

APRENDIZAJE EVOLUTIVO DE SISTEMAS DIFUSOS LINGÜÍSTICOS BASADO EN LA COOPERACIÓN ENTRE LOS OPERADORES Y LAS FUNCIONES DE PERTENENCIA

A.A. Márquez F.A. Márquez A. Peregrín

Dept. de Tecnologías de la Información. Universidad de Huelva 21071 – Huelva,
{amarquez, alfredo.marquez, peregrin}@dti.uhu.es

Resumen

Este trabajo presenta un modelo evolutivo que aprende de forma simultánea los parámetros de los operadores difusos adaptativos del Sistema de Inferencia y la Defuzzificación y el ajuste de las funciones de pertenencia con el propósito de obtener un mayor nivel de cooperación entre ambos elementos: Base de Conocimiento y operadores difusos, que se traduzca en una mejor precisión del modelo manteniendo un buen nivel de interpretabilidad. Por tanto, este trabajo propone un mecanismo para que se pueda dar una sinergia positiva entre dos elementos de diseño diferentes que siendo aprendidos de forma simultánea ofrecen mejores resultados en las aplicaciones.

Palabras Clave: Modelado Difuso Lingüístico, Interpretabilidad, Defuzzificación Adaptativa, Sistema de Inferencia Adaptativo, Ajuste de las Particiones Difusas.

1 INTRODUCCION

El Modelado Difuso Lingüístico (MDL) [3] permite la obtención de modelos más fácilmente interpretables por el ser humano. Sin embargo, la precisión es una propiedad contradictoria con la citada interpretabilidad, por lo que en sistemas complejos, siendo ambas propiedades deseables, suele ser necesario llegar a un compromiso entre ellas. Por este motivo diferentes autores han realizado diferentes propuestas en las que han tratado de dotar al MDL de una mayor precisión preservando el mayor grado posible de interpretabilidad.

Generalmente, de las dos tareas básicas de diseño que son la obtención de la Base de Conocimiento (BC) y la elección de los operadores del Sistema de Inferencia y

la Interfaz de Defuzzificación, la primera es la más ampliamente estudiada.

En relación con el aprendizaje de la BC, una de las técnicas más conocidas para mejorar la precisión de los Sistemas Basados en Reglas Difusas (SBRDs) es el llamado ajuste de las funciones de pertenencia contenidas en la Base de Datos (BD) que consiste en el refinamiento de la misma, manteniendo la Base de Reglas (BR) fija. Concretamente, el proceso de ajuste consiste en una variación en la forma y en la ubicación de su universo de discurso de las funciones de pertenencia predefinidas.

En cuanto a la elección de los operadores difusos, como se indicaba anteriormente existen menos aportaciones, pero recientemente su relevancia es creciente en tanto se ha revelado como un recurso útil. El uso de operadores adaptativos tanto en el Sistema de Inferencia [1] como en la Interfaz de Defuzzificación [8], o en ambos de modo simultáneo [12] ha demostrado ofrecer mejoras significativas en la precisión de diversas aplicaciones.

En [14] se muestra un modelo que aprende tanto la BR, partiendo de la BD fija, como los operadores adaptativos en el Sistema de Inferencia e Interfaz de Defuzzificación, obteniéndose sistemas con un alto nivel de cooperación entre las reglas y los operadores, que se traduce en una mejora significativa de la precisión.

En este trabajo se propone un modelo que al igual que el referido en el párrafo anterior, aprende elementos tanto de la BC como del Sistema de Inferencia y Defuzzificación, pero en este caso eligiendo el ajuste de las funciones de pertenencia que se almacenan en la BD, en un mismo modelo evolutivo, esto es, favoreciendo la cooperación entre los operadores del Sistema de Inferencia y las funciones de pertenencia al ser aprendidos de forma simultánea.

El trabajo se organiza de la siguiente forma: En la Sección 2 se introducen los operadores difusos adaptativos. En la Sección 3 se describe en primer lugar la metodología de aprendizaje del ajuste de las las

funciones de pertenencia mediante algoritmos evolutivos, y posteriormente la propuesta del modelo evolutivo cooperativo que se plantea en este trabajo. En la Sección 4 se desarrolla un breve estudio experimental, y finalmente se incluye una Sección dedicada a las conclusiones.

2 OPERADORES DIFUSOS ADAPTATIVOS

En esta sección se describen los fundamentos del Sistema de Inferencia y de la Interfaz de Defuzzificación adaptativos que serán utilizados posteriormente en la siguiente sección para el modelo propuesto.

2.1. OPERADOR DE CONJUNCIÓN ADAPTATIVO

Los SBRDs para modelado utilizan reglas de tipo SI - ENTONCES de la forma siguiente:

$$R_i: Si X_{i1} es A_{i1} y \dots y X_{im} es A_{im} entonces Y es B_i$$

con $i = 1$ hasta N reglas, siendo X_{i1} hasta X_{im} las entradas e Y la salida, y con A_{i1} hasta A_{im} y B_i los antecedentes y consecuentes respectivamente.

La expresión de la Regla Composicional de Inferencia en el modelado difuso con fuzzificación puntual es la que se muestra en la expresión (1), donde $\mu_{B'}$ es la función de pertenencia del consecuente inferido, $I(\cdot)$ es el operador de implicación, $C(\cdot)$ el operador de conjunción, μ_{A_i} al tratarse de fuzzificación puntual son los puntos de corte de las entradas discretas (x_1, \dots, x_n) del sistema con las funciones de pertenencia de los antecedentes de las reglas, y μ_B el consecuente de la regla.

$$\mu_{B'}(y) = I(C(\mu_{A_1}(x_1), \dots, \mu_{A_n}(x_n)), \mu_B(y)), \quad (1)$$

Por tanto, el sistema de inferencia emplea dos operadores: el de conjunción, $C(\cdot)$ para calcular el grado de emparejamiento, y el de implicación, $I(\cdot)$ para calcular a continuación el consecuente inferido.

En [1 y 11] se desarrollaron estudios sobre los diferentes métodos de introducir parámetros en el Sistema de Inferencia, observándose menor influencia en el uso de operadores de implicación adaptativos frente al de los de conjunción. Por este motivo, en este trabajo se utiliza la parametrización del Sistema de Inferencia exclusivamente mediante el operador de conjunción.

Las T-normas son una familia de operadores frecuentemente utilizados en el diseño de sistemas difusos, modelando tanto el operador de conjunción como el de implicación [10] y ofreciendo buen

comportamiento [7]. En este trabajo, el Sistema de Inferencia Adaptativo se basa en utilizar las t-normas parametrizadas de Dubois, Dombi y Frank recogidas en [15], cuyas expresiones se muestran en (2), (3) y (4).

$$\text{Dubois} \quad T_{\text{Dubois}}(x, y, \alpha) = \frac{x \cdot y}{\text{Max}(x, y, \alpha)} \quad (2)$$

$$\text{Dombi} \quad T_{\text{Dombi}}(x, y, \alpha) = \frac{1}{1 + \alpha \sqrt{\left(\frac{1-x}{x}\right)^\alpha + \left(\frac{1-y}{y}\right)^\alpha}} \quad (3)$$

$$\text{Frank} \quad T_{\text{Frank}}(x, y, \alpha) = \log_\alpha \left[1 + \frac{(\alpha^x - 1)(\alpha^y - 1)}{\alpha - 1} \right] \quad (4)$$

La Tabla 1 muestra la relación entre las cinco t-normas clásicas y las t-normas parametrizadas anteriormente citadas dependiendo del valor del parámetro α .

Tabla 1: Relación entre las t-normas clásicas y las t-normas parametrizadas dependiendo del valor del parámetro.

	$T_{\text{Mín}}$	T_{Ham}	T_{Alg}	T_{Einst}	T_{Acot}	$T_{\text{Drást}}$
T_{Frank}	$\rightarrow 0$		$\rightarrow 1$		∞	
T_{Dombi}	∞	1				$\rightarrow 0$
T_{Dubois}	0		1			

Se podrían considerar dos modelos distintos dependiendo de si se opta por utilizar un sólo parámetro α para toda la expresión del operador de conjunción, o distintos parámetros α_i , uno para cada regla de la BR. Un sólo parámetro permite ajustar globalmente el comportamiento del conectivo mientras que en el caso de los parámetros individuales, éstos ajustan dicho comportamiento localmente para cada regla. El modelo de un sólo parámetro ofrece peores resultados prácticos en cuanto a precisión [1], por lo que en este trabajo se utiliza el modelo de un parámetro para cada regla.

La utilización de operadores de conjunción adaptativos modula la influencia del grado de emparejamiento atenuándolo o acentuándolo. Este cambio no es lineal debido a que las expresiones de las t-normas adaptativas no son lineales. El efecto que producen es pues semejante a un mecanismo para modificar el significado lingüístico de la estructura de las reglas [13], que consiste en relajar la estructura de la regla cambiando el significado de las etiquetas involucradas.

En este caso, el parámetro en el operador de conjunción juega el mismo papel que cambiar las funciones de pertenencia asociadas con los antecedentes de las reglas. Estos antecedentes cambian de la misma forma que lo hace el operador o función aplicada. Las t-normas parametrizadas como operadores de conjunción sólo producen efectos de concentración.

2.2. MÉTODO DE DEFUZZIFICACIÓN ADAPTATIVO

La expresión formal (5) se emplea generalmente para generar métodos de defuzzificación adaptativos,

$$y_0 = \frac{\sum_i^N f(h_i) \cdot V_i}{\sum_i^N f(h_i)}, \quad (5)$$

donde h_i es el grado de emparejamiento, $f(h_i)$ es por tanto un término funcional del grado de emparejamiento y V_i es el máximo valor (MV_i) o el centro de gravedad (CG_i). Se trata de la expresión de un método de defuzzificación que actúa en Modo B, es decir, primero defuzzifica la aportación individual de la inferencia obtenida con cada regla y posteriormente se computa una suma ponderada.

El término funcional puede utilizar un solo parámetro β , o bien un parámetro por cada regla de la BC, β_i , que es la opción utilizada en este trabajo por obtener mejores resultados prácticos [8]. Por otro lado, el término funcional puede ser de tipo producto o potencia, como se muestra en las expresiones (6) y (7).

$$f(h_i) = h_i \cdot \beta_i, \quad f(h_i) = h_i^{\beta_i}. \quad (6), (7)$$

La combinación de estos dos funcionales con el MV o el CG genera cuatro métodos de defuzzificación. El empleo del funcional de tipo potencia con un parámetro para cada regla es equivalente a una de las técnicas empleadas para realizar la modificación lingüística de la estructura de la regla [13]. Por otro lado, el funcional de tipo producto tiene un efecto equivalente al uso de reglas con pesos [5]. A continuación se muestra un conjunto de reglas con pesos, donde éstos son denominados por w_i :

Si X_{11} es A_{11} y... y X_{1n} es A_{1n} entonces Y es B_1 con w_1

Si X_{21} es A_{21} y... y X_{2n} es A_{2n} entonces Y es B_2 con w_2

...

Si X_{n1} es A_{n1} y... y X_{nm} es A_{nm} entonces Y es B_n con w_n

En este trabajo se ha optado por el funcional de tipo producto con un parámetro para cada regla por sus buenos resultados empíricos mostrados en [8], y utilizando CG, cuya expresión es (8).

$$y_0 = \frac{\sum_i^N h_i \cdot \beta_i \cdot CG_i}{\sum_i^N h_i \cdot \beta_i}, \quad (8)$$

Este método de defuzzificación es equivalente al conocido WCOA cuando los parámetros se igualan a la unidad.

3. MODELO EVOLUTIVO COOPERATIVO DE APRENDIZAJE

Como se ha indicado con anterioridad, el aprendizaje de la BC de un SBRD se basa generalmente en la existencia de una BD previamente definida [16]. Generalmente, la BD es obtenida fijando tanto un número de etiquetas para cada variable como los parámetros de dichas etiquetas y distribuyéndolas de modo uniforme en el universo de discurso de la variable. La BD tiene una gran influencia en el comportamiento del SBRD, y prefiarlas para obtener la BR condiciona en gran medida la bondad final de la BC resultante.

Otros enfoques realizan un ajuste a posteriori de la BD [6]. Para ello, aprenden los parámetros o la forma de las funciones de pertenencia, mejorando la precisión del SBRD. No obstante, el proceso de aprendizaje está sesgado por la BD y BR iniciales, quedando fijos aspectos tales como la granularidad de las particiones, o el número y la estructura de las reglas.

En el presente trabajo, en una primera aproximación, hemos tomado este segundo enfoque. Además, con el fin de minimizar la pérdida de interpretabilidad que se produce al realizar el aprendizaje de las etiquetas, los márgenes de variación de éstas se han sesgado para que sólo puedan variar en cierta medida, tal como se comentará posteriormente en la descripción de las asignaciones de valores iniciales de esta parte del cromosoma.

El modelo cooperativo presentado en este trabajo utiliza un modelo evolutivo basado en un Algoritmo Genético (AG) de tipo CHC [9] cuyo cromosoma tiene un esquema de codificación triple ($CS_T + CS_C + CS_D$), donde CS_T codifica las funciones de pertenencia, CS_C los parámetros α_i del operador de conjunción, y CS_D los parámetros β_i del método de defuzzificación.

La parte del cromosoma CS_T utiliza un esquema con codificación real, siendo m_i el número de etiquetas de cada una de las n variables que comprenden la BD,

$$C_i = (a_i^1, b_i^1, c_i^1, \dots, a_i^m, b_i^m, c_i^m), \quad i = 1, \dots, n$$

$$CS_T = C_1 C_2 \dots C_n$$

Cada gen toma valores en un cierto intervalo el cual es obtenido a partir de la BD inicial, de tal forma que, por cada etiqueta lingüística $C_i^j = a^j, b^j, c^j$, el intervalo de variación se calcula de la siguiente forma:

$$[I_{a_j}^l, I_{a_j}^r] = [a_j - (b_j - a_j)/2, a_j + (b_j - a_j)/2]$$

$$[I_{b_j}^l, I_{b_j}^r] = [b_j - (b_j - a_j)/2, b_j + (c_j - b_j)/2]$$

$$[I_{c_j}^l, I_{c_j}^r] = [c_j - (c_j - b_j)/2, c_j + (c_j - b_j)/2]$$

La zona del cromosoma dedicada a los parámetros de los operadores, tiene dos partes:

- La del operador de conjunción, CS_C , que utiliza N parámetros α_i (genes), uno para cada R_i de la BR. Cada gen toma valores en el intervalo $[0,1]$ para la t-norma de Dubois, $(0,10]$ para Dombi y $(0,100]$ para Frank.
- La del método de defuzzificación, CS_D , que utiliza N parámetros β_i , uno para cada R_i de la BR. Cada gen en este caso toma valores en el intervalo $[0,10]$, suficiente para actuar como peso de esa regla.

El AG elegido es un CHC [9] considerado como un modelo evolutivo con un buen balance entre diversidad y convergencia en problemas con grandes espacios de búsqueda.

En cada generación, el algoritmo CHC utiliza una población de padres de tamaño M para generar una población intermedia de M individuos, los cuales son aleatoriamente emparejados y utilizados para generar M' descendientes. Se mantiene entonces una competición por la supervivencia, donde los mejores M cromosomas de las poblaciones de padres y descendientes son seleccionados para constituir la nueva generación.

En la fase de recombinación no se aplica ningún proceso de mutación. En su lugar, cuando la población converge o el proceso de búsqueda deja de progresar adecuadamente (el umbral de cruce llega a 0 y no se generan nuevos descendientes), la población se reinicia. El cromosoma que representa la mejor solución hasta ese momento se utiliza como patrón para generar la nueva población en un cromosoma, mientras que el resto es aleatorio.

Aunque el algoritmo CHC fue concebido para cromosomas con codificación binaria, existen versiones para su uso con cromosomas con codificación real como la empleada en la parte correspondiente a los parámetros de este trabajo, la cual utiliza el BLX- α como operador de cruce ($\alpha=0.5$). El cálculo de la distancia de Hamming se realiza convirtiendo los genes reales en una representación en cadenas y hallando la distancia de Hamming teniendo en cuenta si difiere o no cada uno de los caracteres. Sólo aquellas cadenas con una distancia mayor del umbral son combinados. El umbral se inicializa a $L/4$ siendo L la longitud de la cadena o cromosoma. Cuando ningún descendiente es insertado en la nueva población, el umbral se reduce en una unidad.

El tamaño de la población es de 50 individuos aleatoriamente inicializados, salvo uno, con la siguiente configuración:

- La parte de la BD (funciones de pertenencia), CS_T , con las funciones de pertenencia inicialmente almacenadas en la BD original como el primer individuo, y el resto como se ha indicado se generan aleatoriamente en el correspondiente intervalo de variación.
- La parte de parámetros de los operadores difusos:
 - Operador de conjunción, CS_C , con los N genes inicializados a 0 para la t-norma de Dubois (equivalente a la t-norma del mínimo), a 1 en caso de la de Dombi y en 0.5 en caso de la de Frank.
 - Método de defuzzificación, CS_D , con los N genes que representan los pesos de las reglas inicializados a 1, con objeto de partir con el WCOA.

El modelo evolutivo utiliza un solo umbral, tanto para la parte del cromosoma del ajuste de la BD como para la parte de los operadores difusos.

La función objetivo utilizada es el clásico Error Cuadrático Medio (ECM), mostrado en (9).

$$ECM (S[i])_B = \frac{1}{2} \frac{\sum_{k=1}^r (y_k - S[i](x_k))^2}{P} \quad (9)$$

donde $S[i]$ es el modelo difuso cuyo Sistema de Inferencia utiliza la t-norma de Dubois, Dombi o Frank ($i=1,\dots,3$) como operador de conjunción y la t-norma del mínimo como operador de implicación, y el método de defuzzificación WCOA adaptativo con pesos, utilizando una BD concreta decodificada. Esta medida utiliza un conjunto de datos para evaluación constituido por P pares de datos numéricos $Z_k=(x_k, y_k)$, $k=1,\dots,P$, siendo x_k los valores de las variables de entrada y y_k los correspondientes a las variables de salida.

4. ESTUDIO EXPERIMENTAL

En esta Sección se propone un estudio experimental en el que se utiliza una aplicación real, y se compara la precisión obtenida con distintos modelos difusos cuyos resultados se someten a un estudio estadístico con objeto de mostrar lo significativo de las mejoras de precisión obtenidas con el modelo cooperativo propuesto.

4.1. DESCRIPCIÓN DE LA APLICACIÓN

La aplicación considerada es el problema de distribución del sistema eléctrico español descrito en [4], que consiste en el cálculo de los costes óptimos

teóricos de la línea de media tensión en núcleos urbanos.

Para ello se dispuso de un conjunto de datos de 1059 ciudades que constan de cuatro variables de entrada y una de salida. El dominio de las variables de entrada y salida fue dividido en cinco etiquetas {MP, P, M, G, MG}.

Utilizamos una BR con 65 reglas lingüísticas obtenidas con el método de WM [16] a partir de un conjunto de datos de entrenamiento con el 80% de los datos disponibles, es decir, 847 de las 1059 ciudades, tomados de modo aleatorio.

Se ha utilizado un modelo de validación cruzada con 5 particiones de datos, esto es, 5 particiones aleatorias al 20% (4 de ellas con 212 ejemplos y una con 211), y la combinación de cuatro particiones (80%) como entrenamiento y la partición restante (20%) como test. Así se tienen 5 particiones al 80% y 20% en entrenamiento y test respectivamente.

4.2 METODOLOGÍA DE COMPARACIÓN

La Tabla 2 resume los modelos que se han construido con objeto de comparar su precisión. La primera línea contiene el de referencia, WM. A continuación los tres modelos adaptativos evolutivos, esto es, el modelo con aprendizaje de los operadores difusos (WM + C-D), el modelo con aprendizaje de la BD, es decir, las funciones de pertenencia (WM + T), y el modelo anterior sobre el que se aprenden a posteriori los parámetros de los operadores difusos (WM + T + C-D). Por último aparece el modelo cooperativo presentado en este trabajo (WM + T-C-D), en el que el aprendizaje de las funciones de pertenencia y de los operadores difusos se realiza al mismo tiempo.

Tabla 2. Modelos Difusos del Estudio Experimental

Abreviatura	Aprendizaje Funciones Pertenencia	Conjun. Adapt.	Defuzz. Adapt.
Modelo de Referencia			
WM			
Modelos Adaptativos Evolutivos Op. Difusos			
WM + C _{DUBOIS} -D		•	•
WM + C _{DOMBI} -D		•	•
WM + C _{FRANK} -D		•	•
Modelo Adaptativo Evolutivo BD			
WM + T	•		
Modelos Adaptativos Evolutivos BD + Op. Difusos			
WM + T + C _{DUBOIS} -D	•	•	•
WM + T + C _{DOMBI} -D	•	•	•
WM + T + C _{FRANK} -D	•	•	•
Modelo Evolutivo Cooperativo propuesto			
WM + T- C _{DUBOIS} -D	•	•	•
WM + T- C _{DOMBI} -D	•	•	•
WM + T- C _{FRANK} -D	•	•	•

En los modelos donde existe aprendizaje de los operadores difusos se presenta uno para cada t-norma adaptativa utilizada como conjunción (Dubois, Dombi y Frank).

Para cada una de las 5 particiones generadas, cada modelo evolutivo se ha ejecutado 6 veces, dando lugar a 30 ejecuciones de cada modelo, resultado de la combinación de 6 semillas para el generador de números aleatorios sobre las 5 particiones diferentes de datos obtenidas.

Los distintos modelos se han comparado utilizando el la medida del ECM (9) que se introdujo en la Sección 3, si bien se computa la media aritmética de los 30 resultados. La Tabla 3 muestra las evaluaciones realizadas por el modelo evolutivo CHC para cada modelo difuso.

Tabla 3. Evaluaciones realizadas por el algoritmo evolutivo CHC utilizado

Modelo	Evaluaciones
WM + T	30.000
WM + C-D	30.000
WM + T + C-D	60.000
WM + T-C-D	60.000

4.3 RESULTADOS Y ANÁLISIS

En esta sección se muestran los resultados obtenidos y el análisis de los mismos. En primer lugar los correspondientes al ECM de los modelos difusos utilizados, y en segundo lugar los resultados resumidos del análisis estadístico.

4.3.1 RESULTADOS GENERALES

La Tabla 4 muestra los ECMs de los distintos modelos difusos construidos.

Tabla 4. ECM obtenido con los diferentes modelos para el problema de distribución del sistema eléctrico.

Modelo	ECM _{entrenamiento}	ECM _{prueba}
WM	56135,75	56359,42
WM + C _{DUBOIS} -D	22561,77	25492,77
WM + C _{DOMBI} -D	28627,74	31558,00
WM + C _{FRANK} -D	25691,23	28895,61
WM + T	14962,14	18122,23
WM + T + C _{DUBOIS} -D	8435,83	12203,10
WM + T + C _{DOMBI} -D	9193,79	13020,94
WM + T + C _{FRANK} -D	14759,64	18922,81
WM + T- C _{DUBOIS} -D	5990,18	9101,50
WM + T- C _{DOMBI} -D	7193,25	10913,21
WM + T- C _{FRANK} -D	7124,48	9990,57

En ella podemos observar:

A) En cuanto a los modelos adaptativos evolutivos (WM + T, WM + C-D y WM + T-C-D):

- Como es bien conocido, el ajuste de las funciones de pertenencia es una herramienta muy útil para mejorar de forma importante la precisión del modelo.
- La mejora que presenta aplicar, al resultado del ajuste de las funciones de pertenencia, el aprendizaje de los operadores difusos es también muy relevante, por lo que el ajuste de las funciones de pertenencia y el aprendizaje de los operadores son dos métodos distintos pero compatibles para mejorar la precisión.
- Entre las t-normas parametrizadas, la de Dubois, además de ser computacionalmente más eficiente, es la que mejores resultados obtiene, tal como se observó en [7].

B) En cuanto al modelo de aprendizaje cooperativo evolutivo, cuyos resultados se encuentran en las tres últimas filas (WM + T-C-D):

- La cooperación entre ambos elementos, operadores y funciones de pertenencia, se manifiesta presentando la mejor precisión del estudio experimental.

Tabla 5. Resumen del test estadístico de Tamhane para el problema de distribución del sistema eléctrico

	WM	WM + C-D _{DUBOIS}	WM + C-D _{DOMBI}	WM + C-D _{FRANK}	WM + T	WM + T + C-D _{DUBOIS}	WM + T + C-D _{DOMBI}	WM + T + C-D _{FRANK}	WM + T - C-D _{DUBOIS}	WM + T - C-D _{DOMBI}	WM + T - C-D _{FRANK}
WM	-	-	-	-	-	-	-	-	-	-	-
WM + C _{DUBOIS} -D	+	≈	≈	≈	-	-	-	-	-	-	-
WM + C _{DOMBI} -D	+	≈	≈	≈	-	-	-	-	-	-	-
WM + C _{FRANK} -D	+	≈	≈	≈	-	-	-	-	-	-	-
WM + T	+	+	+	+	≈	-	-	-	-	-	-
WM + T + C _{DUBOIS} -D	+	+	+	+	+	≈	≈	≈	-	-	-
WM + T + C _{DOMBI} -D	+	+	+	+	+	≈	≈	≈	-	-	-
WM + T + C _{FRANK} -D	+	+	+	+	+	≈	≈	≈	-	-	-
WM + T- C _{DUBOIS} -D	+	+	+	+	+	+	+	+	≈	≈	≈
WM + T- C _{DOMBI} -D	+	+	+	+	+	+	+	+	≈	≈	≈
WM + T- C _{FRANK} -D	+	+	+	+	+	+	+	+	≈	≈	≈

4.3.2 ESTUDIO ESTADÍSTICO

Para comparar los resultados obtenidos con los diferentes modelos y confirmar que las mejoras encontradas son significativas, se ha realizado el estudio estadístico que se describe a continuación. En primer lugar se realizó el análisis ANOVA de un factor [2] para cada modelo, siendo el factor los modelos utilizados en los test de los conjuntos de datos. Por extensión, no se han incluido las tablas en este trabajo. Dado que se encontraron diferencias significativas para todos los modelos con respecto a los valores medios asociados, se realizó el test de Tamhane [2] con un coeficiente de confianza del 95%, debido a que la hipótesis de igualdad de varianzas ha sido rechazada en los análisis de todos los modelos.

La Tabla 5 resume los resultados obtenidos en el test de Tamhane para cada aplicación. El signo (+) significa que el modelo de la fila seleccionada mejora al modelo de la columna seleccionada, mientras que el signo (-) significa lo contrario. El signo (≈) significa que son similares. Las tablas deben leerse comenzando por la fila, por ejemplo: *en la Tabla 6, el modelo WM + T-C_{DUBOIS}-D (fila 9) mejora al modelo WM + T (columna 3).*

5. CONCLUSIONES

En este trabajo se ha propuesto un modelo que consigue la cooperación entre las funciones de pertenencia descritas en la BD y los operadores difusos basada en el aprendizaje evolutivo de ambos elementos a la vez. Las descripciones de las etiquetas lingüísticas y la forma en la que deben trabajar los operadores difusos, conjunción y método de defuzzificación para cada regla, son elementos distintos, y el hecho de encontrarse en un mismo cromosoma del algoritmo evolutivo sólo busca la cooperación en el aprendizaje de ambos elementos.

Este trabajo pues pone de manifiesto de nuevo que en el marco del compromiso entre precisión e interpretabilidad, el ajuste de los operadores difusos es también un elemento valioso, y que combinado con el aprendizaje de los elementos de la BC, en este caso con las funciones de pertenencia, se obtienen mejoras significativas en la precisión gracias a la sinergia positiva que adquieren.

El estudio se ha centrado en un método para mejorar la precisión sin introducir a nuestro juicio pérdidas importantes en la interpretabilidad, si bien cuantificarla es un problema abierto y complejo. Igualmente, se tratará de ampliar el estudio con otras aplicaciones que permitan observar los efectos del modelo propuesto en otros casos.

Agradecimientos

Este trabajo está soportado por el proyecto TIN2005-08386-C05-01 del Ministerio de Educación y Ciencia y los proyectos P05-TIC-00531 y P07-TIC-03179 de la Consejería de Innovación Ciencia y Empresa de la Junta de Andalucía.

Referencias

- [1] J. Alcalá-Fdez, F. Herrera, F. Márquez, A. Peregrín. Increasing Fuzzy Rules Cooperation Based On Evolutionary Adaptive Inference Systems. *International Journal of Intelligent Systems* 22:9, Pág. 1035-1064, 2007
- [2] T. W. Anderson. An introduction to multivariate statistical analysis. *John Wiley and Sons*, 1984.
- [3] J. Casillas, O. Cordon, F. Herrera, L. Magdalena (Eds). Accuracy improvements in linguistic fuzzy modeling. *Springer-Verlag*, 2003.
- [4] O. Cordon, F. Herrera, L. Sánchez. Solving electrical distribution problems using hybrid evolutionary data analysis techniques. *Applied Intelligence* 10, Pág.5-24, 1999.
- [5] J.S. Cho, D.J. Park. Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network. *Journal of Intelligent and Fuzzy Systems*, 8, Pág. 99-100, 2000.
- [6] O. Cordon, F. Herrera. A three-stage evolutionary process for learning descriptive and approximate fuzzy logic controller knowledge bases from examples. *Int. J. of Approximate Reasoning* 17:4, Pág.369-407, 1997.
- [7] O. Cordon, F. Herrera, A. Peregrín. Applicability of the Fuzzy Operators in the Design of Fuzzy Logic Controllers. *Fuzzy Sets and Systems*. 86, Pág. 15-41, 1997.
- [8] O. Cordon, F. Herrera, F. Márquez, A. Peregrín. A Study on the Evolutionary Adaptive Defuzzification Methods in Fuzzy Modelling. *Internacional Journal of Hybrid Intelligent Systems*, 1:1, Pág. 36-48, 2004.
- [9] L.J. Eshelman. The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. *Foundations of genetic Algorithms* 1, Pág.265-283, 1991.
- [10] M. Gupta, J. Qi. Theory of T-norms and Fuzzy Inference Methods. *Fuzzy Sets and Systems*, 40, Pág. 431-450, 1991.
- [11] F. Herrera, F. Márquez, A. Peregrín. Genetic Adaptation of Rule Connectives and Conjunction Operators in Fuzzy Rule Based System: An Experimental Comparative. *Third International Conference of the European Society for Fuzzy Logic and Technology Study*. Pág. 100-104, 2003
- [12] F. Herrera, F. Márquez, A. Peregrín. Cooperación de Reglas Difusas basada en la Adapación Evolutiva de los Operadores de Conjunción y los Defuzzificadores. *Coms. del XII Congr. Español sobre Tecnologías y Lógica Fuzzy*. Pág. 231-236, 2004.
- [13] B.D. Liu, C.Y. Chen C-Y., J.Y. Tsao. Design of adaptative fuzzy logic controller based on linguistic-hedge concepts and genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, 31:1, Pág. 32-53, 2001.
- [14] F.A. Márquez, A. Peregrín y F. Herrera Cooperative Evolutionary Learning of Fuzzy Rules and Parametric Aggregation Connectors for Mamdani Linguistic Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*. 15:6 Pág. 1162-1178, 2007.
- [15] M. Mizumoto. Pictorial Representations of Fuzzy Connectives, Part I: Cases of T-norms, T-conorms and Averaging Operators. *Fuzzy Sets and Systems*, 31, Pág. 217-242, 1989.
- [16] L.X. Wang, J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Trans. on Systems, Man, and Cybernetics* 22:6, Pág.1414-1427, 1992.