

La función de pertenencia SDM en la clasificación espacial borrosa

Willin Álvarez

*Universidad Nacional Experimental Rómulo Gallegos. Facultad de Agronomía. Dpto de Instrumentales.
San Juan de los Morros, Estado Guárico.
E-mail: aprendizborroso@yahoo.com*

Resumen

En este artículo se presenta la función de pertenencia SDM (Sthephany-David-Mariana) como una alternativa para superar las dificultades ocasionadas por la aproximación directa en la clasificación espacial usando modelos de pertenencia borrosos o clasificación espacial borrosa. Además, se realiza una pequeña modificación al estimador Bayesiano de los vectores de pertenencia borrosos, a fin de eliminar la mala clasificación que surge cuando la suma de los coeficientes de regresión es negativa. Se ejecutan dos clasificaciones basadas en la teoría de los conjuntos borrosos; una, usando la aproximación directa sin modificar y otra modificando la anterior a través del procedimiento SDM. Posteriormente, se lleva a cabo una comparación entre los resultados de ambas clasificaciones, obteniendo en la clasificación que utiliza el Procedimiento SDM, vectores más fáciles de interpretar y casi insesgados.

Palabras clave: Clasificación borrosa, clasificación espacial borrosa, conjuntos borrosos, agrupamiento borroso, función de pertenencia.

The SDM membership function in the fuzzy spatial classification

Abstract

In this paper the SDM membership (Sthephany-David-Mariana) function is presented as an alternative to overcome the difficulties caused by the direct approach in the Spatial Classification using fuzzy membership models or Fuzzy Spatial Classification. Further, it is realized one small modification at the Bayesian estimator of the fuzzy membership vectors, with the purpose of suppressing the misclassification that appears when the addition of the regression coefficients is negative. Two classifications are executed based on the theory of the fuzzy sets; one using the direct approach without modification and the other modifying the previous one through of SDM procedure. Later, comparisons are established between the results of both classifications, obtaining in the classification with the SDM procedure, classification vectors easier to interpret and almost unbiased.

Keyword: Fuzzy classification, fuzzy spatial classification, fuzzy sets, fuzzy clustering, membership function.

1. INTRODUCCIÓN

Enmarcados en la Teoría Borrosa [1] se considera a un Clasificador Espacial Borroso como un sistema borroso que toma la decisión basada sobre el máximo grado de pertenencia. Esto es, si $\alpha_{f(x)} = \max_L f(x)$ se decide que el patrón X es de la clase C_α , donde $f(x)$ es el grado de pertenencia de la categoría C_L para el patrón X . El resultado o salida del Clasificador Espacial Borroso es una Clasificación Espacial Borrosa. En este ensayo se refiere a la clasificación cuando se tratan dos aspectos fundamentales: La creación de grupos a partir de los datos originales,

y la construcción de reglas que permiten la asignación de nuevos individuos a los grupos.

En cuanto a la creación de grupos a partir de los datos originales [2] se obtienen conglomerados de los datos de la cuenca Sur del Lago de Valencia presentados por [3]; tales conglomerados son obtenidos apoyados en la teoría de conjuntos Fuzzy (conjuntos borrosos) usando para ello un Análisis Cluster, según las ideas de [4].

Por otro lado, [5] proponen la construcción de reglas que permiten la asignación de nuevos individuos a los grupos basándose en dos modelos multivariados Gaussianos de campos aleatorios, aplicando una clasificación espacial que usa modelos de pertenencia borrosos, llamada aquí clasificación espacial borrosa.

Las ideas tratadas en el presente ensayo son un intento de proseguir los trabajos de [2] y [5]. En otras investigaciones se han estudiado la obtención de conglomerados usando las propuestas de [3], como por ejemplo [6], [7], [8] y [9]. Pero la asignación de nuevos individuos según [6] no ha sido explotada en otras literaturas, quizás debido a los inconvenientes que se mostrarán más adelante, y que se intentan superar por medio de un procedimiento propuesto en este ensayo.

2. AGRUPAMIENTO BORROSO

El análisis de datos multivariante se vale del álgebra matricial para la sustentación teórica y la representación de los datos. Por lo general, una matriz de datos se caracteriza porque sus filas representan a los individuos y sus columnas a las variables o características que son medidas sobre los individuos. Por otro lado, si se ha aplicado un método de agrupamiento clásico, obteniendo k grupos o clases, la matriz que define al agrupamiento tendrá k filas y n columnas; las filas describen a los grupos o clases; y las columnas a los individuos. Estas matrices clásicamente tienen un solo uno en sus columnas, indicando así la clase a la cual "pertenece un individuo". La matriz de agrupamiento es ilustrada en la Figura 1.

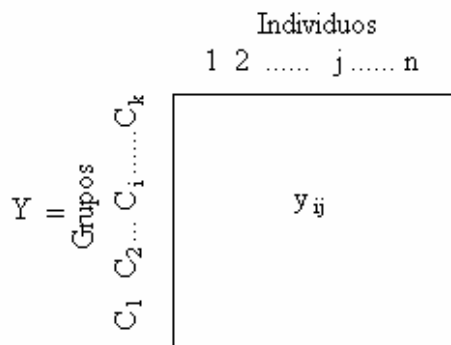


Figura 1. Representación de una matriz resultante de un agrupamiento.

Obsérvese que en un agrupamiento clásico los y_{ij} son 1 ó 0 según el individuo j pertenezca o no al grupo C_i .

Por otro lado, cuando el agrupamiento considerado es borroso la matriz se modifica y las componentes de sus vectores columnas son números entre cero y uno; es decir: $0 < y_{ij} < 1$ y además:

$$\sum_{i=1}^k y_{ij} = 1 \quad j=1, \dots, n \quad [12].$$

Cuando Y es una matriz que caracteriza a un agrupamiento borroso, sus vectores columnas son llamados vectores de agrupamiento borroso, elementos del proceso borroso o vectores de clasificación borrosa [5].

3. CLASIFICADOR QICAR

En [5] se considera como objetivo principal de la clasificación predecir los elementos del proceso borroso $\{y_r\}$ dadas las observaciones $\{x_r\}$. Esto es, caracterizar cada y_r según el sitio r . El modelo matemático propuesto para tal fin está dado por:

$$x_r = \sum y_{rj} \mu_j + e_r = \mu y_r + e_r, \quad r \in D \quad (1)$$

La vía más fácil para encontrar un vector aleatorio y no singular k -dimensional sobre el hiperplano:

$$\sum_{j=1}^k y_j = 1$$

Es trabajar con el vector de diferencias v ($k-1$)-dimensional. Tomando la diferencia a cada sitio r , se obtiene el proceso de diferencia $\{v_r\}$. El modelo particular más usado es el modelo QICAR, para el cual $\{v_r\}$ se distribuye $QICAR_{k-1,d}(\tau^2 G, \{\beta_s\})$; donde QICAR (quasi-intrinsic Gaussian conditional autoregression), es una auto regresión condicional gaussiana cuasi-intrínseca definida para un proceso gaussiano $\{u_r\}$ en $r\hat{Z}^d$ (latice entero), explicado ampliamente en [13]; $\tau > 0$ es un parámetro escalar y G es la matriz balanceada equicorrelacionada, de orden $(k-1) \times (k-1)$, con elementos $g_{ii}=1$ y $g_{ij}=1/2$ si $i \neq j$.

La aproximación directa define un vector de contrastes $v_{(k-1) \times 1}$ como:

$$v_j = y_i - y_k, \quad j = 1, 2, \dots, (k-1) \quad (2)$$

El inverso está dado por:

$$y = A_v + d \quad (3)$$

donde A es una matriz $k \times (k-1)$ con coeficientes.

$$p_{ii} = 1 - k^{-1}, 1 \leq i \leq k-1 \quad (4)$$

$$p_{ij} = -k^{-1}, i \neq j$$

y d es un vector de constantes k -dimensional con coeficientes $d_j = K^{-1}$, $1 \leq i \leq k$.

El estimador QICAR para la aproximación directa está dado por:

$$v_r = \left\{ A' \mu' \Sigma^{-1} \mu A + \tau^{-2} G^{-1} \right\}^{-1} \left\{ A' \mu' \Sigma^{-1} [x_r - \mu d] + \tau^{-2} G^{-1} \Sigma \beta_{s-r} v_s \right\} r \in D \quad (5)$$

Se asume que todos los parámetros $\mu_{pxk}, \Sigma_{pxp}, \tau^2 > 0$, y $\{\beta_\gamma\}$ en (5) se conocen.

El procedimiento de clasificación consiste en aplicar a (5) para posteriormente crear los vectores de clasificación por medio de (3). En [5] se advierte que al aplicar la regla (5) se pueden obtener en el vector de clasificación componentes resultantes con valores negativos y mayores que la unidad, este resultado es un inconveniente para interpretar la matriz de agrupamiento resultante. A continuación se presenta el siguiente ejemplo a fin de ilustrar el problema.

4. APLICACIÓN A DATOS DE SUELO DE CUENCA SUR DEL LAGO DE VALENCIA

La aplicación de (5) y (3) a un problema real requiere de dos conjuntos de datos denotados por $\{x_r\}$ y $\{y_r\}$. Aquí el primer conjunto se basa en datos muestreados en un bloque identificado con el número 24 en un mapa de la depresión del Lago de Valencia, correspondiente a uno de los dos conjuntos de datos utilizados por [3]. El segundo conjunto de datos es la matriz de partición $\hat{u}_{5 \times 48} = \{y_\gamma\}$ que contiene los valo-

res de la función de asociación $\hat{u}_\gamma(x_i)$, $r = 1, 2, 3, 4, 5$ $i = 1, \dots, 48$ esta matriz es el resultado del algoritmo ISO-DATA FUZZY con valores $F_c = 0,373$ y $H_c = 1,252$, aplicado al conjunto $\{x_r\}$ por [3].

4.1.- Dificultades encontradas

Las dificultades encontradas al aplicar el método de clasificación son básicamente debidas a la aproximación directa y al estimador de diferencia QICAR presentado en (5). Estas debilidades se discuten en los párrafos siguientes analizando algunos resultados, para mayor información se sugiere ver a [10].

La Tabla 1 presenta los grados de pertenencia de los sitios de la vecindad 2 (una vecindad está compuesta por cuatro sitios o puntos de muestreo), obsérvese que las filas corresponden a las cinco clases encontradas por [2] y las columnas representan a los cuatro sitios contenidos en vecindad 2.

Tabla 1. Grados de pertenencia de la Vecindad 2. Agrupamiento Borroso obtenido por ISODATA FUZZY[3].

Clase	Sitio 5	Sitio 6	Sitio 7	Sitio 8
1	0.048	0.801	0.453	0.524
2	0.064	0.095	0.254	0.186
3	0.72	0.038	0.128	0.109
4	0.06	0.08	0.046	0.039
5	0.108	0.058	0.119	0.132

La pertenencia de un sitio a una clase está presentada en forma gradual, por ejemplo, el sitio 5 pertenece a la clase 3 en una proporción de 0.72 (72%), 0.048 a la clase 1 (4.8%), 0.064 a la clase 2 (6.4%), 0.06 a la clase 4 (6%) y 0.108 a la clase 5 (10.8%). Nótese que el mayor grado de pertenencia se ha denotado en negritas (Tabla 1).

Dificultad 1: Al caracterizar estos sitios por la aproximación directa estimada por el modelo QICAR (de ahora en adelante ADQICAR) se obtuvo el resultado mostrado en la Tabla 2; véase que en los sitios 5 y 6 hay algunos $y_{ij} < 0$, y particularmente en el sitio 5, $y_{53} = 2.902 > 1$; los sitios 7 y 8 no presentan dificultad alguna; no obstante, todas asignaciones realizadas a las clases son buenas.

Tabla 2. Estimación ADQICAR de la Vecindad 2 aplicada a los datos de ISODATA FUZZY[2].

Clase	Sitio 5	Sitio 6	Sitio 7	Sitio 8
1	-0.563	0.965	0.484	0.548
2	-0.385	0.051	0.268	0.177
3	2.902	-0.00003	0.105	0.134
4	-0.604	-0.036	0.112	-0.016
5	-0.350	0.020	0.031	0.156

Los números en negritas expresan el mayor grado de pertenencia.

Para analizar mejor esta dificultad se presenta la Tabla 3 que contiene la caracterización de la vecindad 2.

Tabla 3. Caracterización de la vecindad 2

Sitio	$\Sigma\beta$	τ^2
5	5.349	0.202
6	1.279	0.116
7	1.15	0.067
8	0.911	0.179

Cuando $S_b = 5.39$ y $\tau^2 = 0.202$ (sitio 5 Tabla 3), los grados de pertenencia son negativos, a excepción de uno, quien a su vez es mayor que la unidad (columna 2 de la Tabla 2). Si S_b y τ^2 disminuyen, por ejemplo 1.279 y 0.116, respectivamente (sitio 6 Tabla 3), existen menos grados de pertenencia negativos y ningún grado es mayor a la unidad (columna 3 de la Tabla 2). Por otro lado, al disminuir S_b y τ^2 como en el sitio 7 se obtiene una buena estimación. Pero se

observa que para disminuciones altas de S_b , acompañados de aumentos considerables en τ^2 habrá por los menos un y_{ij} negativo, ver sitio 8 de las Tablas 2 y 3.

Dificultad 2: Al estimar los grados de pertenencia a los grupos usando el ADQICAR en los 48 sitios se encontró que hay una mala asignación en algunos de estos, con el fin de evidenciarlo se muestra a continuación la Tabla 4 contiene tres columnas de la matriz de agrupamiento a estimar (sitios 13, 23 y 42), resultados del ISODATA FUZZY propuesto por [2], y las tres columnas correspondientes de la matriz de agrupamiento estimada por el ADQICAR de [5]. Estos resultados son utilizados para comparar las componentes de los vectores a estimar con las respectivas componentes de los vectores estimados.

Obsérvese que en cada uno de estos sitios está mal asignado el máximo grado de pertenencia, por ejemplo para el sitio 13 (columna 2 de la Tabla 4) el mayor grado de pertenencia, arrojado por el ISODATA FUZZY está en la clase 1, pero la ADQICAR lo coloca en la clase 4 (columna 3 de la Tabla 4), situación similar ocurre con los otros sitios. Cabe destacar que este problema no fue señalado por [5].

5. CORRECCIÓN A LA ADQICAR. FUNCIÓN DE PERTENENCIA STHEPHANY- DAVID-MARIANA (SDM).

En vista de las dificultades encontradas en la ADQICAR, se realizó un análisis exploratorio de los datos y después de probar varias alternativas para la solución del problema se sugiere modificar a (5) mediante una función de pertenencia la cual se ha definido dentro de la teoría de los conjuntos borrosos y que se presentará próximamente; no sin antes señalar

Tabla 4. Resultados del ISODATA FUZZY propuesto por [2] y el ADQICAR de [5] en los sitios 13, 23 y 42.

Clase	ISODATA FUZZY Sitio 13	ADQICAR Sitio 13 $\Sigma\beta=-0.297$ $\tau^2=0.094$	ISODATA FUZZY Sitio 23	ADQICAR Sitio 23 $\Sigma\beta=-0.47$ $\tau^2=0.036$	ISODATA FUZZY Sitio 42	ADQICAR Sitio 42 $\Sigma\beta=-5.209$ $\tau^2=0.1245$
1	0.308	0.214	0.132	0.231	0.173	0.423
2	0.291	0.201	0.189	0.196	0.58	-1.453
3	0.123	0.156	0.258	0.178	0.093	0.655
4	0.058	0.244	0.135	0.225	0.044	0.852
5	0.21	0.185	0.286	0.170	0.11	0.523

que en [5] se propone una regla de proyección sobre el hiperplano:

$$\sum_{j=1}^k y_j = 1$$

Con la finalidad de interpretar los sitios con valores de pertenencia negativos y aquellos mayores a la unidad; sin embargo, tal propuesta puede ser tomada como una sobresimplificación de la información suministrada por el vector de clasificación borroso, por cuanto éste es llevado a una clasificación clásica para ejecutar su interpretación; por tal motivo se considera inadecuado realizar una clasificación bajo la teoría borrosa si luego ésta será interpretada por medio de un enfoque clásico de clasificación. Debido a lo anterior se sugiere utilizar una función de pertenencia que permita la interpretación de los vectores resultantes de las estimaciones manteniendo el enfoque borroso.

Las funciones de pertenencia presentadas en [2], [11] y [12] no se ajustan al problema de estudio, para valores con igual módulo pero que difieren en signo, por ejemplo -2 y 2 , se asigna el mismo valor de pertenencia, lo cual contradice la interpretación de la clasificación. Por tal razón, se necesita una función que asigne a los y_{ij} negativos bajos grados de pertenencia y que vaya aumentando estos grados a medida que los y_{ij} sean positivos o cercanos a los positivos. Esta función es representada en la gráfica de la Figura 1 y se denota como $q(y_{ij})$, su dominio es $\hat{A}$ y el conjunto de llegada es $[0,1]$, la función queda definida al escoger a , b y c adecuadamente, se ha convenido llamarla función Sthephany- David-Mariana (en honor a los Hijos del autor) y para simplificar función SDM.

En la Figura 2 puede observarse que la función SDM es definida a tramos. A fin de estar acorde con la estructura de los datos se presentan dos casos, en el primer caso existen al menos dos $y_{ij} > 0$ y en el segundo caso existe sólo un $y_{ij} > 0$.

Puesto que los vectores borrosos estimados no deben alcanzar la pertenencia y no pertenencia total, se escoge " a " de manera que sea mayor al máximo grado de pertenencia obtenido de la caracterización de [5] para un sitio particular; o sea, $a > \max \{y_{ij} | \hat{A}\}$ en un sitio r cualquiera. Así también " c " debe ser menor que el mínimo valor de los y_{ij} de la caracterización para asegurar la no pertenencia absoluta, luego $c < \min \{y_{ij} | \hat{A}\}$ en un sitio r cualquiera.

Tomando el criterio individual descrito por [11], se escoge el factor $\frac{k-1}{k-2}$, $k > 2$

(para $k=2$ se sugiere usar la proyección de [5]).

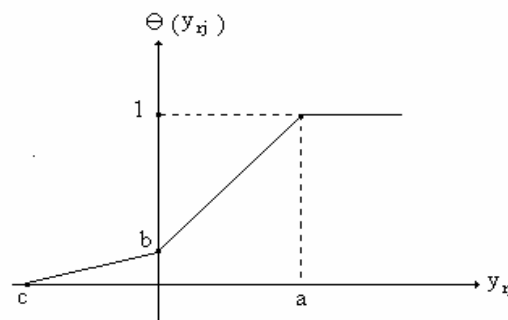


Figura 2. Representación Gráfica de la Función de pertenencia SDM.

Por otro lado, " b " debe manifestar la pertenencia intermedia (punto de mayor borrosidad) de los estimados y_{ij} de cada y_r resultante de la caracterización, para estar acorde a la naturaleza de datos (al menos dos $y_{ij} > 0$ o un $y_{ij} > 0$) se ejecutan dos cálculos de " b " de acuerdo al caso de estudio; en el caso 1 se toma " b " como la razón entre el mínimo valor positivo de los y_{ij} y el máximo valor, y en el caso 2 es la razón entre el máximo y_{ij} negativo y el único y_{ij} positivo.

Entonces los parámetros de $q(y_{ij})$ se calculan como sigue:

Caso 1: Existen al menos dos $y_{ij} > 0$ y $K > 2$.

$$\begin{aligned} a &= \frac{(k-1) \max \{y_{ij}\}}{(k-2)}; y_{ij} \in \mathfrak{R}^+ \\ b &= \frac{(k-1) \min \{y_{ij}\}}{(k-2) \max \{y_{ij}\}}; y_{ij} \in \mathfrak{R}^+ \\ c &= \frac{(k-1) \min \{y_{ij}\}}{(k-2)}; y_{ij} \in \mathfrak{R}^- \end{aligned} \quad (6)$$

Caso 2: Existen sólo un $y_{ij} > 0$ y $K > 2$; a y c son iguales al caso 1, b toma el valor dado por (7).

$$b = \frac{k-1}{k-2} \left| \frac{\max \{y_{ij}, y_{ij} \in \mathfrak{R}^-\}}{\min \{y_{ij}, y_{ij} \in \mathfrak{R}^+\}} \right| \quad (7)$$

Por último para $c < y_{ij} < a$ se toma la pertenencia gradual por medio de dos segmentos de rectas (ver Figura 2); para el segmento de recta donde $c < y_{ij} < 0$ se usa la ecuación de asignación

$$\theta(y_{ij} < 0) = -\frac{b}{c}y_{ij} + b, \quad c \neq 0$$

y para el segmento donde $0 \leq y_n < a$ se utiliza la ecuación de asignación

$$\theta(y_{ij} > 0) = \frac{1-b}{a}y_{ij} + b.$$

Entonces el transformado $q(y_{ij})$ está dado por:

$$\theta(y_{ij}) = \begin{cases} 0; & \text{si } y_{ij} \leq c \\ -\frac{b}{c}y_{ij} + b; & \text{si } c < y_{ij} < 0 \\ \frac{1-b}{a}y_{ij} + b; & \text{si } 0 \leq y_{ij} < a \\ 1; & \text{si } y_{ij} \geq a \end{cases} \quad (8)$$

5.1.- Procedimiento de Corrección SDM

Al intentar lograr el objetivo de presentar los vectores Borrosos en el hiperplano

$$\sum_{j=1}^k y_{rj} = 1$$

Después de ser corregidos por $\theta(y_{ij})$ se pueden presentar dos situaciones;

$$\sum_{j=1}^k y_{rj} < 1 \quad \text{ó} \quad \sum_{j=1}^k y_{rj} > 1,$$

en tales casos debemos hacer correcciones por defecto o exceso respectivamente, a esto se le ha llamado Procedimiento de corrección SDM. A continuación se muestran ambas correcciones.

Corrección por defecto: Se aplica cuando $\sum_{j=1}^k y_{rj} < 1$

Paso 1: Se calcula $\sum_{j=1}^k y_{rj} < 1$

Paso 2: Se halla el defecto total como: $\alpha = 1 - \sum y_n$

Paso 3: Se determina el defecto promedio α/k

Paso 4: Se corrige el sitio r , como $y_{rj} + (\alpha/k)$, con $j = 1, 2, \dots, k$

Corrección por Exceso: Se aplica cuando $\sum_{j=1}^k y_{rj} > 1$

Paso 1: Se calcula $\sum_{j=1}^k y_{rj} > 1$

Paso 2: Se halla el exceso total como $\alpha = \sum y_{rj} - 1$

Paso 3: Se determina el exceso promedio α/k

Paso 4: Si existe algún $y_{rj} < \alpha/k$; $y_{rj} \neq 0$, se hace cero el $\min\{y_{rj}\}$ y $k = k-1$ y se vuelve al paso 1. Si ningún $y_{rj} < \alpha/k$; $y_{rj} \neq 0$, se sigue al paso 5.

Paso 5: Se corrige el sitio r , como $y_{rj} - \alpha/K$; $y_{rj} \neq 0$, $j = 1, 2, \dots, k$.

Es importante señalar que para los transformados sólo se necesitan las ecuaciones siguientes:

$$\theta(y_{ij} < 0) = -\frac{b}{c}y_{ij} + b \quad \text{y} \quad \theta(y_{ij} > 0) = \frac{1-b}{a}y_{ij} + b \quad (9)$$

Nota : En los sitios donde $\sum |y_{rj}| \leq 2$, no es conveniente aplicar a $\theta(y_{ij})$, en estos casos se toma el valor absoluto de todos los y_{rj} para el sitio r en cuestión y es aplicada la corrección por exceso. A la aplicación de la función SDM y posteriormente el procedimiento de corrección SDM se le ha denominado "Procedimiento SDM".

6. RESULTADOS OBTENIDOS Y DISCUSIÓN

A fin de superar las dificultades encontradas en la clasificación usando la ADQICAR en los sitios de muestreo señalados anteriormente de la cuenca Sur del Lago de Valencia, se procedió a corregir de la siguiente manera:

6.1. Corrección de la dificultad 1

La eliminación de la dificultad 1 consistió en aplicar el Procedimiento SDM a los resultados del ADQICAR que así lo ameritaban.

Sitio 5: En la Tabla 5 se puede ver que se han corregido los valores negativos y mayores que 1, y además los y_{ij} son casi insesgados.

Tabla 5. Correcciones efectuadas en el sitio 5 de la vecindad 2 por el Procedimiento SDM.

Clase	ISODATA FUZZY	ADQICAR	Procedimiento S.D.M.
1	0.048	-0.563	0.0322
2	0.064	-0.385	0.0713
3	0.72	2.901	0.7944
4	0.06	-0.604	0.0232
5	0.108	-0.350	0.0789

Los números en negritas expresan el mayor grado de pertenencia

Sitio 6: Se procedió igual que en el caso anterior, los resultados se muestran en la Tabla 6.

Tabla 6. Correcciones efectuadas en el sitio 6 de la vecindad 2 por el Procedimiento SDM.

Clase	ISODATA FUZZY	ADQICAR	Procedimiento SDM
1	0.801	0.965	0.9469
2	0.095	0.051	0.0327
3	0.038	-0.00003	0
4	0.08	-0.036	0.0183
5	0.058	0.020	0.0021

Sitio 7: Aquí no se corrigió ya que no fue necesario.

Sitio 8: Presenta la particularidad de que los valores de $\Sigma|y_{rj}|$ son casi uno, y al usar el Procedimiento SDM completo el vector corregido se acerca mucho a la proyección de [5]; en estos casos se aplica la corrección

por exceso sin pasar por la función SDM tal y como se explica en [10]. Los resultados se pueden ver en la Tabla 7.

Tabla 7. Correcciones efectuadas en el sitio 8 de la vecindad 2 por el Procedimiento SDM.

Clase	ISODATA FUZZY	ADQICAR	Procedimiento SDM
1	0.524	0.5483	0.5421
2	0.186	0.1775	0.1713
3	0.109	0.1336	0.1276
4	0.039	-0.0156	0.00094
5	0.132	0.1561	0.1498

6.2.- Corrección de la dificultad 2

Para eliminar la dificultad 2, se usó el valor absoluto de $\Sigma\beta$ en vez de $\Sigma\beta$ en el estimador del vector de diferencia presentado en la ecuación (5); se aplicó esta modificación en cada sitio que lo ameritaba y se encontraron los siguientes resultados.

En la Tabla 8 se muestra que la dificultad 2 fue eliminada; no obstante, en algunos casos después de suprimirla se presentó la dificultad 1, como se aprecia en el cuadro correspondiente al sitio 42 (columna 7 de la Tabla 8), a estos nuevos resultados de los sitios afectados se les aplicó el Procedimiento SDM, para superar nuevamente a la dificultad 1. Aplicando las correcciones a cada sitio r donde se presentaron las dificultades 1 y 2 se obtiene la matriz de los vectores borrosos corregidos $\{\hat{y}_r\}$, ésta puede encontrarse en [10].

Tabla 8: Uso del valor absoluto de $\Sigma\beta$ en el estimador (5). Corrección efectuada en los sitios 13, 23 y 42.

Clase	ISODATA FUZZY Sitio 13	ADQICAR Sitio 13 $ \Sigma\beta = 0.297$ $t^2=0.094$	ISODATA FUZZY Sitio 23	ADQICAR Sitio 23 $t^2=0.036$ $ \Sigma\beta = 0.47$	ISODATA FUZZY Sitio 42	ADQICAR Sitio 42 $ \Sigma\beta = 5.209$ $t^2=0.1245$
1	0.308	0.272	0.132	0.167	0.173	-0.0002
2	0.291	0.25	0.189	0.186	0.58	2.098
3	0.123	0.115	0.258	0.232	0.093	-0.304
4	0.058	0.172	0.135	0.164	0.044	-0.516
5	0.21	0.191	0.286	0.250	0.11	-0.277

7. CONCLUSIONES Y RECOMENDACIONES

Si el valor de $\Sigma\beta_{s-y}$ es negativo, entonces se realizará una mala Clasificación puesto que los grados de pertenencia de los y_{ij} son mal asignados; esto se corrige usando el valor absoluto de $\Sigma\beta_{s-y}$, en cada sitio.

Cuando $|\Sigma\beta_{s-y}|$ se aproxima a 2 y siendo t^2 muy pequeño los valores de las componentes de los vectores borrosos y_{ij} , son negativos; debe observarse que aún así la asignación de los sitios a un conglomerado es “correcta”. No obstante, con el fin de brindar una mejor interpretación deben hacerse correcciones en las componentes de los vectores de clasificación.

La proyección utilizada por [6] para fines de interpretación, puede ser vista como una sobresimplificación de la información de los vectores borrosos, por tal motivo se recomienda usar el procedimiento SDM.

La función SDM tiene la ventaja de asignar los grados de pertenencia a valores de y_{ij} tanto positivos como negativos, pero destacando un mayor grado de pertenencia a los valores positivos.

Cuando la variabilidad en cada vecindad t^2 es muy pequeña y $0.5 < |\Sigma\beta_{s-y}| < 1.5$ la estimación del vector asociado a un sitio en particular es casi insesgada.

REFERENCIAS BIBLIOGRAFICAS

- [1] Zadeh, L.A.(1965) . “Fuzzy Sets”. Information and Control .Vol.28, pp 338-353
- [2] Balza, M.A. (1995). “Obtención de Conglomerados a la través de la Teoría de Conjuntos Difusos (Fuzzy Sets)”.Tesis Doctoral Universidad Central de Venezuela Facultad de Agronomía.
- [3] Rey, J (1992). “Validación del Uso de Sistema de Clasificación en Base a Fertilidad al Sur-Este de la cuenca del Lago de Valencia”. Tesis para optar al Título de Ingeniero Agrónomo Mención Fitotecnia. Facultad de Agronomía de la Universidad Central de Venezuela.
- [4] Leszczyński, P. Penczek y W. Grochulski (1985). “Sugeno’s Fuzzy measure and Fuzzy clustering”. Fuzzy Sets and Systems. N° 15, pp 147-158.
- [5] Kent, J.T. y Mardia, K.V. (1988). “Spatial Classification using Fuzzy Membership Models”. IEE Transactions on Pattern Analysis and Machine: Intelligence Vol.10 N°5. September.
- [6] Burrough, P.A.; van Gaans P.F. M. y MacMillan R.A. (2000). “High-resolution landform classification using fuzzy k-means”. Fuzzy Sets and Systems **113**, 37 - 52.
- [7] Mazaheri, S.A.; Koppi, A.J. y Mc Bratney (1995). “A Fuzzy allocation scheme for the Australian Great Soil Groups classification system”. European Journal of Soil Science 45, 601 – 612.
- [8] Ménard Michel (2001). “Fuzzy clustering and switching regression models using ambiguity and distance rejects”. Fuzzy Sets and Systems 122, 363 -399.
- [9] Mc Bratney y Gruliter (1992). “A continuum approach to soil classification by modified fuzzy k-means with extragrades. Journal of Soil Science 43, 159 – 175.
- [10] Álvarez, W. (2002). “Clasificación Espacial Borrosa”. Tesis de Maestría en Estadística. Universidad Central de Venezuela Facultad de Agronomía.
- [11] Azorín, P. F. (1979). “Algunas aplicaciones de los conjuntos borrosos a la Estadística”. INE, Madrid.
- [12] Kaufmann A. (1975). “Introduction to the Theory of Fuzzy Subsets” Volumen I. Fundamental Theoretical Elements. Universite de Louvain, Louvain Belgium. Academic Press, INC.
- [13] Künsch Hans R. (1987). “Intrinsic autoregressions and related models on the two - dimensional lattice”. Biometrika, 74,3, pp. 517-24.
- [14] Ramírez G. y Vásquez M. (2002). “Clasificación y Borrosidad”. Comisión de Estudios Interdisciplinarios. Publicaciones Año 4. N13.Junio. Universidad Central de Venezuela