Mejora del equilibrio precisión-interpretabilidad en modelos difusos

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

[1] Fundación CARTIF, Parque Tecnológico de Boecillo. Valladolid (España) {margal,gresai}@cartif.es
[2] Dpto. Ingeniería de Sistemas y Automática, E. I. Industriales. Universidad de Valladolid (España) gresai@eis.uva.es

Resumen

El presente artículo afronta el problema de la búsqueda de equilibrio entre precisión e interpretabilidad en el modelado preciso de sistemas basados en reglas difusas. Tomando como base un modelado neurodifuso, muy preciso pero poco interpretable, se presenta un método de mejora de la solución cuyo objetivo es mejorar la interpretabilidad manteniendo, e incluso mejorando, la precisión.

Se repasan diversas características de interpretabilidad que deben cumplir los sistemas basados en reglas difusas y se presenta una metodología que tiene en cuenta gran parte de esas características, proponiendo diversas formas de cálculo para las mismas y analizando los resultados sobre diversos casos de estudio.

Palabras Clave: Modelado Difuso, Algoritmo Genético, Equilibrio Precisión-Interpretabilidad.

1 INTRODUCCIÓN

Actualmente el modelado de sistemas basados en reglas difusas, y más concretamente la búsqueda de interpretabilidad en dichos modelos, se ha convertido en un campo de investigación en auge [2, 3].

Inicialmente, y en función del objetivo principal que guía el proceso, existen dos tipos básicos de modelado difuso:

- **Modelado difuso lingüístico** (*Linguistic Fuzzy Modeling*), orientado a la generación de modelos difusos cuyas reglas son extraídas del conocimiento del experto, compuestas por variables lingüísticas en las que prima la interpretabilidad de la regla.
- **Modelado difuso preciso** (*Precise Fuzzy Modeling*), orientado a generar modelos cuyas reglas son aproximadas a partir de datos, compuestas por variables borrosas en las que prima la precisión. Este tipo de modelado es el más habitual en aplicaciones técnicas.

Teniendo en cuenta el Principio de Incompatibilidad de Zadeh [17], según el cual la precisión y la interpretabilidad son objetivos contradictorios, lo ideal sería entonces obtener sistemas con buenas prestaciones para ambas características.

El presente artículo propone un método en dos pasos que intenta alcanzar un equilibrio adecuado entre precisión e interpretabilidad, tomando como base el modelado difuso preciso y realizando un postprocesamiento de selección de reglas basado en algoritmos genéticos.

En el apartado 2 se muestra un listado de características de interpretabilidad propias de los sistemas difusos, en el 3 se presenta la metodología propuesta y en los apartados 4 y 5 los casos de estudio utilizados y algunos de los resultados obtenidos. Para finalizar el último apartado presenta las principales conclusiones obtenidas en la investigación y los futuros trabajos a realizar.

2 ÍNDICES DE INTERPRETABILIDAD

Definir la interpretabilidad es una tarea difícil y muy subjetiva. Sin embargo muchos autores [18, 16] han tratado el tema e intentado establecer una serie de características/propiedades que deben tener los sistemas basados en reglas difusas para que sean fáciles de interpretar. Entre otras las principales son:

- **Propiedades de las funciones de pertenencia:**
  - Cobertura (*Coverage*): cada valor del universo de discurso debe pertenecer, al menos, a un término lingüístico/función de pertenencia/conjunto difuso.
  - Sencillez (*Simplicity*): usar mínimo numero de etiquetas/funciones de pertenencia/conjuntos difusos diferentes.
  - Normalidad (*Normality*): cada término lingüístico/función de pertenencia debe alcanzar valor pleno significado (máximo valor) en al menos un elemento del universo de discurso.
  - Distinguibilidad (*Distinguishability*): cada término lingüístico debe tener un significado claro y la función de pertenencia asociada estar claramente definida, de forma que cada una de ellas se distinga de todas las demás.
  - Concisión (*Conciseness*): medida que evalúa la forma y la simetría de los conjuntos difusos.
  - No usar conjuntos difusos irrelevantes, evitando funciones de pertenencia constantes para todo el universo de discurso.
- **Propiedades de la regla (de forma individual):**
  - Reglas Mandani, en las que el consecuente es también un termino lingüístico/conjunto difuso. Estas reglas son más interpretables que las reglas TSK.
  - Sencillez (*Simplicity*): usar mínimo numero de variables (antecedentes y/o consecuentes) diferentes.
  - No usar reglas irrelevantes que se activen en todos los casos.
- **Propiedades del conjunto de reglas (de forma global):**
  - Compactitud (*Compactness*): usar el mínimo numero de reglas que asegure un buen resultado. Implica eliminar reglas redundantes.
  - Consistencia (*Consistency*): todas las posibles combinaciones de antecedentes deben tener solo un consecuente. Es decir que no debe haber reglas incoherentes.
  - Completitud (*Completeness*): cada dato debe activar al menos una regla. Para ello las reglas difusas deben cubrir todo el universo de discurso, de forma que para cualquier entrada se active al menos una regla.

3 METODOLOGIA PROPUESTA

Para intentar alcanzar un equilibrio adecuado entre precisión e interpretabilidad, se propone aquí un método en dos pasos (figura 1) en el que:

1. A partir de cualquiera de los métodos existentes se genera un modelo difuso basado en reglas.
2. Se refina y se mejoran las prestaciones del modelo difuso en cuanto a interpretabilidad, conservando un nivel de precisión adecuado.

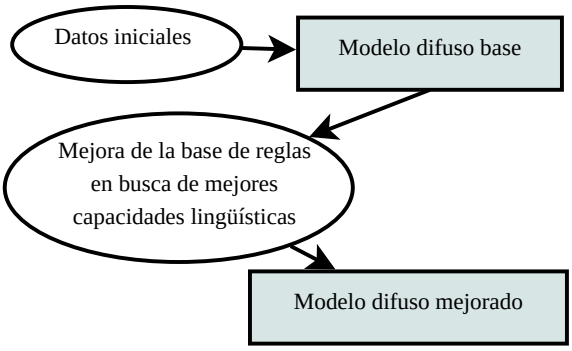

Figura 1: Proceso de mejora.

3.1 FASE I: MODELO DIFUSO BASE

En esta fase del método, a partir de datos o del conocimiento del experto, se aplica cualquier algoritmo de **modelado difuso existente** para generar un modelo difuso base que posteriormente en mejorado/refinado en la fase II.

Las ventajas de poder utilizar cualquier algoritmo de modelado ya existente son evidentes: los algoritmos ya han sido probados, generan modelos difuso precisos y existe gran variedad.

3.2 FASE II: MODELO DIFUSO MEJORADO

Para alcanzar el objetivo de obtener un sistema difuso con un buen equilibrio entre precisión e interpretabilidad en la fase II se emplea una técnica de mejora basada en **algoritmos genéticos multiobjetivo**. Concretamente se definen dos funciones objetivo que guían el proceso de selección de reglas en esta fase, una para medir la precisión y otra para medir la interpretabilidad.

3.2.1 Interpretabilidad

Teniendo en cuenta el listado de características de la sección 2 se establecen una serie de índices que per-

miten medir cuantitativamente la interpretabilidad en un sistema difuso.

- Compactitud medida a través del **numero de reglas** (ecuación 1) como parte de la simplicidad final del modelo, de forma que al disminuir las reglas aumenta la interpretabilidad.

$$Compactitud = Numero\ de\ reglas \tag{1}$$

- Distinguibilidad medida a través de la **similitud**, ya que cuanto menos similares sean entre sí las reglas más fáciles de entender por el ser humano. La similitud de dos reglas cualesquiera se calcula aplicando la medida propuesta en [15] sobre los antecedentes de las reglas (ecuación 2). Posteriormente se calcula la similitud del conjunto de reglas difusas como el promedio o el producto de la similitud en todos los antecedentes (ecuación 3).

$$\begin{aligned} &S_k(R_i,R_j) = \frac{\sum_{i,j} R_{ik}(x) \wedge R_{jk}(x)}{\sum_{i,j} R_{ik}(x) \vee R_{jk}(x)} = \frac{\sum_{i,j} \min(R_{ik},R_{jk})}{\sum_{i,j} \max(R_{ik},R_{jk})} \\ &\forall 1 \leq i < j \leq NumReglas \\ &\forall 1 \leq k \leq NumAntecedentes \end{aligned} \tag{2}$$

$$\begin{aligned} &Similitud = Promedio_{i,j}(Promedio_k(S_k(R_i,R_j))) \\ &Similitud = Producto_{i,j}(Producto_k(S_k(R_i,R_j))) \\ &\forall 1 \leq i < j \leq NumReglas \\ &\forall 1 \leq k \leq NumAntecendentes \end{aligned} \tag{3}$$

  Dicha medida ya ha sido usada con éxito por otros autores en el proceso de generación de reglas con el objetivo de reducir la complejidad de los sistemas difusos resultantes [3, 14, 11]. Sin embargo no es la única forma de calcular la similitud existiendo otras opciones que también pueden usarse [10, 12].

- Eliminación de **redundancias**, utilizando así el menor numero de reglas posibles y facilitando la interpretación final. Se dice que un par de reglas es redundante si tanto sus antecedentes como sus consecuentes son similares. Para calcular la redundancia total del conjunto de reglas se establece un umbral de redundancia β_R $(0 < \beta_R < 1)$ y se aplica la ecuación 4.

$$\begin{aligned} &Redundancia = \frac{Card(S_A(R_i,R_j) > \beta_R y S_C(R_i,R_j) > \beta_R)}{(NumReglas-1)!} \\ &\forall 1 \leq i < j \leq NumReglas \\ &\forall 1 \leq A \leq NumAntecendentes \\ &\forall 1 \leq C \leq NumConsecuentes \end{aligned} \tag{4}$$

- Eliminación de **incoherencias**, aumentando así la consistencia del modelo. Se establece un umbral de incoherencia β_I=1-β_R y se aplica la ecuación 5.

$$\begin{aligned} &Incoherencia = \frac{Card(S_A(R_i,R_j) > \beta_R y S_C(R_i,R_j) < \beta_I)}{(NumReglas-1)!} \\ &\forall 1 \leq i < j \leq NumReglas \\ &\forall 1 \leq A \leq NumAntecendentes \\ &\forall 1 \leq C \leq NumConsecuentes \end{aligned} \tag{5}$$

- Completitud o **cobertura** , de forma que si todas las entradas tienen al menos una salida definida el modelo es más interpretable (ecuación 6).

$$\begin{aligned} &NoCobertura = Promedio_k(NoCoberturaParticion_k) \\ &NoCoberturaParticion_k = \frac{PuntosSinCobertura}{PuntosTotales} \\ &\forall 1 \leq k \leq NumAntecendentes + NumConsecuentes \end{aligned} \tag{6}$$

3.2.2 Método de selección de reglas

Todos los índices anteriores se combinan en dos funciones objetivo que guían el **proceso de selección de reglas** que realiza en algoritmo genético en la fase II.

Por un lado todas las propiedades de interpretabilidad se balancean $(\lambda_i * IndiceInterpretabilidad)$ y se unen (*promedio* o *producto*) en un solo valor de interpretabilidad global del modelo difuso (ecuación 7):

$$\begin{aligned} &Interpretabilidad = f(\lambda_{nr} * NumReglas, \\ &\quad \lambda_s * Similitud, \lambda_r * Redundancia, \\ &\quad \lambda_i * Incoherencia, \lambda_c * NoCobertura) \\ &f \Rightarrow Promedio\ o\ Producto \end{aligned} \tag{7}$$

Por su parte para cuantificar la precisión del modelo basado en reglas difusas se utiliza el error cuadrático medio, medida ampliamente aceptada para medir la exactitud en problemas de modelado:

$$ECM = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i')^2 \tag{8}$$

En todos los casos se considera la diferencia normalizada de los índices propuestos (ecuación 9). Este $Indice_{nor}$ se usa directamente con el operador promedio y como $Indice_{nor} + 1$ para el producto.

$$Indice_{nor} = \frac{Indice_{Actual} - Indice_{Origen}}{Indice_{Actual}} \tag{9}$$

Los individuos se codifican a través de un string binario donde 1 indica la presencia de una regla y 0 su ausencia. Se generan poblaciones iniciales aleatorias con

al menos un individuo igual al sistema difuso inicial y se consideran individuos compuesto por subconjuntos de reglas que evolucionan independientemente.

En todos los casos los operadores de selección, cruce y mutación se ejecutan en sus opciones por defecto y el criterio de parada es el número máximo de generaciones exploradas.

4 CASOS DE ESTUDIO

El método propuesto se valida sobre dos casos de estudio:

1. Horno de gas Box-Jenkins[1] 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 el proceso descrito en [7].
2. Proceso biotecnológico de una estación de depuración de aguas residuales (EDAR) con dos entradas (Caudal de entrada a la planta y Caudal recirculado) y una salida (Substrato).

5 EXPERIMENTACIÓN

A modo de ejemplo en fase I de la metodología propuesta se emplean dos algoritmos neurodifusos de modelado bien conocidos: **FasArt** (*Fuzzy Adaptive System ART based*) [1] y **NefProx**[2] (*Neuro-Fuzzy Function Approximation*) [13]. Para cada caso de estudio generamos varios sistemas difusos iniciales, unos más compactos y otros más complejos.

Posteriormente en la fase II se utiliza el algoritmo genérico multiobjetivo **NSGA-II** [4]. Se realizan 30 ejecuciones diferentes para cada una de las alternativas consideradas con validación cruzada, codificación Gray, poblaciones con $\sqrt{NumReglas}$ individuos y un número máximo de generaciones exploradas de 100.

En cuanto a la medida de interpretabilidad propuesta (ecuación 7) se considera que todos los índices tienen la misma relevancia, por lo que $\lambda = 1$ en todos los casos. Por su parte $\beta_R = 0.8$ [3, 14].

5.1 Resultados

Por falta de espacio solo se presentan en detalle algunas de las soluciones obtenidas más importantes. De esta forma se selecciona un punto de interés del Frente de Pareto, concretamente aquel con mejor equilibrio precisión-interpretabilidad, y se calculan los valores medios de todas las ejecuciones obtenidos al utilizar el promedio en las ecuaciones 3 y 7.

5.1.1 Box Jenkins

Fase I. A partir de los datos del benckmark y mediante validación cruzada se generan dos modelos difusos con FasArt y dos modelos difusos con NefProx sobre los que realizar los experimentos:

- FasArt: BJ1 ($\rho_A = \rho_B = 0.6$ $\gamma_A = \gamma_B = 3$) y BJ2 ($\rho_A = \rho_B = 1$ $\gamma_A = \gamma_B = 12$)
- NefProx: BJ3 (NefProx $imsf/omsf = 5$ $maxrulex = 0$) y BJ4 ($imsf/omsf = 7$ $maxrulex = -1$)

La tabla 1 muestra las prestaciones iniciales de cada uno de estos modelos base. Los modelos BJ1 y BJ3 se caracterizan por ser más compacto, mientras que los modelos BJ2 y BJ4 son más complejos y difusos. Los sistemas compactos tienen un número bajo de reglas con un error pequeño mientras que los sistemas más complejos se caracterizan por tener gran número de reglas y bajo error. La similitud entre reglas puede mejorarse en la mayor parte de los casos, al igual que la redundancia y la incoherencia, y las particiones difusas de partidas suelen ser completas.

Tabla 1: Prestaciones iniciales Box-Jenkins.

	BJ1	**BJ2**	**BJ3**	**BJ4**
Nº Reglas	31	293	86	222
$Error_{tra}$	0.0030	0.00002	0.0040	0.0001
$Error_{tst}$	0.0030	0.00002	0.0040	0.0001
Similitud	0.1835	0.1420	0.3045	0.2073
Redundancia	0	0.00002	0.0216	0.0068
Incoherencia	0	0	0.0077	0.0023
Completitud	100%	100%	99.74%	100%

Fase II. La tabla 2 muestra valores medios del mejor punto de equilibrio obtenido en las ejecuciones realizadas con los modelos de Box-Jenkins. Como se puede observar la interpretabilidad mejora en todos los casos manteniendo un nivel de precisión adecuado.

Se reduce en numero de reglas entre un 5.8% y un 40.7%, la similitud entre 1.1% y 3.6%, la redundancia entre 7.2% y 100%, la incoherencia entre 10.2% y 79.5%. Las particiones difusas se mantienen completas en como mínimo un 99.7%. En cuanto al error este aumenta ligeramente en algunos casos (de 0.00002 a 0.001), pero este incremento es admisible teniendo en cuenta la mejora obtenida en la interpretabilidad y que el valor final del error sigue siendo muy pequeño.

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

[2] http://fuzzy.cs.uni-magdeburg.de/nefprox/

Tabla 2: Mejora modelos Box-Jenkins.

	BJ1	**BJ2**	**BJ3**	**BJ4**
NR	25 −19.4%	276 −5.8%	51 −40.7%	188 −15.3%
E_{tra}	0.0031 3.1%	0.001 283.3%	0.0040 −0.3%	0.001 9.2%
E_{tst}	0.0031 3.1%	0.0001 283.3%	0.0040 −0.3%	0.0001 9.2%
S	0.1769 −3.6%	0.1395 −1.8%	0.2950 −3.1%	0.2050 −1.1%
R	0 -	0 −100%	0.0173 −20.2%	0.0063 −7.2%
I	0 -	0 -	0.0016 −79.5%	0.0026 10.2%
C	100%	100%	99.7%	100%

5.1.2 EDAR

Fase I. A partir de los datos de la planta y mediante validación cruzada se generan tres modelos difusos con FasArt y uno con NefProx:

- FasArt: D1 ($\rho_A = \rho_B = 0.7$ $\gamma_A = \gamma_B = 10$), D2 ($\rho_A = \rho_B = 0.9$ $\gamma_A = \gamma_B = 10$) y D3 ($\rho_A = \rho_B = 1$ $\gamma_A = \gamma_B = 10$)
- NefProx: D4($imsf/omsf = 5$ $maxrulex = -1$)

Al igual que antes se tienen modelos más compactos (D1, D2, D4) y modelos más complejos (D3). La tabla 3 muestra las prestaciones iniciales de cada uno de los modelos base.

Tabla 3: Prestaciones iniciales EDAR.

	D1	**D2**	**D3**	**D4**
N^o Reglas	9	30	237	11
$Error_{tra}$	0.0164	0.0119	0.0109	0.0139
$Error_{tst}$	0.0172	0.0124	0.0115	0.0146
Similitud	0.3464	0.3668	0.4494	0.2896
Redundancia	0	0	0.0317	0
Incoherencia	0	0	0.0011	0
Completitud	100%	100%	100%	100%

Fase II. La tabla 4 muestra valores medios del mejor punto de equilibrio obtenido en las ejecuciones realizadas con los modelos de la EDAR. En todos los casos se obtienen sistemas que mejoran tanto la precisión como la interpretabilidad del modelo base de forma simultánea. De forma individual se mejoran todos los índices de interpretabilidad (número de reglas en 71.2%, similitud en 60.8%, redundancia en 8.3% e incoherencia en 82.3%) a excepción de la completitud que disminuye ligeramente (hasta un 82.85%). El error disminuye hasta en un 29%.

Tabla 4: Mejora modelos EDAR.

	D1	**D2**	**D3**	**D4**
NR	3.03 −66.3%	8.91 −70.3%	117.63 −50.4%	3.17 −71.2%
E_{tra}	0.0116 −29.0%	0.0110 −7.5%	0.0106 −2.4%	0.0117 −15.4%
E_{tst}	0.0123 −28.7%	0.0122 −1.2%	0.0115 0.4%	0.0123 −15.8%
S	0.3396 −1.9%	0.3606 −1.7%	0.4406 −2.0%	0.1136 −60.8%
R	0 -	0 -	0.0291 −8.3%	0 -
I	0 -	0 -	0.0002 −82.3%	0 -
C	82.85%	91.98%	95.92%	89.74%

6 CONCLUSIONES Y TRABAJO FUTURO

El presente articulo presenta una metodología en dos pasos que busca obtener sistemas difusos con un buen equilibrio precisión-interpretabilidad. Los autores ya han expuesto previamente esta metodología en [6, 5], aunque aquí se utilizan índices de interpretabilidad optimizados y otros casos de estudio.

Las principales ventajas del método en relación a otros ya existentes [8, 9] son:

- Metodología abierta, sencilla y fácilmente implantable con cualquier alternativa.
- Mejora simultanea de precisión e interpretabilidad.
- Utilización de índices más complejos para definir el concepto de interpretabilidad.
- Sistema Mandani en vez de Takagi-Sugeno.
- Problemas de modelado en vez de clasificación.

Los experimentos realizados con dos casos de estudio diferentes han demostrado que es posible obtener un mejor balance precisión-interpretabilidad tanto en modelos compactos, y por tanto más difícil de conseguir una mejora, como en modelos mucho más complejos, demostrándose que la metodología y criterios propuestos son suficientemente válidos.

Cara al futuro más inmediato están en marcha las siguientes iniciativas:

- Ampliación de la metodología a un mayor número de casos de estudio.
- Realización de test estadísticos no paramétricos.
- Mejora y refinamiento de los índices de interpretabilidad utilizados.

Referencias

[1] 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.

[2] J. Casillas, O. Cordón, Herrera F., 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 Linguistic Fuzzy Modeling: An Overview, pages 3–24. Springer-Verlag, Berlin Heilderberg, 2003.

[3] J. Casillas, O. Cordón, Herrera F., 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.

[4] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transacions on Evolutionary Computation*, 6(2):182–197, 2002.

[5] M. Galende, G. Sainz, and M.J. Fuente. Accuracy-interpretability balancong in fuzzy models based on multiobjective genetic algortihm. In *Proceedings of European Control Conference 2009 (ECC'09)*, pages 3915 – 3920, Budapest, Hungary, 23 - 26 August 2009.

[6] 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.

[7] 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.

[8] H. Ishibuchi, T. Murata, and I. B. Türksen. Single-objective and two-objective genetic algorithms for selecting linguistic rules for pattern classification problems. *Fuzzy Sets and Systems*, 89(2):135 – 150, 16 July 1997.

[9] H. Ishibuchi, T. Nakashima, and T. Murata. Three-objective genetics-based machine learning for linguistic rule extraction. *Information Sciences*, 136(1-4):109 – 133, August 2001.

[10] Y. Ji, W. Seelen, and B. Sendhoff. On generating FC^3 fuzzy rule systems from data using evolution strategies. *IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics*, 29(6):829–845, December 1999.

[11] F. Jimenez, A. F. Gómez-Skarmeta, G. Sanchez, H. Roubos, and R. Babuska. *Interpretability Issues in Fuzzy Modelling*, volume 128 of *Studies in Fuzziness and SoftComputing*, chapter Accurate, Transparent and Compact Fuzzy Models by Multi-Objective Evolutionary Algorithms, pages 431–451. Springer-Verlag, Berlin Heildelberg, 2003.

[12] Y. Jin. Fuzzy Modeling of High-Dimensional Systems: Complexity Reduction and Interpretability Improvement. *IEEE Transactions on Fuzzy Systems*, 8(2):212–221, April 2000.

[13] D. Nauck and R. Kruse. Neuro-fuzzy systems for function approximation. In *4th International Workshop Fuzzy-Neuro-Systeme '97 (FNS'97). Proceedings in Artificial Intelligence*, pages 316–323, Soest, 1997.

[14] M. Setnes. *Interpretability Issues in Fuzzy Modelling*, volume 128 of *Studies in Fuzziness and SoftComputing*, chapter Simplification and Reduction of Fuzzy Rules, pages 278–302. Springer-Verlag, Berlin Heildelberg, 2003.

[15] 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.

[16] T. Suzuki and T. Furuhashi. *Interpretability Issues in Fuzzy Modelling*, volume 128 of *Studies in Fuzziness and SoftComputing*, chapter Conciseness of Fuzzy Models, pages 569–586. Springer-Verlag, Berlin Heildelberg, 2003.

[17] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, 3(1):28 – 44, January 1973.

[18] S.-M. Zhou and J. Q. Ganb. Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets and Systems*, 159:3091 – 3131, 2008.