación (24):

$$K_C = \frac{1}{1,2 \cdot 2} = 0,42; \quad (24)$$

se obtiene: $K_C = 0,42$ y $T_I = 6$ min.

Al transformar T_I del controlador estándar a paralelo se obtiene por medio de la ecuación (25):

$$K_I = \frac{K_C}{T_I} = \frac{0,42}{6} = 0,07 \text{ rep/min.} \quad (25)$$

A continuación en la Figura 19 se muestra los resultados de las simulaciones al realizar un cambio de setpoint de un 50 a 60 %.

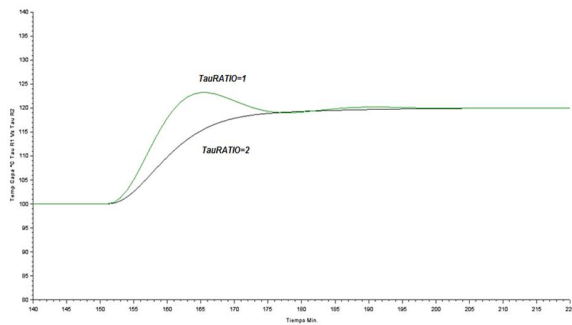


Figura 19: Respuesta del lazo de control aplicando método de síntesis directa

Como se observa en la Figura 19, la respuesta del proceso para $\tau_{ratio} = 2$ se comporta de forma bastante estable y sin sobre impulso.

5.2. Simulación del lazo de control aplicando el método organizado aproximado de ensayo y error

En esta simulación se coloca el lazo de control en lazo cerrado, es decir, en modo automático, colocando

el selector del diagrama de bloque del sistema Xcos en la posición 1. Al igual que en el punto anterior, este método se aplica al lazo de control con respuesta oscilatoria mostrado en la Figura 16. Como punto principal hay que analizar los datos de la respuesta del proceso con el objeto de evaluar la condición reflejada en la ecuación (26)

$$1,5 < \frac{P}{T_I} < 2. \quad (26)$$

Al analizar la respuesta original del proceso con Excel se encontraron los siguientes resultados mostrados en la Figura 20:

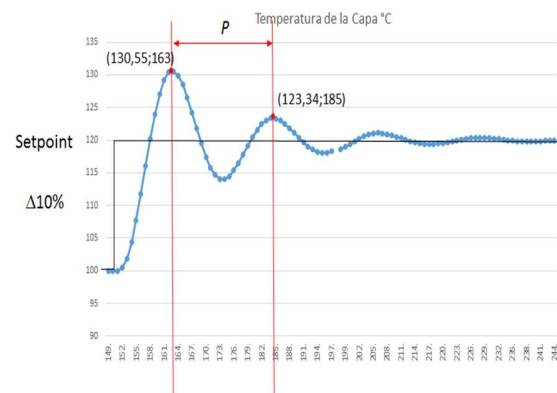


Figura 20: Análisis de la respuesta original del proceso oscilatorio

Al realizar la validación de la relación P/T_I se encontró lo siguiente: $P = 185 - 163 = 22$ min

Como el parámetro $K_I = 0,27$ rep/min del controlador PID de Xcos corresponde a un controlador PID paralelo, se debe convertir a forma estándar para hallar T_I y determinar la relación P/T_I . Al realizar los cálculos, según las ecuaciones (27), (28), (29) y (30), se obtiene:

$$\frac{K_C}{T_I} = K_I, \quad (27)$$

$$T_I = \frac{1,27}{0,27} = 4,7 \text{ min/rep,} \quad (28)$$

$$\frac{P}{T_I} = \frac{22 \text{ min}}{4,7 \text{ min/rep}} = 4,68 > 2; \quad (29)$$

$$\begin{aligned} \text{Decay R} &= \frac{123,34 - 120}{130,55 - 120} \\ &= 0,32 > 1/4. \end{aligned} \quad (30)$$

Como era de esperarse, la respuesta del lazo de control se encuentra totalmente fuera de rango en los

parámetros y se requiere el entonamiento. Se efectúa un primer ajuste del T_I empleando este método organizado aproximado de ensayo y error; para lograr este ajuste se debe disminuir la relación P/T_I de tal forma que sea menor a 2. Se inicia aumentando T_I a 12 min/rep, en este caso la relación sería (ecuación (31))

$$\frac{P}{T_I} = \frac{22}{12} = 1,83 < 2. \quad (31)$$

Entonces de acuerdo a lo expresado en la ecuación (32),

$$K_I = \frac{K_C}{T_I} = \frac{1,27}{12} = 0,1 \text{ rep/min} \quad (32)$$

Los parámetros PI del controlador serían: $K_C = 1,27$ $K_I = 0,1$ rep/min. En la Figura 21 se muestra el comportamiento del lazo de control al realizar una variación en el setpoint de 60 a 50 %.

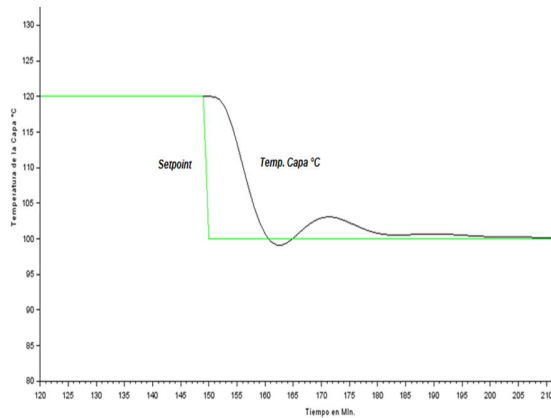


Figura 21: Respuesta del lazo de control aplicando método organizado aproximado de ensayo y error

Al analizar los datos de la respuesta del lazo de control en Excel se obtuvieron los resultados mostrados en la Figura 22.

A continuación se presenta la verificación del primer entonamiento: $P = 182 - 163 = 19$ min, de acuerdo con la ecuación (33)

$$\frac{P}{T_I} = \frac{19 \text{ min}}{12 \text{ min/rep}} = 1,58 < 2 \quad (33)$$

Al analizar los resultados se obtiene: Decay $R \approx 0$ y $1,5 < P/T_I < 2$.

Aunque se observa un leve sobre-impulso en la respuesta del lazo de control, se cumplen las condiciones de estabilidad y se logra realizar la sintonización con tan solo un intento de entonamiento.

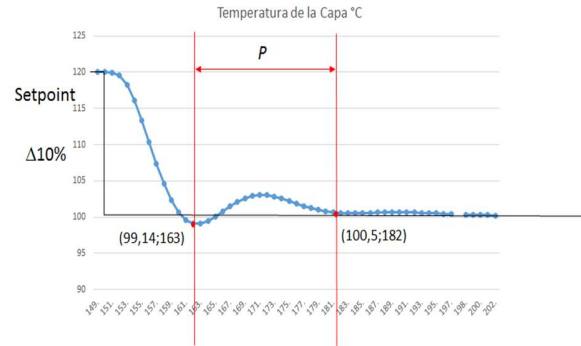


Figura 22: Análisis de respuesta del proceso con el primer intento de entonamiento

6. Evaluación comparativa de los métodos de sintonización

En este punto se evalúan en detalle las ventajas y desventajas entre el método de síntesis directa y el método organizado aproximado de ensayo y error, con el objeto de determinar la conveniencia de la aplicación de cada uno en el proceso industrial estudiado. Es evidente que al observar la respuesta original oscilatoria del lazo de control al realizar cambios de setpoint presentado en la Figura 16, se obtiene una notable mejora al emplear ambos métodos de sintonización de controladores PI. En la Figura 23 es posible obtener los resultados comparativos entre ambos métodos.

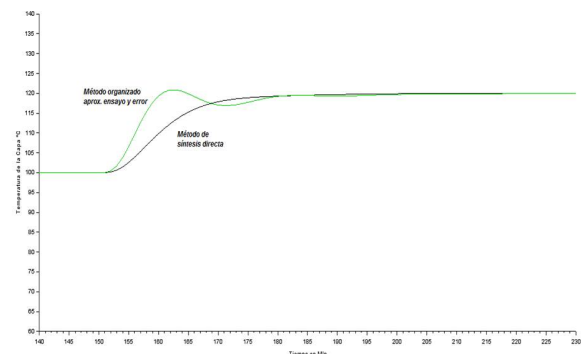


Figura 23: Resultados comparativos de los métodos de sintonización

Al detallar la respuesta del proceso en cada caso, se observa que el ajuste del controlador mediante el método

organizado aproximado de ensayo y error ocasiona un leve sobre impulso en comparación al ajuste realizado a través de síntesis directa. Se observa claramente un comportamiento más estable al realizar el ajuste del controlador mediante el método de síntesis directa. Por otro lado, al analizar el comportamiento de la salida del controlador $U(t)$ mostrado en la Figura 24, se observa que el comportamiento del controlador al realizar los ajustes mediante el método de síntesis directa tiene un comportamiento más estable en comparación al ajuste a través del método organizado aproximado de ensayo y error, el cual presenta una señal de salida más brusca.

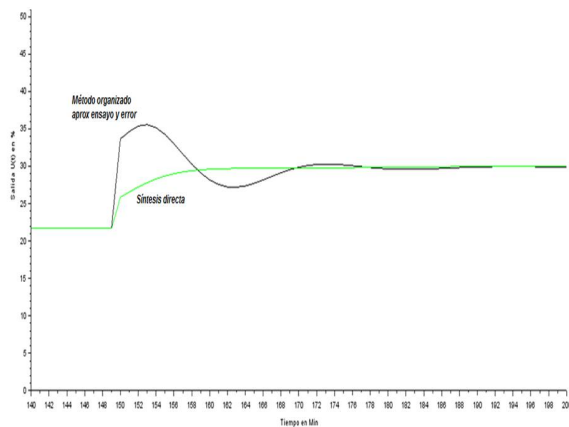


Figura 24: Resultado comparativo de la salida del controlador con ambos métodos

También se evaluó el comportamiento del proceso ante la variable de perturbación, en la Figura 25 se presentan los resultados de la simulación al realizar una variación de la temperatura del aire de entrada de 40 a 30 °C.

Se observa que el controlador ajustado mediante el método organizado aproximado de ensayo y error presentó una mejor respuesta ante la perturbación, ya que permitió estabilizar la temperatura de la capa de forma más rápida en comparación al ajuste realizado mediante el método de síntesis directa.

7. Conclusiones

La identificación experimental de procesos es una herramienta poderosa para aproximar el modelo matemático de un proceso alrededor de su punto de operación, lo cual representa una gran ventaja sobre la obtención del modelo matemático mediante sus ecuaciones o principios fundamentales, ya que en ocasiones es casi imposible conseguir toda la

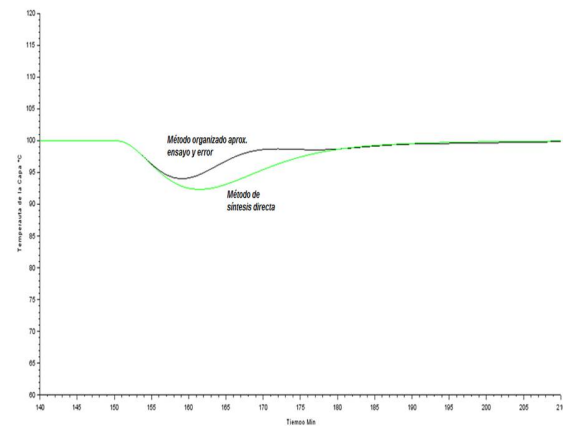


Figura 25: Resultado comparativo de los métodos ante variable de perturbación

información requerida para determinarlo. El método de identificación presentado en este trabajo es bastante eficiente, práctico y fácil de implementar, siempre y cuando se dispongan de las herramientas adecuadas de recolección de datos.

Al comparar los métodos de sintonización organizado aproximado de ensayo y error y síntesis directa es posible concluir que ambos métodos funcionan adecuadamente y permiten realizar el entonamiento de los lazos de control de forma organizada con pocos intentos de ajustes, lo cual resulta en menos perturbaciones en el proceso durante las sesiones de sintonización. El método de síntesis directa resulta ser un método bastante optimizado para procesos autorregulados, permitiendo respuestas más estables tanto de la variable del proceso como de la salida del controlador, lo cual permite conservar la integridad del elemento final de control.

En el caso del proceso de regeneración catalítica resulta más conveniente sintonizar el controlador empleando el método de síntesis directa ya que permite un control bastante fino y estable de la temperatura de la capa. Una ventaja resaltante del método organizado aproximado de ensayo y error es que por ser un método en lazo cerrado, no es necesario realizar la prueba escalón sino que puede aplicarse dicho método tomando los datos del curso normal de operación llevada a cabo por el operador.

El éxito en la aplicación de ambos métodos dependerá en gran manera del uso de herramientas de recolección de datos que permitan analizar el comportamiento de las tendencias del proceso. Si no se dispone de estas herramientas, como alternativas se puede utilizar

el entonamiento por ensayo y error basado en la experiencia o usar el Auto-Tuning y afinar manualmente los parámetros del controlador en caso de ser requerido. Además se recomienda implementar una estrategia de control en cascada para minimizar las perturbaciones por causa de las variaciones de temperatura del aire de entrada.

8. Referencias

- [1] C. Smith and A. Corripio, *Control Automático de Procesos Teoría y Práctica*. Editorial Limusa, Grupo Noriega Editores, 1994, México.
- [2] K. Starr, *Single loop control methods*. Westerville, Ohio: ABB Inc, 2015.
- [3] P. Chau, "Process control a first course with Matlab," in *Cambridge Series in Chemical Engineering*. Cambridge University Press, 2002.
- [4] H. Wade, "Trial and error: An organized procedure," *InTech Magazine*, vol. 52, no. 5, pp. 38–42, 2005.
- [5] A. Aguado Behar and M. Martínez Iranzo, *Identificación y Control Adaptativo*. Madrid: Prentice Hall, Pearson Educación, S.A., 2003.
- [6] T. Marlin, *Process Control Designing Processes and Control Systems for Dynamic Performance*. New York: McGraw-Hill, Inc., 2000.
- [7] V. Alfaro, "Identificación de procesos sobreamortiguados utilizando técnicas de lazo abierto," *Ingeniería*, vol. 11, no. 2, pp. 11–25, 2001.
- [8] V. Alfaro, "Métodos de sintonización de controladores PID que operan como reguladores," *Ingeniería*, vol. 12, no. 2, pp. 21–36, 2002.
- [9] V. Alfaro, "Identificación de modelos de orden reducido a partir de la curva de reacción del proceso," *Ciencia y Tecnología*, vol. 24, no. 2, pp. 197–216, 2006.
- [10] S. Whan-Sung, J. Lee, and I. Lee, *Process Identification and PID Control*. Singapore: IEEE Press, John Wiley & Sons, 2009.
- [11] A. Creus, *Instrumentación Industrial*. Marcombo Ediciones Técnicas S.A., 2005.
- [12] R. Perry and D. Green, *Perry's Chemical Engineers' Handbook*. New York: McGraw-Hill Inc., 1984.
- [13] H. Wade, *Basic and Advanced Regulatory Control: System Design and Application*. International Society of Automation (ISA), 2017.
- [14] F. Shinskey, *Process Control System Application, Design and Tuning*. New York: McGraw-Hill, Inc., 1988.
- [15] T. Weber, *An Introduction to Process Dynamics and Control*. New York: John Wiley & Sons, 1973.
- [16] K. Ogata, *Modern Control Engineering*. New Jersey: Pearson Education, Inc., Prentice Hall, 2010.
- [17] C. Grimholt and S. Skogestad, "Optimal PI and PID control of first-order plus delay processes and evaluation of the original and improved SIMC rules," *Journal of Process Control*, vol. 70, pp. 36–46, 2018.
- [18] B. Vargas, "Efectos del retardo en el control de lazo cerrado de plantas sobreamortiguadas," *Revista de la Facultad de Ingeniería Industrial*, vol. 15, no. 1, pp. 110–119, 2012.
- [19] D. Chen and D. Seborg, "PI/PID Controller Design Based on Direct Synthesis and Disturbance Rejection," *Industrial and Engineering Chemistry Research*, vol. 41, pp. 4807–4822, 2002.
- [20] C. Cornell, *Control Systems Engineer Technical Reference Handbook*. USA: International Society of Automation (ISA), 2012.
- [21] N. Battikha, *Condensed Handbook of Measurement and Control*. USA: International Society of Automation (ISA), 2007.
- [22] H. Wade, "Aplicación de un procedimiento para sintonizar controladores por el método de prueba y error," *Revista Industria al día*, vol. 76, 2009.

Índice de materias

Symbols

K-means, 180

A

agrupamiento basado en vectores de soporte, 180

Análisis fractal, 4

análisis termodinámico, 111

Aprendizaje Profundo, 152

aprendizaje profundo, 141

autoencoders, 180

C

calidad del agua, 47

carburo de boro, 111

carga, 83

coeficientes de concentración y acumulación, 83

contaminantes, 83

control por realimentación, 193

corriente de armadura, 165

D

degradación, 83

diseño, 130

ductilidad, 121

E

EEG, 141

ENSO, 59

entorno natural, 47

esfuerzo máximo, 121

estadísticas de orden superior, 141

estatus ambiental, 47

Extracción de Características, 152

F

fallas, 165

flujo de escombros, 35

funciones cópula, 59

G

generador sincrónico, 165

grafito, 23

I

IBER, 35

identificación de modelo, 193

indicador acumulativo de contaminación, 83

L

líquido radioactivo, 23

M

máquinas de vectores de soporte, 165

mantenimiento predictivo, 165

material compuesto, 111

medio urbano, 83

microestructura, 121

modelamiento de flujos, 35

modelo probabilístico, 59

monitoreo, 47

O

objeto de actividades, 47

P

PM_{2,5}, 4

Procesamiento Natural de Lenguaje, 152

Pronósticos, 4

propiedades geoquímicas de suelos, 69

pulsos electromagnéticos de nanosegundos, 23

R

radiación, 23

RAMMS, 35

Raspberry Pi, 152

recocido, 121

red neuronal convolucional pre-entrenada Inception,
141

Redes Neuronales, 152

regeneración catalítica, 193

rodamientos, 165

rotura de presas, 35

S

síntesis autopropagante a temperaturas elevadas, 111

Scilab-XCOS, 193

Señales EEG epilépticas, 180

sequía agrícola, 59

Series de tiempo, 4

servicios auxiliares, 130

sintonización de PID, 193

sistema, 47

sistemas de respaldo, 130
subestaciones, 130
suelos subtropicos de Rusia, 69

T

tenacidad, 121
transformación en entorno urbano, 69

Z

zonas funcionales, 83

Subjects index

Symbols

K-means, 180

A

agricultural drought, 59

annealing, 121

armature current, 165

Artificial Neural Network, 152

autoencoder, 180

auxiliary services, 130

B

backup systems, 130

bearings, 165

BHFP determination technique, 97

boron carbide, 111

bottom-hole flowing pressure, 97

C

catalytic regeneration, 193

composite, 111

concentration and accumulation factors, 83

copula functions, 59

D

dam breaking, 35

debris flow, 35

Deep Learning, 152

deep learning, 141

degradation, 83

design, 130

E

EEG, 141

elongation, 121

ENSO, 59

environmental status, 47

Epileptic EEG signals, 180

F

failures, 165

Features Extraction, 152

feedback control, 193

flow modeling, 35

Forecasts, 4

Fractal Analysis, 4

functional areas, 83

G

geochemical properties of soils, 69

graphite, 23

H

higher order statistics, 141

I

IBER, 35

L

load, 83

M

microstructure, 121

model identification, 193

modelado multinivel, 97

modelo estadístico multivariable, 97

monitoring, 47

multilevel modeling, 97

multivariate statistical model, 97

N

nanosecond electromagnetic pulses, 23

natural environment, 47

Natural Language Processing, 152

O

object of activities, 47

oil production well, 97

P

PID tuning, 193

PM_{2,5}, 4

pollutants, 83

pozo de producción, 97

pre-trained convolutional neural network Inception,
141

predictive maintenance, 165

presión de flujo del fondo del pozo, 97

probabilistic model, 59

R

radiation, 23

radioactive liquid, 23

RAMMS, 35
Raspberry Pi, 152
regression analysis, 97

S

Scilab-XCOS, 193
self-propagating high-temperature synthesis, 111
soil of subtropics of Russia, 69
substations, 130
support vector machines (SVM), 165
SVC, 180
synchronous generator, 165
system, 47

T

técnicas de determinación de BHFP, 97
thermal dynamic analysis, 111
Time Series, 4
total pollution index, 83
toughness, 121
transformation in urban conditions, 69

U

ultimate tensile stress, 121
urban environment, 83

W

water quality, 47

Índice de autores (*Author index*)

A

Abdulkarimova, Roza G., 111
 Abreu, Santiago, 130
 Aguilar-Fernández, Mario, 4
 Aguilera, Krishna, 121
 Arteaga-Bravo, Francisco, 193
 Ataya, Juan C., 130
 Azarov, Valery N., 47

B

Barrientos-Ramírez, Sergio, 4
 Batkal, Aisulu N., 111
 Bondarenko, Vladimir L., 47

C

Cabrera, Juan, 59
 Chourio, Rafael, 152

D

Díaz-Salas, Abelardo M., 35
 Donoso, Caterine, 121
 Dubovitskaya, Tatyana V., 83

F

Fomenko, Sergey M., 111

G

Galkin, Sergey Vl., 97
 Galkin, Vladislav Ign., 97
 García-Jarquín, Brenda, 4
 Guerra, Ledy, 165
 Guevara, Yonder, 130
 Guevara-Pérez, Edilberto, 35, 59

H

Heredia, Reinaldo, 121

K

Kaverina, Natalia V., 83
 Khetsuriani, Elguja D., 47
 Kondaurov, Roman A., 83
 Krymsky, Valerii V., 23
 Kurolap, Semen A., 83

L

López-Sánchez, Víctor M., 4
 Lanza, Gerardo, 193

M

Malyukova, Lyudmila S., 69
 Mejía, Jesús, 59
 Merlín-Rodríguez, Arturo, 4
 Monagas, Eva, 130
 Montilla, Guillermo, 141, 180
 Montilla, Ricardo, 141

N

Nazarenko, Natalya N., 83

O

Otero, Antonio, 130

P

Pérez, Egilda, 141, 180
 Plotnikova, Natalya V., 23
 Ponomareva, Inna N., 97

R

Ramírez, Jesús, 165
 Rastegaev, Alexander V., 97
 Romero, Trino, 121
 Ruiz, César Rodolfo†, 130

S

Sáenz, Laura, 121
 Sanz, Wilmer, 152
 Seidualiyeva, Aizhan J., 111
 Seijas, César, 141, 180
 Semenova, Elena A., 47
 Singaña, Marcos, 121
 Sparano, Aniello, 165

T

Terán, Rubén, 165
 Tolendiuly, Sanat, 111

V

Vidal-Moreno, James D., 35
 Villazana, Sergio, 141, 180
 Vinogradov, Pavel M., 83

Y

Yliasov, Allaberdy I., 47

Z

Zakharikhina, Lalita V., 69

Apéndices

Línea editorial

Misión científica

Revista Ingeniería UC es el principal organismo de difusión científica y tecnológica de la Facultad de Ingeniería de la Universidad de Carabobo.

El objetivo de Revista Ingeniería UC es la difusión de trabajos científicos tecnológicos en todas las áreas de la ingeniería, así como también ciencias afines aplicadas a la ingeniería.

Revista Ingeniería UC está adscrita a la Dirección de Investigación de la Facultad de Ingeniería. Es considerada por todos sus profesores como uno de los patrimonios más importantes de la Facultad de Ingeniería. Su primer volumen aparece en diciembre de 1992. Fundada por el Dr. Edilberto Guevara en 1992 (actualmente Editor Honorario).

Todos los artículos son revisados por el Comité Editorial y arbitrados, a doble ciego, por el Comité Técnico y por especialistas en la materia.

Revista Ingeniería UC es publicada cada cuatro meses.

Cobertura temática

Revista Ingeniería UC considerará para su difusión trabajos *originales e inéditos* (Los autores deben enviar la carta de compromiso), en español e inglés. Manuscritos en las áreas de la ingeniería de todas sus especialidades, pero en particular en las áreas de industrial, eléctrica, mecánica, civil, química y telecomunicaciones; así como de las ciencias básicas aplicadas a la ingeniería: matemáticas, física, química, computación, biología, ciencias ambientales, bioingeniería, biotecnología, estadística, etc.

Foro de Revista Ingeniería UC

La audiencia a la cual está dirigida consta de todos los investigadores en las áreas de ciencia, tecnología, ingeniería y matemáticas. En particular a los investigadores, profesores y estudiantes de Doctorado, Maestría y pregrado en las áreas de ingeniería y ciencias básicas afines. Esta audiencia es extendida a los gerentes de las áreas de innovación y desarrollo tecnológico del sector industrial. El foro cubre las comunidades a nivel local, Universidad de Carabobo y sus estados de influencia; nivel nacional, toda Venezuela; así como a nivel internacional.

Editorial line

Scientific mission

Revista Ingeniería UC is the main body for scientific and technological diffusion at Engineering Faculty of Carabobo University.

The Revista Ingeniería UC objective is the diffusion of scientific – technological works in all areas of engineering, as well as related sciences applied to engineering.

Revista Ingeniería UC is subscribed to Research Direction of Engineering Faculty. It is considered by all its professors as one of the most important patrimonies of the Faculty of Engineering. Its first volume appears in December of 1992. Founded by Dr. Edilberto Guevara in 1992 (currently Honorary Editor).

Revista Ingeniería UC is published every four months.

Thematic coverage

Revista Ingeniería UC will consider for its diffusion original and unpublished works (The authors must send the commitment letter), in Spanish and English. Manuscripts in engineering areas of all its specialties, but particularly in the areas of industrial, electrical, mechanical, civil, chemical and telecommunications; as well as the basic sciences applied to engineering: mathematics, physics, chemistry, computation, biology, environmental sciences, bioengineering, biotechnology, statistics, etc.

Revista Ingeniería UC forum

The target audience is made up of all researchers in the areas of science, technology, engineering and mathematics. In particular the researchers, professors and students of Doctorate, Master's and undergraduate in the areas of engineering and related basic sciences. This audience is extended to innovation and technological development areas managers of the industrial sector. This forum has covers local communities, University of Carabobo and their state of influence; national level, all of Venezuela; as well as international level.

All articles are reviewed by Editorial Committee and refereed, double-blinded, by Technical Committee and by specialists in the field.

Políticas de ética y publicación

Introducción

Revista Ingeniería UC, como publicación científica cuatrimestral incluye tres números (Abril-Agosto-Diciembre) por cada volumen en un año, alcanza la generación continua de contenidos como principal organismo de difusión académica, científica y tecnológica de la Facultad de Ingeniería de la Universidad de Carabobo. El idioma principal de la revista es el español, manejando el inglés como segundo idioma y todos los artículos incluyen el título, resumen y las palabras clave en ambos.

Comprometidos en mantener los más altos estándares éticos, se efectúan acciones en contra de cualquier práctica anti-ética o por negligencia profesional. El plagio está estrictamente prohibido y nuestros autores y demás investigadores expresan formalmente que sus trabajos son originales, inéditos, y no están siendo sometidos a proceso de arbitraje en otras revistas, por otra parte, queda expresamente prohibida la copia o el plagio, parcial o en su totalidad, de otras obras. Todo trabajo para ser publicado cuenta con la exigencia de manera explícita a sus autores de la normativa ética de la Revista que se plasma en una *Carta de Compromiso*.

Prácticas anti-éticas y negligencia profesional

Por considerar tanto aspectos científicos como tecnológicos, se debe tomar en cuenta la necesidad de velar por la veracidad de toda información suministrada incluyendo datos técnicos, diagramas, tablas y figuras, diseños experimentales, metodologías, entre otros, siendo de vital importancia la cita a las referencias respectivas indicando con los estándares habituales cuando la información es tomada textualmente. Se consideran como faltas graves que conllevan un proceso disciplinario sancionatorio las siguientes:

- a) **Copias o plagio:** presentar datos o ideas como propios, de manera parcial o en su totalidad, provenientes de otras obras.
- b) **Falsificación o adulteración:** la modificación y cambio de datos presentados, omitir información de forma tal que la investigación no esté completamente presentada.

- c) **Presentación de datos no verídicos:** fabricación, manipulación y presentación de información que no se desprende de datos reales derivados de la investigación.
- d) **Segmentación de la investigación:** implica la fragmentación de un estudio en diversas partes con la intención de generar más de un producto con el mismo trabajo para su publicación, incluso en diferentes revistas, sin reconocer la fuente original.
- e) **Alteración de autor y coautores:** la inclusión de algún miembro que no ha participado en la investigación, así como la exclusión de alguno que si lo ha realizado. De igual forma, cualquier modificación relacionada con cambios en autores y coautores, debe estar debidamente justificada; luego de finalizado el proceso de edición no será considerada.
- f) **Alteración del proceso de arbitraje:** cualquier intento que afecte el proceso de arbitraje por pares a doble ciego. Se prohíbe de manera explícita la comunicación entre un árbitro y los autores del artículo.

Proceso de revisión y arbitraje

Los autores serán guiados a través de cuatro fases que conllevan la revisión preliminar, el proceso de arbitraje por pares doble ciego por árbitros externos, el proceso de edición y montaje y por último, corrección de redacción y estilo. El contenido completo de todos los manuscritos será sometido a un proceso de revisión objetiva por parte de expertos en el área, quienes siguiendo para ello las pautas del Formato de Evaluación comunicarán los resultados obtenidos, permitiendo al Editor-Jefe conocer las recomendaciones formuladas y el veredicto respectivo. Además, el proceso de revisión es de carácter estrictamente confidencial, y los revisores no deben tener conflicto de intereses.

La revisión de los manuscritos permitirá evaluar la calidad técnica de estos en base a su originalidad y presencia de avances significativos en el campo de la ingeniería y ciencias afines, atendiendo a la

pertinencia del contenido, el enfoque metodológico, los hallazgos obtenidos y a las referencias que, con respecto a contribuciones anteriores, se presentan. Tomando como premisa la debida organización de la estructura del manuscrito y el cumplimiento de las normas de presentación.

De acuerdo a los resultados de la revisión, se propone al Editor-Jefe:

- Publicación, sin cambios.
- Publicación, después de cambios menores, quedando comprometidos los autores a realizar los ajustes respectivos de acuerdo a la evaluación.
- Publicación, después de cambios mayores, quedando comprometidos los autores a realizar los ajustes y modificaciones requeridas según los resultados de la evaluación, sujeto a una nueva revisión.
- Publicación de una versión corta del manuscrito como Nota Técnica, en cuyo caso los autores serán notificados y ejecutarán las modificaciones a que tuviera lugar.
- Publicación en otra revista, en cuyo caso se realizará la recomendación a los autores.
- Rechazo del manuscrito para su publicación, en cuyo caso no se recibirán nuevas versiones.

Una vez concluida la revisión, el manuscrito se someterá al proceso de edición, en donde con la intención de asegurar la calidad de la publicación se podrá solicitar la condensación, supresión o incorporación de tablas, figuras, y ecuaciones, teniendo en cuenta que el comité editorial se reserva el derecho de efectuar las modificaciones pertinentes sin afectar el contenido, así como también la potestad de decidir respecto a su publicación si los autores no responden satisfactoriamente a las observaciones planteadas. El manuscrito editado, será remitido a los autores para su aprobación.

De la responsabilidad de los autores

Los autores que envían sus manuscritos para su publicación en Revista Ingeniería UC, deben someterlos a un proceso de revisión y arbitraje por pares doble ciego, y están obligados a:

- Firmar la *Carta de Compromiso*, manifestando que el contenido del trabajo enviado a la revista es original e inédito: no ha sido publicado, no es duplicado ni redundante, además no está siendo sometido simultáneamente a arbitraje para su publicación por ningún otro medio de difusión, y por último que los datos presentados son originales y verídicos. Adicionalmente, el autor y los coautores ceden los derechos patrimoniales a la Revista Ingeniería UC, pero mantienen sus derechos como autores intelectuales.
- Deben evitarse o reconocerse los envíos anteriores del manuscrito a otras revistas o publicaciones, en el caso de haber sido enviado es necesario la presentación de una comunicación de la otra publicación donde se deje sin efecto el proceso de arbitraje y las motivaciones consideradas.
- Respecto a materiales/datos citados o utilizados de otras investigaciones deben documentarse cuidadosamente y citarse como referencia e indicar cuando son tomados textualmente. No es permitida la duplicación de trabajos anteriores.
- Especificar de manera clara la información relativa a las fuentes de financiamiento para el desarrollo de la investigación.
- Responder de manera oportuna a las comunicaciones que tenga lugar con el Comité Editorial.
- Cuando un error sea detectado, el autor está obligado a retractarse y efectuar las correcciones pertinentes.
- Respecto a la identificación de autores, se debe indicar el nombre tanto del autor como de los coautores, que hayan contribuido significativamente en la investigación y que estén involucrados con esta, evitando incluir personas que no estén vinculadas a la misma. Siendo además miembros de una institución de educación superior o de una estructura de investigación reconocida, suministrando información detallada relacionada con su filiación, correo electrónico de contacto y cualquier otro elemento que permita su identificación.

De la responsabilidad de los revisores-árbitros

Una vez recibido el manuscrito, acompañado de las Normas de Publicación y el Formato de Evaluación

previsto por Revista Ingeniería UC, los árbitros están en la obligación de informar de manera expedita, si se acepta o rechaza la revisión del mismo explicando los motivos. En el caso de que los árbitros tengan interés intelectual o material relacionado al manuscrito bajo revisión, estos deberán inhibirse de manera voluntaria evitando así Conflicto de Intereses.

Los árbitros deberán rechazar la revisión de manuscritos, cuando estos estén fuera de su área de experticia o de especialización técnica y científica. Por otro lado, los árbitros o revisores deben completar debidamente el Formato de Evaluación previsto así como también realizar las observaciones que considere pertinentes sobre el manuscrito evaluado, lo cual incluye todas aquellas propuestas tendentes a mejorar cuando sea el caso aspectos metodológicos, la presentación y análisis de resultados y la incorporación de investigaciones de referentes de importancia que hayan sido omitidos por los autores o que debieran ser tomados en consideración, tomando en cuenta que el envío de dicha información al Editor-Jefe es de carácter obligatorio y estrictamente confidencial.

De la responsabilidad de los miembros del Comité Editorial

El Editor-Jefe de Revista Ingeniería UC, realizará una revisión preliminar al recibir el manuscrito, atendiendo a las características de la línea editorial correspondiente. Una vez superada dicha revisión, se procederá a proponer a los revisores o árbitros de acuerdo al área de conocimiento.

El Editor-Jefe será el responsable de enviar las comunicaciones respectivas a los autores, con respecto a las observaciones emitidas por los árbitros o revisores, y conducir el proceso para que se tomen medidas correctivas. Así mismo, decidirá sobre la aceptación o no del manuscrito en atención a los comentarios de los árbitros o revisores designados. El Comité Editorial tratará la información relacionada con los artículos con criterios de estricta confidencialidad.

El Comité Editorial velará por el cumplimiento de las prácticas de ética. Además realiza un esfuerzo por identificar e impedir la publicación de artículos que involucren malas conductas durante la investigación. De igual manera, se debe tratar cualquier tipo de acusación o señalamiento de manera responsable y dar inicio al proceso de sustanciación del expediente con las averiguaciones a las que se tenga lugar y se decidirá si es necesario un proceso disciplinario que conlleva:

- Amonestación: se envía a los autores una carta con las conclusiones del proceso disciplinario.
- Artículo retractado: se incluirá una carta de exposición de motivos con las conclusiones del proceso disciplinario, y el artículo será marcado con una nota de retractación.
- Suspensión: se le notifica a los autores que no serán recibidos manuscritos para su publicación mientras dure la sanción correspondiente.

El Comité Editorial es responsable de la publicación de cualquier nota, crítica razonable, fe de errata o disculpas con respecto a los manuscritos publicados en números anteriores.

Preservación digital

Revista Ingeniería UC utiliza para la preservación digital de sus artículos, el resguardo con archivo comprimido en formato L^AT_EX, además a través de Marcalyc 2.0 realiza el marcaje obteniendo archivo XML según estándar JATS.

Licencia Creative Commons

Licencia de Creative Commons Reconocimiento – No Comercial – Sin Obras derivadas 4.0 Internacional



Ethics and Publication Policies

Introduction

Revista Ingeniería UC, as a quarterly scientific publication includes three issues (April-August - December) for each volume in a year, reaches the continuous generation of contents as principal media of diffusion academic, scientific and technological of the Faculty of Engineering at Carabobo University. The primary language of the journal is Spanish, managing English as a second language, all articles include title, abstract and keywords in both.

Also, committed to maintain the highest ethical standards, are performed actions against any anti-ethics practice or professional negligence. Plagiarism is strictly forbidden and our authors and other researchers must express formally that their manuscripts are original, unpublished, and not have been submitted to arbitration process in other journals simultaneously. In the other hand, is expressly forbidden the copying or plagiarism, partially or in its entirety, of other works. An article to be published has as requirement knowledge by the authors of the ethic policies, that are reflected in the signing of “*Compromise Letter*”.

Anti-ethics practices and professional negligence

Due scientific and technological aspects, Revista Ingeniería UC needs to ensure the veracity of all information provided including technical data, diagrams, tables and figures, experimental designs, methodologies, among others, being vital the cite of the respective references indicating with the usual standards when the information is taken verbatim. Considering as serious faults involving a disciplinary process:

- a) **Copy or plagiarism:** present data or ideas as own, partially or entirety, from other works.
- b) **Falsification or adulteration:** modification or change presented data. Omitting information in such way that the investigation is not completely presented.
- c) **Presentation of untrue data:** manufacturing and handling of presented information that does not arise from real data derived from research.
- d) **Research segmentation:** fragmentation of a study in different parts with the intention of generating more than one product with the same work for publication, even in different journals, without recognizing original source.
- e) **Alteration of author and coauthors:** Exclusion of members participating in the investigation or inclusion of members without a direct role. Also, any alteration of author-coauthors, must be justified and will not be considered after has been closed edition process.
- f) **Alteration of arbitration process:** any attempt that affects the arbitration process (double blind peer-review). Communication between an arbitrator and authors of manuscript, is strictly prohibited.

Review process and arbitration

The authors will be guided through four phases including a preliminary review, arbitration process (peer-review, double blind by external reviewers), final edition and assembly and correction of style and redaction. All the content will be sent to an objective reviewing process by experts in knowledge field, who following the guidelines of the Evaluation Format will communicate the results obtained, allowing the Editor-Chief to know the recommendations made and the respective verdict. In addition, the review process is strictly confidential, and the reviewers should not have interest conflict.

The revision of the manuscripts will allow to evaluate technical quality based on originality and presence of significant advances in the field of engineering and related sciences, considering the relevance of content, methodological approach and findings obtained and references to previous contributions. Taking as a premise organization of the structure of manuscript and compliance with presentation rules.

Based on review process results, is proposed to the Editor-Chief:

- Publication, without changes.

- Publication, after minor changes, the authors are committed to make the respective adjustments according to the evaluation.
- Publication, after major changes, the authors are committed to make the adjustments and modifications required according to the evaluation results, subject to a new revision.
- Publication of a short version of the manuscript as a *emphTech Note*, in which case the authors will be notified and will execute the modifications that may take place.
- Publication in another journal, in which case the recommendation to the authors will be made.
- Rejection of the manuscript for publication, in which case new versions will not be received.

Once the review is completed, the manuscript will be elevated to the editing process, where with the intention of ensuring the quality of the publication, may be requested condensation, deletion or incorporation of tables, figures, and equations, considering that the editorial committee could make relevant modifications without affecting the content, as well as would decide about publication if the authors do not respond satisfactorily to the comments made. The edited manuscript will be submitted to the authors for approval.

About the responsibility of authors

Authors who submit their manuscripts for publication in *Revista Ingeniería UC*, are in knowledge that them will be subjected to a double-blind peer review and arbitration process, and are obliged to:

- Sign the “*Compromise Letter*”, stating that the content of manuscript sent to the journal is original and unpublished: it has not been published, it is not duplicated or redundant, it is not being simultaneously submitted to arbitration for publication by another means of diffusion, and finally that the data presented are original and true. Furthermore, the author and coauthors yield patrimonial rights to the journal, but maintain their rights as intellectual authors.
- Previous submits of the manuscript to other journals should be avoided or recognized, and it is necessary to present a communication from the

other journal where the arbitration process are left without effect and the motivations considered.

- Should be carefully documented and cited as references, all materials/data used from other research, indicating when they are taken verbatim. Duplication of previous works is not allowed.
- The authors must clearly specify the information related to sources of financing for the development of research.
- Authors should respond timely all communications of Editorial Committee.
- When an error is detected, the author is obliged to retract and make the appropriate corrections.
- Respect to the identification of authors, should be indicated the name of author and coauthors, that have significantly contributed on the research and have been involved in its development, avoiding to include persons who are not linked to it. The authors and coauthors must be members of a higher education institution or a recognized research structure, providing detailed information related to their filiation, contact email and any other element that allows their identification.

About the responsibility of reviewers – arbitrators

Once upon the manuscript is received, with the Normative of Publication and Evaluation Format of *Revista Ingeniería UC*, reviewers are obligated to inform immediately, if accept or refuse review the document and must explain the motives about it. In the event that arbitrators have intellectual or material interest related to the manuscript under review, must be voluntarily inhibited them, thus avoiding Conflict of Interest.

The reviewers must reject the revision of manuscripts, when they are outside their area of expertise or technical and scientific specialization. In other hand, the arbitrators or reviewers must complete the expected Evaluation Form, as well as make the observations that they consider pertinent about the manuscript evaluated, which includes, as the case may be, all those proposals aimed at improving the presentation and analysis of results, and the incorporation of investigations of relevant referents, taking into account that the sending of information to the Editor-Chief is mandatory and strictly confidential.

About the responsibility of Editorial Committee members

The Editor–Chief of Revista Ingeniería UC, will carry out a preliminary review upon receiving the manuscript, considering the characteristics of the corresponding editorial line. Once this revision has been completed, it will proceed to propose the reviewers or arbitrators according to the area of knowledge.

The Editor–Chief will be responsible for sending the respective communications to the authors, with respect to the observations issued by the reviewers or arbitrators, and conduct the process to corrective measures are taken. Likewise, it will decide on the acceptance or not of the manuscript in response to the comments of the appointed arbitrators or reviewers. The Editorial Committee will manage the information related to the articles with confidentiality criteria.

The Editorial Committee will ensure compliance with ethical practices. In addition, the editorial committee makes an effort to identify and prevent the publication of articles that involve misconduct during the investigation. At the same time, any type of accusation must be treated responsibly, initiate the inquiries for the substantiation file, through which it will be decided if a disciplinary process is necessary, that implicates:

- Admonish: a letter with the conclusions of the disciplinary process is sent to the authors.
- Retracted article: it will include a letter of explanation of the reasons with the conclusions of the disciplinary process, and the article will be marked with a retraction note.
- Suspension: the authors are notified that manuscripts will not be received for publication during the corresponding sanction.

The Editorial Committee is responsible for the publication of any note, reasonable criticism, erratum or apology regarding the manuscripts published in previous issues.

Digital preservation

Revista Ingeniería UC uses to digital preservation of its articles, saving zip file with L^AT_EX format, also via Marcalyc 2.0 makes marked file generating XML JATS file.

License Creative Commons

License de Creative Commons Attribution – No Commercial – No derivatives. 4.0 International



Normas para la presentación de artículos

Generales

Línea editorial

Revista Ingeniería UC, considerará para su difusión trabajos *originales e inéditos* en ingeniería, en particular las áreas de industrial, eléctrica, mecánica, civil, química y telecomunicaciones; así como de las ciencias afines.

Tipos de trabajos

- a Artículos de investigación inéditos con un máximo de veinte (20) páginas.
- b Notas técnicas con un máximo de cinco (5) páginas.
- c Artículos de actualización científica que resuman el Estado del Arte de un área específica de la ingeniería con un máximo de doce (12) páginas.
- d Artículo invitados especiales con un máximo de veinte (20) páginas.
- e Cartas al editor.
- f Artículos de tendencia tecnológica, estudios de prospectiva de I+D+i a medio y largo plazo de alta utilidad en el campo industrial.

Al enviar un artículo a Revista Ingeniería UC, el autor está comprometido formalmente con que el trabajo consignado es original e inédito, de igual manera manifiesta su conocimiento de las normas de la revista y acepta que sea sometido al proceso de arbitraje.

Estilo

La redacción de los trabajos puede realizarse en idioma castellano o inglés. El trabajo original debe ser redactado en formato \LaTeX o en su defecto en Microsoft Word.

Los trabajos en Word deben emplear una fuente de la familia Times, estar almacenado en disco compacto (CD) o en su defecto enviarse por vía electrónica al correo de la revista revistaing@uc.edu.ve.

Los trabajos en \LaTeX deben incluir los archivos pdf, tex y una carpeta comprimida con las figuras en eps (o png)

Figuras

Las figuras pueden ser en escalas de grises claros o a color, en formato vectorializado preferiblemente png o eps (también se admiten los formatos CompuServe gif o jpg, si estas son de alta calidad y trabajó en word) deben añadirse en archivos independientes y numeradas. Las leyendas o descripciones de la figuras no pueden estar embutidas en éstas, deben ser incluidas en el texto del trabajo.

Tablas:

Las tablas no pueden ser resaltadas por ningún tipo de color. Solamente los textos a resaltar mediante “negritas”.

Toda tabla y/o figura deben ser numerada en arábigo (1, 2, ...), citada y suficientemente comentada en el texto del trabajo. La cita sería en la forma: “en la Figura 3”, “en la Tabla 2”, por ejemplo (sin abreviar). Toda tabla o figura debe tener un ancho máximo de 17,5 cm.

No es estilo de la revista frases como: “en la Figura siguiente” o en la “Tabla anterior”

El orden a seguir para la redacción del trabajo es el siguiente:

1. Portada.
2. Introducción.
3. Metodología o Desarrollo de la investigación.
4. Análisis y discusión de resultados.
5. Conclusiones.
6. Referencias.

La portada debe contener:

- Título del trabajo en castellano y en inglés, con un máximo de 20 palabras.

- Nombre(s) del autor(es) y su dirección(es) institucionales completa(s), dirección postal, correo electrónico del autor para correspondencia (sólo será publicado el correo electrónico del autor para correspondencia).
- Resumen del trabajo en castellano y en inglés (Abstract) con una extensión máxima de 200 palabras.
- Al final tanto del resumen como del abstract debe agregarse entre tres (3) y cinco (5) palabras clave.

Los encabezamientos de cada sección se escriben tipo título, mayúsculas sólo en la primera letra.

citas

Las citas de referencias en el texto, contendrán el nombre del autor principal seguido de corchetes con el número correspondiente a la referencia, por ejemplo: Engelbrecht [11], o simplemente el número de la referencia bibliográfica [11], sin citar al autor.

Referencias

Para los artículos en L^AT_EX se usará el estilo de bibliografía “IEEEtran”.

Las referencias bibliográficas se escribirán en orden de citación, deben ser completas y contener todos y cada uno de los datos para identificarla. Se debe incluir por cada referencia la información bibliográfica (DOI, URL, o captura de pantalla) que permita su verificación.

Se pueden citar:

- Artículos de revistas.
- Artículos de memorias de congresos.
- Artículos en colecciones.
- Capítulos de libro.
- Memorias de congresos.
- Libros
- Tesis doctorales
- Trabajos especiales de grado.
- Informes técnicos.

- Manuales técnicos.

No se permiten direcciones electrónicas ni páginas web.

En el caso de artículos de revista contendrán:

- Autor(es) (Inicial del nombre y Apellido).
- “Título”.
- *Nombre de la revista*,
- volumen,
- número,
- páginas y
- año de publicación

Por ejemplo:

- [1] F. Chen, Z. Ji, and Q. Qi, “Effect of pore size and layers on filtration performance of coalescing filters with different wettabilities,” *Separation and Purification Technology*, vol. 201, no. 7, pp. 71–78, 2018.

En caso de libros ha de incluirse: Autor(es), *título*, lugar de publicación : editorial, año de publicación.

Por ejemplo:

- [2] P. Mangonon, *Ciencia de materiales selección y diseño*. Atlacomulco, México: Pearson Educación México, 2001.

En caso de colección editada: Autor(es), “Título”, en *título*, Editor(es), año, Eds., editorial, año de publicación, volumen, capítulo, número de páginas.

Por ejemplo:

- [3] T. Thanapalasingam, F. Osborne, A. Birukou, and E. Motta, “The Semantic Web – ISWC 2018,” in *17th International Semantic Web Conference, Monterey, CA, USA, October 8–12, 2018, Proceedings, Part II*, D. Vrandečić, K. Bontcheva, M. C. Suárez-Figueroa, V. PresuttiIrene, I. Celino, M. Sabou, L.-A. Kaffee, and E. Simperl, Eds. Springer, 2018, vol. 11137 LNCS, ch. Ontology-based recommendation of editorial products, pp. 341–358.
- [4] D. Nilesh and M. Nagle, “The new cryptography algorithm with high throughput,” in *2014 International Conference on Computer Communication and Informatics*, Coimbatore, 2014, pp. 1–5.

Se recomienda a los autores tener en cuenta las normas internacionales de nomenclatura para la utilización de símbolos, unidades y abreviaturas. Adicionalmente, el uso de coma (,) como separador decimal

Notas Finales

Los artículos serán sometidos al proceso de arbitraje, doble ciego por pares por arbitros externos, previo a su publicación. La Revista Ingeniería UC como órgano de difusión y divulgación de las actividades científicas de la Facultad de Ingeniería de la Universidad de Carabobo siempre se ha pronunciado por un Acceso Abierto No comercial (AANC), siendo una institución sin fines de lucro.

Authors Information

General

Editorial line

Revista INGENIERÍA UC will consider for publication original and unpublished contributions in engineering, in particular areas of industrial, electrical, mechanical, civil, chemical and telecommunications, and applied science.

Type of Manuscripts:

- a Unpublished research articles with a maximum length of twenty (20) pages.
- b Technical Notes, with a maximum of five (5) pages.
- c State of the Art Articles (Update Articles) of a specific engineering field (maximum of twenty (20) pages).
- d Articles sent by invited guests, maximum of twenty (20) pages.
- e Letter to Editor.
- f Technological trend articles, medium and long-term R & D + i prospective studies of high utility in the industrial field.

Submitting an article to Revista Ingeniería UC, the author is formally committed to the consigned paper is original and unpublished, he manifest his knowledge of Revista Ingeniería UC standards and that his work will be refereed.

Article style

Articles may be sent either in Spanish or English. The original work should be written using \LaTeX or in Microsoft WORD.

The Microsoft WORD works must employ font of Times family, to be stored in compact disk (CD) or send electronically by email to revistaing@uc.edu.ve.

The works in \LaTeX must include the files pdf, tex and a folder with the figure in png, eps or pdf.

The figures can be in gray scale, or color, preferable in format png or eps (also admit the formats CompuServe gif or jpg if these are of high quality) must to be added in records separated and numbered. Figures caption or descriptions can not be into these, must be includes in the text of the work and in a apart file called Figure Caption.

The table can not be protruded by any type of color. Only the texts to be produced by means of "bold type". Also must to be included in others files numerated and it is necessary to include a file with table caption.

All table and/or figure must be numbered in arabic (1, 2, 3, . . .) and cited in the text of the work. The quote would be in the manner: "In Figure 3", "in Table 2", for example (without abbreviating). All board or figure must have a wide maximum of 17,5 cm.

The writing sequence for the articles shall be as following:

1. Title page.
2. Introduction.
3. Methodology or Research development.
4. Analysis and Discussion of results.
5. Conclusions.
6. References.

The Title page should contain:

- Title of the work in Spanish and English, maximum 20 words.
- Full name(s) of author(s) with Full Institutional Address(es) (Address, Telephone, Email).
- Summary of the work in Spanish and in English (Abstract) with a maximum of 200 words for Unpublished, State of the Art articles and Technical Notes.
- At the end of the Summary in Spanish and Abstract in English three to five (3 to 5) Key Words should be added.

Text references must contain the name of the main author followed by the corresponding number of reference; for example: Engelbrecht [11], or simply [11] with no author name.

For L^AT_EX work will use bibliography style “IEEE-tran”.

References will be written in order of citation, complete and should contain: author(s) (in caps and lower case), “full title”, *journal*, volume, number, pages and year of publication. For each reference, bibliographic information is required to checking (DOI, URL, or screenshot). For example:

- [1] F. Chen, Z. Ji, and Q. Qi, “Effect of pore size and layers on filtration performance of coalescing filters with different wettabilities,” *Separation and Purification Technology*, vol. 201, no. 7, pp. 71–78, 2018.

In the case of Books: Author(s) (in caps and lower case), *title*, place: publisher and year of publication should be included. For example:

- [2] P. Mangonon, *Ciencia de materiales selección y diseño*. Atlacomulco, México: Pearson Educación México, 2001.

Work in an edited collection: Author(s) (in caps and lower case), “title,” in: *booktitle*, Editor(s) (in caps and lower case), publisher, year, volume, chapter and pages. For example:

- [3] T. Thanapalasingam, F. Osborne, A. Birukou, and E. Motta, “The Semantic Web – ISWC 2018,” in *17th International Semantic Web Conference, Monterey, CA, USA, October 8–12, 2018, Proceedings, Part II*, D. Vrandečić, K. Bontcheva, M. C. Suárez-Figueroa, V. PresuttiIrene, I. Celino, M. Sabou, L.-A. Kaffee, and E. Simperl, Eds. Springer, 2018, vol. 11137 LNCS, ch. Ontology-based recommendation of editorial products, pp. 341–358.
- [4] D. Nilesh and M. Nagle, “The new cryptography algorithm with high throughput,” in *2014 International Conference on Computer Communication and Informatics*, Coimbatore, 2014, pp. 1–5.

It is recommended to the authors to follow the International Nomenclature Norms (symbols, units and abbreviations). Additionally, use comma(,) as decimal separator.

Final Notes

Articles will be submitted for the reviewing process by external reviewers before they can be published. Revista Ingeniería UC as an organ of diffusion and divulgation of scientific activities of Engineering Faculty at Carabobo University, has always been pronounced for Open Access Non-Commercial policy, being a non-profit institution.

Carta de compromiso

Envíe junto con su trabajo la siguiente carta al Comité Editorial de Revista Ingeniería UC:

Ciudadanos
Director Editor en jefe y demás
Miembros del Comité Editorial
Revista INGENIERÍA UC
Presente.

Por medio de la presente envío a Ud.(s) el manuscrito del trabajo titulado:

para que sea sometido a evaluación para la publicación.

Manifiesto que:

Este trabajo es original e inédito: no ha sido publicado, no es duplicado, ni redundante; no está siendo sometido simultáneamente a arbitraje para su publicación por ningún medio de difusión, que los datos son originales y verídicos.

El autor y los coautores ceden los derechos de autor a la Revista INGENIERÍA UC, pero mantienen sus derechos como autor intelectual.

El trabajo, tanto en su texto como las tablas y figuras ha sido elaborado de acuerdo a las Instrucciones para los Autores publicadas por Revista INGENIERÍA UC, y que las referencias están directamente relacionadas con el trabajo.

Se designa como autor de correspondencia al autor o coautor que lo indique, con quien el Comité Editorial mantendrá comunicación a través del correo electrónico revistaing@uc.edu.ve, quien será responsable ante autores y coautores y dará respuesta rápida a los requerimientos del Comité Editorial.

No se conocen conflictos de intereses, y de haberlos los autores y coautores están obligados a indicarlo en el original, junto a la fuente de financiamiento.

Firma:

Nombre y apellido:

El autor para correspondencia:

Compromise letter

Send together with your paper the following letter to Editorial Committee:

Citizens
Director, Chief Editor and other
Members of Editorial Committee
Revista INGENIERÍA UC
Present.

I hereby send you the manuscript of the work entitled:

to be submitted for evaluation to publication.

I declare that:

This work is original and unpublished: it has not been published, it is not duplicated, nor redundant; it is not being simultaneously submitted to arbitration for publication by any diffusion means, that the data are original and truthful.

The author and coauthors assign the copyright to Revista INGENIERÍA UC, but maintains their rights as an intellectual author.

The work, in its text as the tables and figures has been elaborated according to the Authors Instructions published by Revista INGENIERÍA UC, and that the references are directly related to the work.

The author or coauthor who indicates it is designated as correspondence author, with whom Editorial Committee will maintain communication via email revistaing@uc.edu.ve, who will be responsible for the authors and coauthors, and will respond quickly to the requirements of the Editorial Committee.

There are no known interest conflicts, and if there are the authors and coauthors are obliged to indicate in the original, next to funding source.

Signature:

First name and surname:

Correspondence author:

Índices de Revista Ingeniería UC

Revista Ingeniería UC está indizada en:



Universidad de Carabobo

Autoridades

Jessy Divo de Romero

Rectora

Ulises Rojas

Vicerrector Académico

José Ángel Ferreira

Vicerrector Administrativo

Pablo Aure

Secretario

Autoridades de la Facultad de Ingeniería

Prof. Manuel Elías Jiménez Bahri

Decano

Prof^a. Carmen Guédez
Directora Esc. Ing. Industrial

Prof^a. Cristina Cordero Peraza
Asistente al Decano

Prof. Reny Marin
Director Consejo Facultad

Prof^a. Ledy Guerra
Directora Esc. Ing. Eléctrica

Prof. Carlos Brito
Director de Asuntos Estudiantiles

Prof^a. Lissette Jiménez
Directora de Postgrado

Prof^a. Vanessa Hurtado
Directora Esc. Ing. Mecánica

Prof. Edson Martínez Oberto
Director Esc. Ing. Civil

Prof. Luis E. Di Stefano
Director de Administración

Prof^a. Ixmit López
Directora Esc. Ing. Química

Prof. Carlos Cochiarella
Director de Asuntos Profesorales

Prof. Pablo Baricelli
Director del Centro de Investigaciones Químicas

Prof. Ángel D. Almarza M.
Director de Investigación y Producción Intelectual

Prof^a. María Teresa Cruz
Directora de Docencia y Desarrollo Curricular

Prof. César O. Seijas Fossi
Director del Centro de Procesamiento de Imágenes

Prof. César Cuperto Ruiz P.
Director Esc. Ing. de Telecomunicaciones

Prof^a. Neyda Severian
Directora Estudios Básicos

Prof^a. Marlin Gómez
Directora de Extensión

Prof. Eduardo Vargas
Director del Instituto de Matemática y Cálculo Aplicado

Prof^a. Lin Hurtado
Directora de Tecnologías de la Información y Comunicación

Prof. Donato Magarelli
Director de Biblioteca

REVISTA INGENIERÍA UC

La Revista Ingeniería UC es una publicación periódica de la Facultad de Ingeniería de la Universidad de Carabobo, adscrita a la Dirección de Investigación. Es arbitrada, indizada en REVENCYT (Venezuela) ; Actualidad Iberoamericana (CIT-Chile); IET/INSPEC (Unitec Kingdom UK); LATIDEX-Catálogo (México); REDALYC (México); Periódica (México) y está incluida en Ulrich's International Periodical Directory (USA), de Publicación Internacional; Emerging Sources Citation Index, Thomson Reuters; MIAR-Matriz de información para el análisis de revistas; RedIB Red Iberoamericana de Innovación y Conocimiento Científico; DOAJ Directory of Open Access Journals

Se publica cada cuatro meses Abril, Agosto, Diciembre. Se aceptan trabajos en Castellano e Inglés. Todos los trabajos son revisados por el Comité Editorial, Arbitrados por el Comité Técnico y por especialistas en la materia.

Dirección postal: Facultad de Ingeniería, Universidad de Carabobo, Bárbula-Valencia, Código Postal 2008, estado Carabobo, Venezuela.

Correo-e: revistaing@uc.edu.ve

Correo-e: revistaing@gmail.com

Página Web: <http://servicio.bc.uc.edu.ve/ingenieria/revista/>; <https://www.redalyc.org/revista.oa?id=707/>

**Revista Ingeniería UC, ISSN 1316–6832, Online ISSN: 2610-8240
Enero–Abril 2021, Volumen 28
Número 1**

Revista Ingeniería UC se encuentra indizada en:
Revencyt (Venezuela) – índice de revistas venezolanas de ciencia y tecnología;
Actualidad Iberoamericana (Chile) CIT;
IET/INSPEC (United Kingdom); Latindex (México); Redalyc (México);
Periódica (México);
Emerging Sources Citation Index, Thomson Reuters;
MIAR–Matriz de información para el análisis de revistas;
RedIB Red Iberoamericana de Innovación y Conocimiento Científico;
DOAJ Directory of Open Access Journals.

Contenido general: páginas V y IX.



Revista INGENIERÍA UC
Impreso en Publicaciones de la
Facultad de Ingeniería de la
Universidad de Carabobo,
Campus de Bárbula,
Venezuela