Criterios para el empleo de transformaciones ortogonales en la mejora de modelos difusos. Aplicación a modelos basados en ART

M.I. Rey[1] **M. Galende**[2] **G.I. Sainz**[2,3]

[1] INDOMAUT, S.L. Pol.Ind.San Cristóbal. Valladolid (España), i.rey@indomaut.com
[2] Fundación CARTIF, Parque Tecnológico de Boecillo. Valladolid (España) {margal,gresai}@cartif.es
[3] Dpto.Ingeniería de Sistemas y Automática, E.I.Industriales Universidad de Valladolid (España), gresai@eis.uva.es

Resumen

El propósito de este trabajo es la mejora de modelos difusos a través de transformaciones ortogonales, tratando de encontrar un equilibrio entre la precisión del modelo y su interpretabilidad. El modelado difuso preciso genera modelos con una buena precisión pero no tiene en cuenta otros aspectos básicos de la lógica difusa, como la interpretabilidad de las reglas, de ahí que estos modelos contengan un excesivo número de reglas, redundancia, incoherencia, etc. Así, el procedimiento de mejora de la simplificación se realiza intentando mantener las ventajas del modelo original (precisión) y tratando de mejorar otros aspectos bajo la idea de interpretabilidad: compactitud, distinguibilidad, etc. Para ello, se utiliza una metodología para simplificar y ordenar la base de reglas del sistema difuso haciendo uso de las transformaciones ortogonales y definiendo unos criterios que rijan el proceso de mejora de los modelos difusos. Esa propuesta se ha probado en dos casos de estudio: Motor CC y BoxJenkins.

Palabras Clave: Modelado difuso, Balance Precisión-Interpretabilidad, Transformaciones Ortogonales

1 INTRODUCCIÓN

En la actualidad es muy común la aplicación de sistemas lógico difusos en problemas reales [15]. El estudio de un sistema difuso se suele hacer a través de su conjunto de reglas difusas, que nos "describen" el comportamiento de dicho sistema, teniendo en cuenta la *Precisión* y la *Interpretabilidad*.

Si vemos modelos difusos propuestos en la literatura científica, analizados bajo estos puntos de vista, la gran mayoría de ellos muestran una buena precisión pero muy mala interpretabilidad. A partir de aquí se definen dos tipos de modelado difuso [1, 5, 7]:

- Modelado difuso preciso (*Precise Fuzzy Modeling*), donde la precisión es el principal objetivo del proceso pero la interpretabilidad es mala. Este tipo de modelado es el más habitual en aplicaciones técnicas.
- Modelado difuso lingüístico (*Linguistic Fuzzy Modeling*), donde la interpretabilidad es muy buena pero la precisión es muy pobre.

En nuestro caso, partimos de modelados precisos y queremos conseguir modelados lingüísticos. Para reducir la complejidad del modelo se utilizan transformaciones ortogonales y a continuación se aplican un conjunto de criterios que se definen en este artículo, teniendo en cuenta aspectos basados en la interpretabilidad como la redundancia o incoherencia. Se considera como ejemplo un sistema neuro-difuso FasART (Fuzzy Adaptive System Adaptive Resonance Theory based) con las ventajas que tiene un algoritmo de modelado ya existente (se podría utilizar cualquier otro) y se estudian dos casos: Motor CC y Box-Jenkins.

Cabe destacar que se ha aplicado a modelos Mandani ya que, si revisamos la literatura científica vemos que la mayoría de los trabajos están basados en TSK.

El artículo está organizado del siguiente modo: en primer lugar, una breve descripción del modelado difuso junto con transformaciones ortogonales, los criterios propuestos y las definiciones de precisión e interpretabilidad. En segundo lugar una breve descripción del sistema neuro-difuso FasART. En tercer lugar la metodología usada en el artículo, la descripción de los casos de estudio y los principales resultados obtenidos. Y por último las conclusiones más importantes a las que se ha llegado.

2 MODELADO DIFUSO: PRECISIÓN VS. INTERPRETABILIDAD

Cuando se revisan los algoritmos de modelado difuso, se encuentra el *modelado difuso preciso*, generado por datos, cuyas reglas están compuestas por variables borrosas en las que prima la precisión y el *modelado difuso lingüístico*, guiado por el conocimiento de expertos, que genera modelos difusos cuyas reglas están compuestas por variables lingüísticas en las que prima la interpretabilidad. El modelado difuso preciso intenta obtener mejor nivel de interpretabilidad mejorando este aspecto a partir de las reglas difusas [4, 7], mientras que la aproximación lingüística intenta mejorar la precisión [5, 7].

En general, la mejora de la interpretabilidad está basada en la reducción y simplificación de las reglas y sus componentes. Una de las principales aproximaciones que aparecen en la literatura científica es la simplificación basada en transformaciones ortogonales que es lo que nos ocupa en este trabajo [4, 18, 19, 21]. Se consideran tres métodos de transformaciones ortogonales [11]: SVD-QR (QR Singular Value Descomposition) [20, 21], QR-P (Pivoted QR Descomposition) [20, 21] y OLS (Orthogonal Least Square) [8, 20, 21], y se incorporan varios criterios para selección de reglas que nos llevan a un modelo reducido. Para comprobar estos criterios se utilizan el error del modelo y la interpretabilidad de las reglas teniendo en cuenta la redundancia, incoherencia y similitud de las mismas.

2.1 *Transformaciones Ortogonales*

Lo principal de las transformaciones ortogonales es descubrir la dimensionalidad intrínseca de los datos [11, 18], que en el caso de modelos difusos se traduce en encontrar las reglas más relevantes del sistema. Se consideran dos grupos:

- Métodos de búsqueda de rango, donde el rango efectivo se obtiene a partir de los valores singulares de la matriz de disparo: SVD-QR [20, 21].
- Métodos de evaluación individual, donde la contribución de cada regla se evalúa a partir del orden de las mismas: OLS [8, 20, 21].

El método QR-P [20, 21] se puede considerar como un método intermedio que genera un ordenamiento de las reglas con el cálculo de SVD.

2.2 *Criterios*

Un aspecto crítico de las aproximaciones ortogonales es determinar el número adecuado de valores singulares correctos que escoge cada algoritmo y por tanto, el número de reglas asociadas.

En este trabajo se proponen tres criterios:

1. Consiste en eliminar las reglas no relevantes: se toma el conjunto de reglas del modelo y se eliminan aquellas cuyos valores singulares o varianzas sean menores a un valor umbral, $\beta_1 \leq 8\%$.

$$\begin{gathered} ValorSingular/Varianza_{Regla_j} \leq \beta_1 \\ \Rightarrow Regla_j \notin ModeloReducido \end{gathered} \tag{1}$$

2. Se trata de encontrar un "gap" entre las reglas y asegurarse de que las reglas seleccionadas tienen suficiente información: se comprueba si un valor singular o varianza es el doble del siguiente y se ve si la suma de los valores considerados es superior a un determinado valor umbral, $\beta_2 \geq 20\%$.

$$\begin{aligned} &\text{For } (Regla_1 ... Regla_n) \\ &ValorSing/Var_n \geq 2 * ValorSing/Var_{n+1} \\ &\textstyle\sum_{j=1}^{j=n} ValorSingular/Varianza_{Regla_j} \geq \beta_2 \\ &\Rightarrow Regla_1, .., Regla_n \in ModeloDifuso \end{aligned} \tag{2}$$

3. Se trata de seleccionar las reglas mas relevantes: se consideran las reglas cuyos valores singulares o varianzas son mayores a un valor umbral, $\beta_3 \geq 80\%$.

$$\begin{gathered} ValorSingular/Varianza_{Regla_j} \geq \beta_3 \\ \Rightarrow Regla_j \in ModeloReducido \end{gathered} \tag{3}$$

Los valores de estos umbrales, β_1, β_2 y β_3 están analizados en la sección de experimentos de este trabajo a partir de los valores de precisión e interpretabilidad.

2.3 *Precisión*

Para la precisión se utiliza el error cuadrático medio, medida ampliamente aceptada para medir la exactitud en problemas de modelado:

$$ECM = \frac{1}{N} \sum_{i=1}^{N} (Y_i - Y'_i)^2 \tag{4}$$

2.4 *Interpretabilidad*

En este artículo la interpretabilidad está basada en los conceptos de completitud y distinguibilidad [6, 12, 13, 14, 16], expresados como sigue:

$$\gamma_1 < Similitud(A_i, A_{i+1}) < \gamma_2$$

$$\begin{aligned} &Similitud(A_i, A_{i+1}) = \tfrac{|A_i \cap A_{i+1}|}{|A_i \cup A_{i+1}|} = \\ &= \frac{\sum_i^j [\mu_{a_i}(x_j) \wedge \mu_{a_{i+1}}(x_j)]}{\sum_i^j [\mu_{a_i}(x_j) \vee \mu_{a_{i+1}}(x_j)]} \end{aligned} \tag{5}$$

Si γ_1 es pequeña pero no cero, la completitud está garantizada y si γ_2 tiene un valor superior a 1, se obtiene una buena distinguibilidad. Por otra parte, también hay que tener en cuenta la incoherencia y la redundancia entre las reglas (en [9] se da una definición de estos índices). En el caso que nos ocupa, estos índices son usados para evaluar los criterios propuestos y conseguir una reducción de las reglas a partir de las transformaciones ortogonales. Cuando una regla tiene varios antecedentes, la similitud entre dos reglas se puede calcular de diferentes maneras: la suma normalizada de la similitud de los antecedentes (eq. 6), el máximo de la similitud de los antecedentes (eq. 7) o el mínimo de la similitud en los antecedentes (eq. 8) (expresiones también usadas cuando se usan consecuentes):

$$\begin{array}{l} S_A(R_i,R_j) = \frac{\sum_k S_{A_k}(R_i,R_j)}{NumAntecedentes} \\ \forall 1 \leq i < j \leq NumReglas \\ \forall 1 \leq k \leq NumAntecedentes \end{array} \tag{6}$$

$$\begin{array}{l} S_A(R_i,R_j) = \max_k S_{A_k}(R_i,R_j) \\ \forall 1 \leq i < j \leq NumReglas \\ \forall 1 \leq k \leq NumAntecedentes \end{array} \tag{7}$$

$$\begin{array}{l} S_A(R_i,R_j) = \min_k S_{A_k}(R_i,R_j) \\ \forall 1 \leq i < j \leq NumReglas \\ \forall 1 \leq k \leq NumAntecedentes \end{array} \tag{8}$$

Finalmente, la similitud de todas las reglas difusas se calcula como el número de pares de reglas (*Card*) cuya similitud es mayor que un determinado umbral β_S (eq. 9) o como la media aritmética de la similitud de los antecedentes (eq. 10):

$$\begin{array}{l} Similitud = \frac{Card(S_A(R_i,R_j) > \beta_S)}{(NumReglas-1)!} \\ \forall 1 \leq i < j \leq NumReglas \end{array} \tag{9}$$

$$\begin{array}{l} Similitud = Media(S_A(R_i,R_j)) \\ \forall 1 \leq i < j \leq NumReglas \end{array} \tag{10}$$

Desde el punto de vista del balance precisión-interpretabilidad, las expresiones más utilizadas son el Máximo y la Media.

2.5 *Sistema Neuro-Difuso FasART*

El modelo FasART [2, 17] es un modelo neuro-difuso basado en la teoría de resonancia adaptativa (ART): Fuzzy ARTMAP [3]. FasART introduce una equivalencia entre la función de activación de cada neurona FasART y una función de pertenencia. FasART sería el equivalente a un sistema basado en reglas difusas Mamdani con fuzificación por singleton, inferencia por producto y defuzificación a través de la media de los centros de los conjuntos difusos. Se puede encontrar una descripción completa de este modelo en [2, 17].

3 METODOLOGÍA

En este artículo se propone un método en dos pasos:

1. Generación de un modelo difuso, usando cualquiera de los algoritmos descritos en la bibliografía científica para modelado difuso preciso o lingüístico. En nuestro caso se ha utilizado el sistema neuro-difuso FasART pero se podría utilizar cualquier otro.
2. Mejora del modelo difuso, a través del conjunto de reglas y funciones de pertenencia, obteniendo un modelo más compacto y más interpretable. En este trabajo se utilizan las transformaciones ortogonales y criterios descritos en 2.2 para tal fin.

4 EXPERIMENTOS Y RESULTADOS

El método propuesto se valida en dos casos de estudio:

1. Motor eléctrico CC Admira DR-300 [1] con tres entradas (Tensión (U), ΔU, Intensidad (I_{arm})) y una salida (Velocidad de Giro (ω)).
2. Datos Box-Jenkins [2] con seis entradas ($x(k-1), x(k-2), x(k-3), y(k-1), y(k-2), y(k-3)$) y una salida ($y(k)$) según lo descrito en [10].

Para cada caso de estudio tenemos dos modelos, uno compacto y otro complejo, para estimar la bondad de la propuesta realizada con los cuatro casos. De este modo y de acuerdo con la metodología propuesta, *en primer lugar* se generan los modelos difusos para el Motor CC y Box-Jenkins, ambos generados con FasART, y en *segundo lugar* se mejoran estos modelos usando las transformaciones ortogonales y los criterios basados en interpretabilidad. Para ver esta interpretabilidad, se define la siguiente función de coste donde cada parámetro se toma normalizado:

$$\begin{array}{rl} F_1 = & Error + NumReglas + SimilitudReglas \\ & +IncoherenciaReglas + RedundanciaReglas \\ & +NumParticionesCompletas(Cobertura) \\ & = E + NR + S + I + R + C \end{array}$$

Por otra parte, los parámetros FasART considerados a la hora de generar ambos modelos son:

Caso de estudio A. Motor CC:

1. Modelo 1A, con parámetros FasART $\rho_A = \rho_B = 0.9$ y $\gamma_A = \gamma_B = 11$, y prestaciones iniciales: NR: 16, E: 0.0022, S: 0.0138, R: 0.0077, I: 0.0000 y Cobertura completa.

[1] http://www.amira.de/dr300_engl.html

[2] http://www.stat.wisc.edu/~reinsel/bjr-data/gas-furnace

2. Modelo 1B, generado con parámetros FasART $\rho_A = \rho_B = 1$ y $\gamma_A = \gamma_B = 11$, y prestaciones iniciales: NR: 505, E: 0.0045, S: 0.3142, R: 0.0000, I: 0.0000 y Cobertura completa.

Caso de estudio B. Horno de gas Box-Jenkins:

1. Modelo 2A, con parámetros FasART $\rho_A = \rho_B = 0.6$ y $\gamma_A = \gamma_B = 3$, y prestaciones iniciales: NR: 31, E: 0.0030, S: 0.0688, R: 0.0193, I: 0.0494 y Cobertura completa.
2. Modelo 2B, con parámetros FasART $\rho_A = \rho_B = 1$ y $\gamma_A = \gamma_B = 12$, y prestaciones iniciales: NR: 293, E: 0.00002, S: 0.00009, R: 0.00002, I: 0.0000 y Cobertura completa.

A partir de aquí, los mejores resultados obtenidos teniendo en cuenta transformaciones ortogonales, criterios y la función de coste son:

4.1 *CRITERIO I*

Para testear β_1 se ha tomado el rango $8\% - 20\%$ para todos los métodos de transformaciones ortogonales citados (SVD-QR, QRP y OLS), y se ha evaluado la función de coste F_1 para los dos casos de estudio. En general los mejores resultados se han obtenido para $\beta_1 = 8\%$ aunque hay alguna excepción con $\beta_1 = 16\%$:

4.1.1 Modelo 1A

- **SVD-QR:** NR: 16 $\rightarrow$ 6, E: 0.0022 $\rightarrow$ 0.0065 , S: 0.3142 $\rightarrow$ 0.2866. No Redundancia ni Incoherencia y la Cobertura es completa.
- **QRP:** NR 16 $\rightarrow$ 9, E: 0.0022 $\rightarrow$ 0.0043, S: 0.3142 $\rightarrow$ 0.3037. No Redundancia ni Incoherencia y la Cobertura es completa.
- **OLS:** NR: 16 $\rightarrow$ 5, E: 0.0022 $\rightarrow$ 0.0080, S: 0.3142 $\rightarrow$ 0.3044. La Redundancia, Incoherencia y Cobertura siguen como en el modelo inicial.

4.1.2 Modelo 1B

- **SVD-QR:** NR: 505 $\rightarrow$ 190, E: 0.0045 $\rightarrow$ 0.0076 , S: 0.0138 $\rightarrow$ 0.123, R: 0.0077 $\rightarrow$ 0.0050. No Incoherencia y Cobertura se mantiene completa.
- **QRP:** NR: 505 $\rightarrow$ 201, E: 0.0045 $\rightarrow$ 0.0070, S: 0.0138 $\rightarrow$ 0.0117, R: 0.0077 $\rightarrow$ 0.0049. No Incoherencia y Cobertura se mantiene completa.
- **OLS:** NR: 505 $\rightarrow$ 6, E: 0.0045 $\rightarrow$ 0.3191, S: 0.0138 $\rightarrow$ 0.029, R: 0.0077 $\rightarrow$ 0.0021. No Incoherencia y Cobertura completa excepto en dos casos.

4.1.3 Modelo 2A

- **SVD-QR**, NR: 31 $\rightarrow$ 19, E: 0.0030 $\rightarrow$ 0.00361, S: 0.0688 $\rightarrow$ 0.0643, R: 0.0193 $\rightarrow$ 0.0175, I: 0.0494 $\rightarrow$ 0.0467. Las particiones se mantienen completas.
- **QRP**, NR: 31 $\rightarrow$ 19, E: 0.0030 $\rightarrow$ 0.0036, S: 0.0688 $\rightarrow$ 0.0643, R: 0.0193 $\rightarrow$ 0.0175, I: 0.0494 $\rightarrow$ 0.0468. La Cobertura permanece completa.
- **OLS**, NR: 31 $\rightarrow$ 18, E: 0.0030 $\rightarrow$ 0.0060, S: 0.0688 $\rightarrow$ 0.0182, R: 0.0193 $\rightarrow$ 0.0175, I: 0.0494 $\rightarrow$ 0.0468. Las particiones siguen completas.

4.1.4 Modelo 2B

- **SVD-QR**, NR: 293 $\rightarrow$ 289, E: 0.00002 $\rightarrow$ 0.00003, S: 0.1423 $\rightarrow$ 0.1397, R: 0.00002 $\rightarrow$ 0.00002. No Incoherencia y Cobertura completa.
- **QRP**, NR: 293 $\rightarrow$ 288, E: 0.00002 $\rightarrow$ 0.00002, S: 0.00009 $\rightarrow$ 0.0001, R: 0.00002 $\rightarrow$ 0.00002. No Incoherencia y Cobertura completa.
- **OLS**, NR: 293 $\rightarrow$ 180, E: 0.00002 $\rightarrow$ 0.0252, S: 0.00009 $\rightarrow$ 0.00012, R: 0.00002 $\rightarrow$ 0.0000. No Incoherencia y Cobertura completa.

En general este criterio ha permitido simplificar los modelos difusos: El número de reglas ha disminuido y la interpretabilidad se ha incrementado. Esto supone una subida del error pero como los valores de error siguen siendo muy pequeños no se tiene en cuenta. En el caso del *Modelo 2B* los resultados han sido diferentes al resto de los modelos: el número de reglas no se ha decrementado mucho con SVD-QR y QRP mientras que con OLS los resultados son similares a los *Modelos 1A, 1B* y *2B*.

4.2 *CRITERIO II*

Este criterio es el más similar al *"gap"* entre valores singulares, o equivalentes, descritos en otros trabajos. Para los *Modelos 1A* y *1B*, la reducción en el número de reglas ha sido extremo lo cuál no está bien desde el punto de vista lingüístico. En el caso de los *Modelos 2A* y *2B* la reducción ha sido mínima. Para un rango de β_2 de $20\% - 40\%$ los mejores resultados se han obtenido para $\beta_2 = 20\%$ y son:

4.2.1 Modelo 1A

- **SVD-QR**, NR: 16 $\Rightarrow$ 1, E: 0.0022 $\Rightarrow$ 0.4761; S: 0.4000 $\Rightarrow$ 0.0000; R: 0.0417 $\Rightarrow$ 0.0000, I: 0.2083 $\Rightarrow$ 0.0000; Cobertura completa.
- **QRP**, NR: 16 $\Rightarrow$ 1, E: 0.0022 $\Rightarrow$ 0.4761; S: 0.4000 $\Rightarrow$ 0.0000; R: 0.0417 $\Rightarrow$ 0.0000, I: 0.2083 $\Rightarrow$ 0.0000; Cobertura completa.

- **OLS**, no se han obtenido resultados.

4.2.2 Modelo 1B

- **SVD-QR, QRP** no se han obtenido resultados.
- **OLS**, NR: 505 ⇒ 1, E: 0.0045 ⇒ 0.4761; S: 0.0138 ⇒ 0.0000; R: 0.0077 ⇒ 0.0000, I: 0.2083 ⇒ 0.0000; Cobertura completa.

4.2.3 Modelo 2A

Los resultados obtenidos para **SVD-QR**, **QRP** and **OLS** han sido iguales: NR: 31 ⇒ 30, E: 0.0030 ⇒ 0.0029; S: 0.0000 ⇒ 0.0000; R: 0.0417 ⇒ 0.0000, I: 0.2083 ⇒ 0.0000; Cobertura completa.

4.2.4 Modelo 2B

- **SVD-QR, QRP**, no obtenido resultados.
- **OLS**, NR: 293 ⇒ 290, E: 0.00002 ⇒ 0.00003; S: 0.0000 ⇒ 0.0000; R: 0.0000 ⇒ 0.0000, I: 0.2083 ⇒ 0.0000; Cobertura completa.

4.3 *CRITERIO III*

Para testear β_3 se ha tomado el rango $60\% - 90\%$ y los mejores resultados se han obtenido para $\beta_3 = 80\%$ donde podemos ver que ha disminuido el número de reglas aumentando la interpretabilidad con valores de error totalmente aceptables:

4.3.1 Modelo 1A

- **SVD-QR**, NR: 16 ⇒ 6, E: 0.0022 ⇒ 0.0065; S: 0.3142 ⇒ 0.2866; R: 0.0000 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa excepto una partición.
- **QRP**, NR: 16 ⇒ 7, E: 0.0022 ⇒ 0.0060; S: 0.3142 ⇒ 0.0192; R: 0.0000 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa excepto una partición.
- **OLS**, NR: 16 ⇒ 5, E: 0.0022 ⇒ 0.0080; S: 0.0415 ⇒ 0.0107; R: 0.0000 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa.

4.3.2 Modelo 1B

- **SVD-QR**, NR: 505 ⇒ 210, E: 0.0045 ⇒ 0.0063; S: 0.0138 ⇒ 0.0111; R: 0.0077 ⇒ 0.0046, I: 0.0000 ⇒ 0.0000; Cobertuta completa.
- **QRP**, NR: 505 ⇒ 236, E: 0.0045 ⇒ 0.0049; S: 0.0138 ⇒ 0.0098; R: 0.0077 ⇒ 0.0044, No Incoherencia y Cobertura completa.
- **OLS**, NR: 505 ⇒ 192, E: 0.0045 ⇒ 0.0079; S: 0.0138 ⇒ 0.0149; R: 0.0021 ⇒ 0.0019, I: 0.0077 ⇒ 0.0063; Cobertuta completa.

4.3.3 Modelo 2A

- **SVD-QR**, NR: 31 ⇒ 19, E: 0.0030 ⇒ 0.0036; S: 0.0000 ⇒ 0.0000; R: 0.0000 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa.
- **QRP**, NR: 31 ⇒ 17, E: 0.0030 ⇒ 0.0055; S: 0.1835 ⇒ 0.1770; R: 0.0000 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa.
- **OLS**, NR: 31 ⇒ 18, E: 0.0030 ⇒ 0.0064; S: 0.0000 ⇒ 0.0000; R: 0.0000 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa.

4.3.4 Modelo 2B

- **SVD-QR**, NR: 293 ⇒ 188, E: 0.0002 ⇒ 0.0276; S: 0.1508 ⇒ 0.1258; R: 0.0095 ⇒ 0.0090, I: 0.0963 ⇒ 0.0773; Cobertuta completa.
- **QRP**, NR: 293 ⇒ 204, E: 0.00002 ⇒ 0.0133; S: 0.000009 ⇒ 0.00005; R: 0.00002 ⇒ 0.0000, I: 0.0000 ⇒ 0.0000; Cobertuta completa.
- **OLS**, NR: 293 ⇒ 145, E: 0.00002 ⇒ 0.0416; S: 0.000009 ⇒ 0.00005; No R ni I y C completa.

5 CONCLUSIONES

Este trabajo define un camino a seguir para mejorar el balance precisión-interpretabilidad en modelos basados en reglas difusas, haciendo uso de transformaciones ortogonales, de conceptos de interpretabilidad y del conjunto de criterios descritos para reducir el número de reglas, criterios que son una alternativa a los gráficos en los que se busca un "gap" entre dos valores singulares, que es lo utilizado en otros trabajos.

Esta propuesta se ha aplicado para dos casos de estudio (Motor CC y Box-Jenkins) obteniendo para cada caso dos modelos con diferentes complejidades los cuales se han mejorado posteriormente con tres métodos de transformación ortogonal: SVD-QR, QR-P y OLS.

A continuación se han aplicado los tres criterios definidos de manera que en los criterios I y III, excepto en el modelo 2B, el número de reglas se ha reducido incluso más del 50% lo que lleva a que otros aspectos de la interpretabilidad como similitud, redundancia, etc, también hayan mejorado incluso se hayan eliminado en algunos casos. Esta mejora ha implicado una subida del error pero no se han considerado importante puesto que los valores de error siguen siendo muy pequeños.

Por otra parte, el criterio II ha generado peores resultados desde el punto de vista de la interpretabilidad, ya que, se han puesto unas restricciones muy severas a la hora de buscar el "gap" lo que ha hecho que la reducción del número de reglas sea extremo.

Referencias

[1] R. Alcalá, J. Alcalá-Fdez, M.J. Gacto, and F. Herrera. Hybrid learning methods to get the interpretability-accuracy trade-off in fuzzy modeling. *Soft Computing*, 10 (9):717–734, 2006.

[2] J. M. Cano Izquierdo, Y. A. Dimitriadis, E. Gómez Sánchez, and J. López Coronado. Learnning from noisy information in FasArt and Fasback neuro-fuzzy systems. *Neural Networks*, 14(4-5):407–425, May 2001.

[3] G.A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, 3(4):698–713, September 1992.

[4] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. *Accuracy Improvements in Linguistic Fuzzy Modelling*, volume 129 of *Studies in Fuzziness and SoftComputing*, chapter Accuracy Improvements to Find the Balance Interpretability-Accuracy in Fuzzy Modeling: An Overview, pages 3–24. Springer-Verlag, Berlin Heildelberg, 2003.

[5] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena. *Interpretability Issues in Fuzzy Modelling*, volume 128 of *Studies in Fuzziness and SoftComputing*, chapter Interpretability Improvements to Find the Balance Interpretability-Accuracy in Fuzzy Modeling: An Overview, pages 3–22. Springer-Verlag, Berlin Heildelberg, 2003.

[6] M.Y. Chen and D.A. Linkens. Rule-base self-generation and simplification for data-driven fuzzy models. *Fuzzy sets and systems*, (142):243–265, 2004.

[7] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena. *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*, volume 19 of *Advances in Fuzzy Systems - Applications and Theory*. World Scientific, Singapore., 2001.

[8] S. Destercke, S. Guillaume, and B. Charnomordic. Building an interpretable fuzzy rule base from data using orthogonal least squares-application to a depollution problem. *Fuzzy Sets and Systems*, 158:2078 – 2094, 2007.

[9] M. Galende, G. I. Sainz, M. J. Fuente, and A. Herreros. Interpretability-accuracy improvement in a neuro-fuzzy ART based model of a DC motor. In *Proceedings of the 17th IFAC World Congress*, pages 7034 – 7039, Seoul, Korea, 6-11 July 2008.

[10] A.E. Gaweda and J.M. Zurada. Data-driven linguistic modeling using relational fuzzy rules. *IEEE Transactions on Fuzzy Systems*, 11(1):121–134, 2003.

[11] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press (Third edition), 1996.

[12] S.H. Huang and H. Xing. Extract intelligible and concise fuzzy rules from neural networks. *Fuzzy Sets and Systems*, (132):233–243, 2002.

[13] Y. Jin. Fuzzy modelling of high-dimensional systems: Complexity reduction and interpretability improvement. *IEEE Transactions on Fuzzy Systems*, 8(2):212–221, April 2000.

[14] Y. Jin and B. Sendhoff. Extracting interpretable fuzzy rules from RBF networks. *Neural Processing Letters*, 17(2):149–164, 2003.

[15] F. Karray and C. De Silva. *Soft Computing and Intelligent Systems Design, Theory, Tools and Applications*. Addison Wesley Publising, 2004.

[16] R. Mikut, J. Jäkel, and L. Gröll. Interpretability issues in data-based learning of fuzzy systems. *Fuzzy Sets and Systems*, 150:179 – 197, 2005.

[17] G.I. Sainz Palmero, Y. Dimitriadis, J.M. Cano Izquierdo, E. Gómez Sánchez, and E. Parrado Hernández. ART based model set for pattern recognition: FasArt family. In H. Bunke and A. Kandel, editors, *Neuro-fuzzy pattern recognition*, pages 147–177. World Scientific Pub. Co., 2000.

[18] M. Setnes and R. Babuska. Rule base reduction: Some comments on the use of orthogonal transforms. *IEEE Transactions on Systems, Man and Cybernetics. Part C: Applications and reviews*, 31(2):199 – 206, May 2001.

[19] M. Setnes, R. Babuska, U. Kaymak, and H.R. Lemke. Similarity measures in fuzzy rule base simplification. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics*, 28(3):376 – 386, June 1998.

[20] M. Setnes and H. Roubos. Ga-fuzzy modeling and classification: Complexity and performance. *IEEE Transactions on Fuzzy Systems*, 6(8):509 – 522, May 2000.

[21] J. Yen and L. Wang. Simplifying fuzzy rule-based models using orthogonal transformation methods. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics*, 29(1):13 – 24, February 1999.