

# UNIVERSITAT POLITÈCNICA DE CATALUNYA



**Departament d'Enginyeria de Sistemes, Automàtica i  
Informàtica Industrial**

## **APRENDIZAJE DE PARTICIONES DIFUSAS PARA RAZONAMIENTO INDUCTIVO**

**TESIS DOCTORAL**

Jesús Antonio Acosta Sarmiento

Barcelona, Diciembre de 2006





# **APRENDIZAJE DE PARTICIONES DIFUSAS PARA RAZONAMIENTO INDUCTIVO**

MEMORIA QUE PRESENTA

**JESUS ANTONIO ACOSTA SARMIENTO**

PARA OPTAR AL GRADO DE DOCTOR

DICIEMBRE 2006

DIRECTORES

**Dra. ÀNGELA NEBOT CASTELLS    Dr. JOSEP M<sup>a</sup> FUERTES ARMENGOL**

**DEPARTAMENT D'ENGINYERIA DE SISTEMES, AUTOMÀTICA I  
INFORMÀTICA INDUSTRIAL**

**FACULTAT  
D'INFORMÀTICA DE  
BARCELONA (FIB)**

**UNIVERSITAT  
POLITÈCNICA DE  
CATALUNYA (UPC)**



La memoria titulada **Aprendizaje de particiones difusas para Razonamiento Inductivo**, que presenta D. Jesús Antonio Acosta Sarmiento para optar al grado de Doctor, ha sido realizada en el Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial (ESAI) de la *Universitat Politècnica de Catalunya* bajo la dirección de los doctores D. Àngela Nebot Castells y D. Josep M<sup>a</sup> Fuertes Armengol.

Barcelona, Diciembre de 2006.

El Doctorando

Los Directores

Fdo. J. Acosta

Fdo. A. Nebot

Fdo. J. Fuertes



Esta memoria está dedicada a mis padres  
y a mis abuelos (Q.E.P.D.)

Siempre los tengo presente y  
me dan fortaleza para superar los obstáculos.  
Están dentro de mi corazón.



## AGRADECIMIENTOS

A ***Dios Todopoderoso***, por darme energía, la sabiduría y el entendimiento necesario para lograr tan ansiada meta.

A ***mis hermanos***, por apoyarme siempre en mis planes y compartir nuestros triunfos, los quiero mucho.

A ***todos mis tíos, a la Flia. Guzmán, a la Flia. Escorriola y a la gran Flia. Arteaga***, por su confianza e interés tanto en mi formación profesional como humanística. Espero no haberles defraudado.

A ***mis primos y sobrinos***, porque espero que este logro sirva de ejemplo y ánimo en sus vidas.

A ***mis directores, Àngela Nebot y Josep M. Fuertes***, por haber confiado en mí y el esfuerzo que han tenido que dedicar para que esta memoria llegara a ser una realidad.

A ***“Perico”***, por brindarme su amistad, ayudarme y colaborar en el trabajo realizado.

A ***mis amigos y compañeros*** tanto del ***Departament d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial (ESAI)*** como del ***Departament de Llenguatges i Sistemes Informàtics (LSI)***, por todo el apoyo que me han demostrado y estar allí cuando los necesité.

A ***mis amigos y compañeros*** del ***Instituto Universitario de Tecnología “Alonso Gamero” (IUTAG)***, en especial a ***Elio Hernández, Pablo Martínez, Antonio Fuentes, Ibeth Sivila e Irelia Delgado***, por apoyar mi plan de formación en el extranjero y confiar en mí en todo momento.

A ***Juan G., Carla, Doris, Daniel, Carlos A., Nohelwin, Yenny, Izaymar, Maria, Osmar, Juan Z., Antonio, Carlos R., Evelyn P., José Q., Keyla, Evelyn R., Pedro y Rafael***, quienes en algunos casos han sido mucho más que amigos, los aprecio.

A ***mis demás familiares y a la gente que sinceramente me aprecia (de corazón)***, que sin sus palabras de aliento no hubiese tenido las fuerzas y la voluntad suficiente para alcanzar este triunfo tan ansiado en mi vida.

*Gracias a Todos*



# Indice General

	Pág.
<b>Indice de Tablas</b>	ix
<b>Indice de Figuras</b>	xviii
<b>Lista de Acrónimos</b>	xxiii
<b>Introducción</b>	1
<b>CAPITULO I. Razonamiento Inductivo Difuso</b>	
1.1 Fusificación.....	12
1.2 Modelado cualitativo.....	14
1.3 Simulación cualitativa.....	18
1.4 Defusificación.....	21
<b>CAPITULO II. Antecedentes y Estado del Arte</b>	
2.1 Trabajos previos.....	25
2.2 Modelado e identificación de sistemas	
2.2.1 Tipos de modelos matemáticos.....	34
2.2.2 Técnicas de Soft Computing usadas para la identificación de sistemas.....	34
2.3 Sistemas Basados en Reglas Difusas Descriptivos	
2.3.1 Base de conocimiento.....	36
2.3.2 Interfaz de fusificación.....	37
2.3.3 Sistema de inferencia.....	38
2.3.4 Interfaz de defusificación.....	39
2.4 Sistemas neurodifusos	
2.4.1 Taxonomía de los SNDs.....	41
2.4.2 Ventajas de los SNDs.....	42
2.4.3 Limitaciones de los SNDs.....	43
2.5 Algoritmos genéticos	
2.5.1 Representación de las soluciones.....	48
2.5.2 Mecanismo de selección.....	48
2.5.3 Operador de cruce.....	50
2.5.4 Operador de mutación.....	52
2.5.5 Características principales de los AGs.....	53
<b>CAPITULO III. Un Método de Aprendizaje de la Granularidad para Razonamiento Inductivo</b>	
3.1 Algoritmo genético para encontrar de forma automática la granularidad: AG1.....	57

	Pág.
3.2 Aspectos principales	
3.2.1 Codificación de las soluciones.....	60
3.2.2 Población inicial.....	60
3.2.3 Evaluación de los cromosomas.....	61
3.2.4 Operadores genéticos	
3.2.4.1 Selección.....	62
3.2.4.2 Cruce.....	63
3.2.4.3 Mutación.....	63
3.3 Comentarios.....	64
<b>CAPITULO IV. Un Método de Aprendizaje de las Funciones de Pertenencia para Razonamiento Inductivo</b>	
4.1 Algoritmo genético para encontrar de manera automática la forma de las funciones de pertenencia con un número fijo de clases por variable: AG2.....	68
4.2 Aspectos principales	
4.1.1 Codificación de las soluciones.....	69
4.1.2 Población inicial.....	70
4.1.3 Evaluación de los cromosomas.....	71
4.1.4 Operadores genéticos	
4.1.4.1 Selección.....	71
4.1.4.2 Cruce.....	71
4.1.4.3 Mutación.....	72
4.3 Comentarios.....	72
<b>CAPITULO V. Un Método de Aprendizaje Conjunto de la Granularidad y Funciones de Pertenencia para Razonamiento Inductivo</b>	
5.1 Algoritmo genético para encontrar de manera automática conjuntamente la granularidad y la forma de las funciones de pertenencia: AG3.....	75
5.2 Aspectos principales	
5.2.1 Codificación de las soluciones.....	76
5.2.2 Población inicial.....	77
5.2.3 Evaluación de los cromosomas.....	78
5.2.4 Operadores genéticos	
5.2.4.1 Selección.....	78
5.2.4.2 Cruce.....	78
5.2.4.3 Mutación.....	79
5.2.4.3 Reinicialización.....	80
5.3 Comentarios.....	80

**CAPITULO VI. Resultados Experimentales**

6.1 Caso de estudio: Sistema nervioso central humano	
6.1.1 Descripción del problema.....	86
6.1.2 Trabajos previos.....	89
6.1.3 Modelos obtenidos por los <i>métodos AG1, AG2 y AG3</i> .....	90
6.1.4 Análisis de resultados y conclusiones.....	104
6.2 Caso de estudio: Costos óptimos de la línea de media tensión en núcleos urbanos	
6.2.1 Descripción del problema.....	108
6.2.2 Trabajos previos.....	110
6.2.3 Modelos obtenidos por los <i>métodos AG1, AG2 y AG3</i> .....	113
6.2.4 Análisis de resultados y conclusiones.....	124
6.3 Caso de estudio: Concentraciones de ozono en Austria	
6.3.1 Descripción del problema.....	129
6.3.2 Trabajos previos.....	132
6.3.3 Modelos obtenidos por los <i>métodos AG1, AG2 y AG3</i> .....	137
6.3.4 Análisis de resultados y conclusiones.....	169

**CAPITULO VII. Problema de Predicción de Concentraciones de Ozono en la Ciudad de México**

7.1 Descripción del problema.....	182
7.2 Trabajos previos.....	187
7.3 Modelos obtenidos por los <i>métodos AG1, AG2 y AG3</i> .....	195
7.4 Análisis de resultados y conclusiones.....	215

**CAPITULO VIII. Una Aportación Adicional. Método Multiobjetivo de Aprendizaje para Razonamiento Inductivo**

8.1 Acerca de los algoritmos genéticos multiobjetivo.....	225
8.2 Generación de particiones difusas adecuadas usando algoritmos genéticos multiobjetivo ( <i>AGMO3</i> ).....	237
8.3 Aspectos principales del método propuesto	
8.3.1 Codificación de las soluciones.....	238
8.3.2 Población inicial.....	239
8.3.3 Evaluación de los cromosomas.....	239
8.3.4 Operadores genéticos	
8.3.4.1 Selección.....	240
8.3.4.2 Cruce.....	240
8.3.4.3 Mutación.....	241
8.4 Comentarios.....	242

	Pág.
8.5 Caso de estudio: Concentraciones de ozono en la Ciudad de México.	242
<b>Conclusiones</b> .....	253
<b>Referencias Bibliográficas</b> .....	262

## Indice de Tablas

N°	Descripción	Pág.
1	Resumen de las principales referencias consideradas sobre SBRDEs.....	31
2	Comparación entre las técnicas de RNA y Lógica Difusa.....	40
3	Errores de predicción MSE de los modelos del SNC usando metodologías NARMAX, TDNN y RNN (valor medio de los 6 conjuntos de datos de prueba para cada controlador).....	90
4	Parámetros de los AGs para el problema de modelado del sistema nervioso central (SNC) humano.....	91
5	Resultados de la granularidad para el controlador RC usando las funciones objetivo: a) 1-Q y, b) Predicción del 25% de la señal de entrenamiento ( $MSE_{entr}$ ). Método EFP para las funciones de pertenencia. (* Solución óptima).....	92
6	Resultados de la granularidad para el controlador TV y RP usando las funciones objetivo: a) 1-Q y, b) Predicción del 25% de la señal de entrenamiento ( $MSE_{entr}$ ). Método EFP para las funciones de pertenencia. (* Solución óptima).....	94
7	Resultados de la granularidad para el controlador CM y RCA usando las funciones objetivo: a) 1-Q y, b) Predicción del 25% de la señal de entrenamiento ( $MSE_{entr}$ ). Método EFP para las funciones de pertenencia. (* Solución óptima).....	95
8	Resultados de las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1).....	97
9	Resultados de las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo $MSE_{entr}$ (Número de clases por variable suministrado por el AG1).....	98
10	Resultados de la granularidad y las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo 1-Q....	101
11	Resultados de la granularidad y las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo $MSE_{entr}$ .....	102

Nº	Descripción	Pág.
12	MSE <sub>prue</sub> de los mejores resultados obtenidos (desde la perspectiva de las funciones objetivo) usando: AG1 con ambas funciones objetivo, FIR DEFECTO y otras metodologías inductivas.....	105
13	MSE <sub>prue</sub> de los mejores resultados obtenidos (desde la perspectiva de las funciones objetivo) usando: AG3, AG1+AG2, AG1+EFP, FIR DEFECTO y otras metodologías inductivas. Se muestran únicamente los resultados de los AG con función objetivo MSE <sub>entr</sub> .....	107
14	Notación considerada para las variables del problema de estimación de los costos de mantenimiento de la línea de media tensión.....	110
15	Errores de predicción ECM de los modelos obtenidos del problema de estimación de costos de mantenimiento de la línea de media tensión usando tanto métodos clásicos como técnicas híbridas evolutivas de análisis de datos.....	113
16	Parámetros de los AGs para el problema de estimación de costos óptimos de mantenimiento de la línea de media tensión en núcleos urbanos.....	114
17	Resultados de granularidad para el problema de estimación de costos óptimos de mantenimiento usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 20% de datos del conjunto de entrenamiento (MSE <sub>entr</sub> ). Método EFP. (X significa cualquier valor entre [2..9]).....	115
18	Resultados de las funciones de pertenencia para el problema de estimación de costos óptimos de mantenimiento usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1).....	117
19	Resultados de las funciones de pertenencia para el problema de estimación de costos óptimos de mantenimiento usando la función objetivo MSE <sub>entr</sub> (Número de clases por variable suministrado por el AG1).....	118
20	Resultados de la granularidad y las funciones de pertenencia para el problema de estimación de costos de mantenimiento de la línea de media tensión usando la función objetivo 1-Q.....	122
21	Tiempo consumido por el AG3 para realizar las treinta ejecuciones para cada criterio de parada. Funciones objetivo 1-Q y MSE <sub>entr</sub> .....	121
22	Resultados de la granularidad y las funciones de pertenencia para el problema de estimación de costos de mantenimiento de la línea de media tensión usando la función objetivo MSE <sub>entr</sub> .....	123

N°	Descripción	Pág.
23	ECM <sub>prue</sub> de los mejores resultados obtenidos (desde la perspectiva de las funciones objetivo) usando: AG3, AG1+AG2, AG1+EFP, FIR DEFECTO y otros métodos (clásicos + híbridos evolutivos).....	128
24	Modelos neuronales obtenidos por [WW99] para la predicción de concentraciones locales máximas de ozono en Austria. <b>Partición Inicial</b> ...	135
25	Máscaras FIR para la predicción de concentraciones locales máximas de ozono usando la granularidad (3,2,2,2) buscada de manera manual. Funciones de pertenencia uniformes (Método EFP). <b>Partición Inicial</b> .....	136
26	Máscaras FIR para la predicción de concentraciones locales máximas de ozono usando la granularidad (3,2,2,2) buscada de manera manual. Funciones de pertenencia uniformes (Método EFP). <b>Particiones: Caso 1 y Caso 2</b> .....	136
27	Modelos neuronales obtenidos por [GNRA+03] para la predicción de concentraciones locales máximas de ozono. <b>Partición Inicial</b> .....	137
28	Parámetros de los AGs para el problema de estimación de las concentraciones locales máximas de ozono en Austria.....	137
29	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable obtenida manualmente en [GNRA+03]). <b>Partición Inicial</b> .....	140
30	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable obtenida manualmente en [GNRA+03]). <b>Partición: Caso 1</b> .....	141
31	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable obtenida manualmente en [GNRA+03]). <b>Partición: Caso 2</b> .....	142
32	Tiempo consumido por el MANUAL + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Distribuciones de Datos (Ozono Austria): <b>Partición Inicial, Caso 1 y Caso 2</b> . Funciones objetivo: 1-Q y MSE <sub>entr</sub> .....	143

Nº	Descripción	Pág.
33	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo $MSE_{entr}$ (Número de clases por variable obtenida manualmente en [GNRA+03]). <b>Partición Inicial</b> .....	144
34	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo $MSE_{entr}$ (Número de clases por variable obtenida manualmente en [GNRA+03]). <b>Partición: Caso 1</b> .....	145
35	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo $MSE_{entr}$ (Número de clases por variable obtenida manualmente en [GNRA+03]). <b>Partición: Caso 2</b> .....	146
36	Resultados de la granularidad para la estimación de la concentración de ozono en Austria usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 8% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. <b>Partición Inicial</b> . (X significa cualquier valor entre [ 2..9]).....	149
37	Resultados de la granularidad para el problema de estimación de la concentración de ozono en Austria usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 8% de los datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. <b>Partición Caso 1</b> . (X significa cualquier valor entre [ 2..9]).....	150
38	Resultados de la granularidad para el problema de estimación de la concentración de ozono en Austria usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 8% de los datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. <b>Partición Caso 2</b> . (X significa cualquier valor entre [ 2..9]).....	151
39	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). <b>Partición Inicial</b> .....	152
40	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). <b>Partición Caso 1</b> .....	153

N°	Descripción	Pág.
41	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). <b>Partición Caso 2</b> .....	154
42	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo $MSE_{entr}$ (Número de clases por variable suministrado por el AG1). <b>Partición Inicial</b> .....	155
43	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo $MSE_{entr}$ (Número de clases por variable suministrado por el AG1). <b>Partición Caso 1</b> .....	156
44	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo $MSE_{entr}$ (Número de clases por variable suministrado por el AG1). <b>Partición Caso 2</b> .....	157
45	Tiempo consumido por el AG1 + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Distribuciones de Datos (Ozono Austria): <b>Partición Inicial, Caso 1 y Caso 2</b> . Funciones objetivo: 1-Q y $MSE_{entr}$ .....	159
46	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo 1-Q. <b>Partición Inicial</b> .....	161
47	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo 1-Q. <b>Partición Caso 1</b> .....	162
48	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo 1-Q. <b>Partición Caso 2</b> .....	163
49	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo $MSE_{entr}$ . <b>Partición Inicial</b> .....	164
50	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo $MSE_{entr}$ . <b>Partición Caso 1</b> .....	165

Nº	Descripción	Pág.
51	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo $MSE_{entr}$ . <b>Partición Caso 2</b> .....	166
52	Tiempo consumido por el AG3 para realizar las treinta ejecuciones por cada criterio de parada. Distribuciones de Datos (Ozono Austria): <b>Partición Inicial, Caso 1 y Caso 2</b> . Funciones objetivo: 1-Q y $MSE_{entr}$ .....	168
53	$ECM_{prue}$ de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para la <b>Partición Inicial</b> usando: AG1+EFP, MANUAL+AG2, AG1+AG2, AG3, FIR DEFECTO, FIR MANUAL y otros métodos basados en RN.....	172
54	$ECM_{prue}$ de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para la <b>Partición Caso 1</b> usando: AG1+EFP, MANUAL+AG2, AG1+AG2, AG3, FIR DEFECTO, FIR MANUAL.....	174
55	$ECM_{prue}$ de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para la <b>Partición Caso 2</b> usando: AG1+EFP, MANUAL+AG2, AG1+AG2, AG3, FIR DEFECTO, FIR MANUAL.....	176
56	Relación de parámetros registrados por las estaciones del SIMAT, 2006....	186
57	Interpretación del IMECA.....	187
58	Landmarks usados por [NM06] en el proceso de fusificación para las variables de entrada y de salida.....	189
59	Máscaras obtenidas por [NM06] para el modelo mensual <b>Enero</b> .....	189
60	Máscaras obtenidas por [NM06] para el modelo mensual <b>Agosto</b> .....	192
61	Parámetros de los AGs para el problema de la estimación de concentraciones locales máximas de ozono en México DF.....	195
62	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable proporcionado por expertos en [NM06]). <b>Mes Enero</b> .....	198
63	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable proporcionado por expertos en [NM06]). <b>Mes Agosto</b> .....	199

N°	Descripción	Pág.
64	Tiempo consumido por el EXPERTOS + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Caso Estudio: Ozono México- <b>Mes Enero</b> . Funciones objetivo: 1-Q y $MSE_{entr}$ .....	197
65	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo $MSE_{entr}$ (Número de clases por variable proporcionado por expertos en [NM06]). <b>Mes Enero</b> .....	200
66	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo $MSE_{entr}$ (Número de clases por variable proporcionado por expertos en [NM06]). <b>Mes Agosto</b> .....	201
67	Resultados de granularidad para la estimación de la concentración de ozono en México usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 25% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. <b>Mes Enero</b> . (X significa cualquier valor entre [ 2..9]).....	202
68	Resultados de granularidad para la estimación de la concentración de ozono en México usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 25% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. <b>Mes Agosto</b> . (X significa cualquier valor entre [ 2..9]).....	203
69	Tiempo consumido por el AG1 para realizar las treinta ejecuciones por cada criterio de parada. Caso de Estudio: Ozono México- <b>Meses (Enero y Agosto)</b> . Funciones objetivo: 1-Q y $MSE_{entr}$ .....	204
70	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). <b>Mes Enero</b> .....	205
71	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). <b>Mes Agosto</b> .....	206
72	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo $MSE_{entr}$ (Número de clases por variable suministrado por el AG1). <b>Mes Enero</b> .....	207

Nº	Descripción	Pág.
73	Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo $MSE_{entr}$ (Número de clases por variable suministrado por el AG1). <b>Mes Agosto</b> ....	208
74	Tiempo consumido por el AG1 + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Caso Estudio: Ozono México- <b>Mes Enero</b> . Funciones objetivo: 1-Q y $MSE_{entr}$ .....	209
75	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo 1-Q. <b>Mes Enero</b> .....	210
76	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo 1-Q. <b>Mes Agosto</b> .....	212
77	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo $MSE_{entr}$ . <b>Mes Enero</b> .....	213
78	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo $MSE_{entr}$ . <b>Mes Agosto</b> .....	214
79	Tiempo consumido por el AG3 para realizar las treinta ejecuciones por cada criterio de parada. Caso Estudio: Ozono México- <b>Mes Enero</b> . Funciones objetivo: 1-Q y $MSE_{entr}$ .....	211
80	$ECM_{prue}$ de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para el <b>Mes de Enero</b> usando: AG1+EFP, EXPERTOS+AG2, AG1+AG2, AG3, FIR EXPERTOS.....	217
81	$ECM_{prue}$ de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para el <b>Mes de Agosto</b> usando: AG1+EFP, EXPERTOS+AG2, AG1+AG2, AG3, FIR EXPERTOS.....	220
82	Parámetros del AG multiobjetivo para el problema de la estimación de concentraciones locales máximas de ozono en México DF.....	243
83	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 2000 evaluaciones. Mes Enero.....	244

N°	Descripción	Pág.
84	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 4000 evaluaciones. Mes Enero.....	245
85	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 8000 evaluaciones. Mes Enero.....	246
86	Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 16000 evaluaciones. Mes Enero.....	247
87	Tiempo consumido, número de generaciones y número de soluciones no dominadas obtenidas por ejecución y criterio de parada del AG multiobjetivo. Caso Estudio: Ozono México, mes Enero.....	248

## Índice de Figuras

N°	Descripción	Pág.
1	Representación esquemática de la metodología FIR.....	11
2	Proceso de fusificación de la metodología FIR.....	12
3	Ejemplo de una partición difusa fuerte.....	14
4	Proceso de obtención de la matriz I/O (base de reglas patrón).....	18
5	Esquema del proceso de simulación cualitativa de FIR.....	19
6	Procedimiento general para llevar a cabo una identificación de sistemas.....	33
7	Estructura básica de un Sistema Basado en Reglas Difusas de tipo Mamdani.	36
8	Estructura básica de un Algoritmo Genético.....	47
9	Ejemplo de aplicación del mecanismo de selección.....	49
10	Ejemplo de aplicación del operador de cruce simple en un punto.....	50
11	Ejemplo de aplicación del operador de mutación.....	52
12	Proceso de fusificación de la metodología FIR, variable presión.....	68
13	Diagrama simplificado del modelo del sistema cardiovascular, compuesto del sistema hemodinámico (SH) y del Sistema Nervioso Central (SNC).....	86
14	Señales de entrenamiento de la entrada PSC y los controladores RCA y RP..	87
15	Señales de entrenamiento de los controladores CM, TV y RC.....	88
16	Mejor y peor predicción usando los modelos FIR para el controlador RC con granularidad (9,3). Función objetivo 1-Q (errores $mva=1.614 \times 10^{-26}\%$ ref=6.66%). Método AG1.....	93
17	Mejor y peor predicción usando los modelos FIR para el controlador RC con granularidad (2,5). Función objetivo $MSE_{entr}$ (errores $mva=6.8793 \times 10^{-29}\%$ ref=0.34%). Método AG1.....	93
18	Mejor y peor predicción usando los modelos FIR para el controlador VT con granularidad (8,7) y distribución de landmarks CSP:(0.21,0.17,0.19,0.07,0.13,0.07,0.11,0.05)-VT:(0.06,0.08,0.05,0.05,0.06,0.62,0.08). Función objetivo 1-Q (errores $mva=0.11\%$ ref=12.46%). Método AG2.....	99

N°	Descripción	Pág.
19	Mejor y peor predicción usando los modelos FIR para el controlador VT con granularidad (2,5) y distribución de landmarks CSP:(0.38,0.62)-VT:(0.11, 0.14,0.31,0.22,0.22). Función objetivo $MSE_{entr}$ (errores mva=0.05% ref=0.81%). Método AG2.....	99
20	Mejor y peor predicción usando los modelos FIR para el controlador RCA con granularidad (3,2) y distribución de landmarks: CSP:(0.72,0.23,0.05)-RCA:(0.92, 0.08). Función objetivo 1-Q (errores mva=2.69% ref=11.19%). Método AG3.....	103
21	Mejor y peor predicción usando los modelos FIR para el controlador RCA con granularidad (8,5) y distribución de landmarks: CSP:(0.10,0.12,0.10, 0.11,0.11,0.15,0.14,0.17)- RCA:(0.11,0.42,0.32,0.09,0.06). Función objetivo $MSE_{entr}$ (errores mva=9.41e-30% y ref=1.24e-29%). Método AG3.....	103
22	Señales de entrenamiento para las variables de entrada AOE, AOC y ESC....	111
23	Señales de entrenamiento para la variable de entrada LTC y la de salida CMT.....	112
24	Mejores predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento. Función objetivo 1-Q. $ECM_{prue}=5111$ y $ECM_{prue}=2805$ , respectivamente. Método AG1.....	116
25	Mejores predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento. Función objetivo $MSE_{entr}$ . $ECM_{prue}=3305$ y $ECM_{prue}=3375$ , respectivamente. Método AG1.....	116
26	Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidad (2,2,3,2,2). Mejor y peor solución de proporción de datos (20000 evaluaciones). Función objetivo 1-Q. $ECM_{prue}=2727$ y $ECM_{prue}=5094$ , respectivamente. Método AG2.....	119
27	Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidad (2,2,4,9,7). Mejor y peor solución de proporción de datos (20000 evaluaciones). Función objetivo $MSE_{entr}$ . $ECM_{prue}=2973$ y $ECM_{prue}=3005$ , respectivamente. Método AG2...	120
28	Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidades (2,8,3,2,2) y (2,8,4,2,2). Mejor y peor solución (320000 evaluaciones). Función objetivo 1-Q. $ECM_{prue}=2728$ y $ECM_{prue}=5132$ , respectivamente. Método AG3.....	124

N°	Descripción	Pág.
29	Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidades (2,2,6,7,8) y (3,2,3,5,9). Mejor y peor solución (320000 evaluaciones). Función objetivo $MSE_{entr}$ . $ECM_{prue}=2880$ y $ECM_{prue}=2970$ , respectivamente. Método AG3.....	125
30	Señales de entrenamiento para las variables de entrada TMP y NUB. <b>Partición Inicial</b> .....	133
31	Señales de entrenamiento para la variable de entrada VV y la variable de salida $O_3$ . <b>Partición Inicial</b> .....	134
32	Predicción del mejor modelo FIR para la <b>Partición Inicial</b> . Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (3,2,2,2) con $Q=0,9707$ y distribución de landmarks: TMP:(0.23,0.46, 0.31)-NUB:(0.72,0.28)-VV:(0.35,0.65)- $O_3$ :(0.94,0.06). $RMS_{prue}=9.6817$ . Método MANUAL+AG2.....	139
33	Predicción del mejor modelo FIR para la <b>Partición Caso 1</b> . Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (3,2,2,2) con $Q=0.9247$ y distribución de landmarks: TMP:(0.31,0.39,0.30)-NUB:(0.46,0.54)-VV:(0.53,0.47)- $O_3$ :(0.05,0.95). $RMS_{prue}=8.1732$ . Método MANUAL+AG2.....	143
34	Predicción del mejor modelo FIR para la <b>Partición Caso 2</b> . Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Método MANUAL+AG2. Granularidad (3,2,2,2) con $Q=0.7967$ y distribución de landmarks:TMP:(0.60,0.22,0.18)-NUB:(0.37,0.63)-VV:(0.71,0.29)- $O_3$ :(0.71, 0.29). $RMS_{prue}=9.6190$ . Método MANUAL+AG2.....	147
35	Predicción del mejor modelo FIR para la <b>Partición Inicial</b> . Problema concentraciones de ozono en Austria. Función objetivo $MSE_{entr}$ . Granularidad (3,2,2,2) con $Q=0.4851$ y distribución de landmarks: TMP:(0.36,0.45,0.19)-NUB:(0.55,0.45)-VV:(0.52,0.48)- $O_3$ :(0.65,0.35). $RMS_{prue}=8.8040$ . Método MANUAL+AG2.....	147
36	Predicción del mejor modelo FIR para la <b>Partición Caso 1</b> . Problema concentraciones de ozono en Austria. Función objetivo $MSE_{entr}$ . Granularidad (3,2,2,2) con $Q=0.6511$ y distribución de landmarks: TMP:(0.41,0.31,0.28)-NUB:(0.57,0.43)-VV:(0.40,0.60)- $O_3$ :(0.67,0.33). $RMS_{prue}=8.8942$ . Método MANUAL+AG2.....	148

N°	Descripción	Pág.
37	Predicción del mejor modelo FIR para la <b>Partición Caso 1</b> . Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (7,2,2,2) con $Q=0.5350$ empleo del EFP para determinar las funciones de pertenencia. $RMS_{prue}=8.6116$ . Método AG1.....	150
38	Predicción del mejor modelo FIR para la <b>Partición Caso 1</b> . Problema concentraciones de ozono en Austria. Función objetivo $MSE_{entr}$ . Granularidad (7,2,5,5) con $Q=0.4431$ y distribución de landmarks: TMP:(0.11,0.17,0.13,0.18,0.18,0.14,0.09)-NUB:(0.52,0.48)-VV:(0.11,0.14,0.22,0.12,0.41)-O <sub>3</sub> :(0.23,0.18,0.17,0.14,0.28). $RMS_{prue}=9.0861$ . Método AG1+AG2.....	158
39	Predicción del mejor modelo FIR para la <b>Partición Caso 2</b> . Problema concentraciones de ozono en Austria. Función objetivo $MSE_{entr}$ . Granularidad (9,2,2,3) con $Q=0.6148$ y distribución de landmarks: TMP:(0.12,0.09,0.10,0.16,0.06,0.12,0.08,0.22,0.05)-NUB:(0.64,0.36)-VV:(0.56,0.44)-O <sub>3</sub> :(0.23,0.35,0.42). $RMS_{prue}=9.6089$ . Método AG1+AG2....	158
40	Predicción del mejor modelo FIR para la <b>Partición Inicial</b> . Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (3,7,3,2) con $Q=0.7593$ y distribución de landmarks: TMP:(0.33,0.33,0.34)-NUB:(0.15,0.15,0.15,0.15,0.10,0.15)-VV:(0.33,0.33,0.34)-O <sub>3</sub> :(0.80,0.20). $RMS_{prue}=8.8196$ . Método AG3.....	167
41	Predicción del mejor modelo FIR para la <b>Partición Caso 1</b> . Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (8,5,3,2) con $Q=0.7512$ y distribución de landmarks: TMP:(0.17,0.07,0.24,0.06,0.11,0.22,0.07,0.06)-NUB:(0.20,0.20,0.20,0.20,0.20)-VV:(0.40,0.07,0.53)-O <sub>3</sub> :(0.84,0.16). $RMS_{prue}=8.8088$ . Método AG3.....	167
42	Predicción del mejor modelo FIR para la <b>Partición Inicial</b> . Problema concentraciones de ozono en Austria. Función objetivo $MSE_{entr}$ . Granularidad (3,3,4,3) con $Q=0.5847$ y distribución de landmarks: TMP:(0.13,0.23,0.64)-NUB:(0.47,0.31,0.22)-VV:(0.32,0.13,0.27,0.28)-O <sub>3</sub> :(0.05,0.66,0.29). $RMS_{prue}=8.8591$ . Método AG3.....	168
43	Predicción del mejor modelo FIR para la <b>Partición Caso 1</b> . Problema concentraciones de ozono en Austria. Función objetivo $MSE_{entr}$ Granularidad (4,2,7,3) con $Q=0.4637$ y distribución de landmarks: TMP:(0.15,0.15,0.28,0.42)-NUB:(0.68,0.32)-VV:(0.12,0.22,0.12,0.12,0.12,0.14,0.16)-O <sub>3</sub> :(0.36,0.31,0.33). $RMS_{prue}=8.6461$ . Método AG3.....	169

N°	Descripción	Pág.
44	Señales de entrenamiento para las variables de entrada HD, TMP y HR. <b>Mes Enero</b> . Nota: Los valores negativos corresponden a valores perdidos.....	190
45	Señales de entrenamiento para las variables de entrada VV, DR y la variable de salida O <sub>3</sub> . <b>Mes Enero</b> . Nota: Los valores negativos corresponden a valores perdidos.....	191
46	Señales de entrenamiento para las variables de entrada HD, TMP y HR. <b>Mes Agosto</b> . Nota: Los valores negativos corresponden a valores perdidos.....	193
47	Señales de entrenamiento para las variables de entrada VV, DR y la variable de salida O <sub>3</sub> . <b>Mes Agosto</b> . Nota: Los valores negativos corresponden a valores perdidos.....	194
48	Predicción del mejor modelo FIR para <b>Mes Enero</b> . Problema concentraciones de ozono en México. Función objetivo MSE <sub>entr</sub> . Granularidad (3,2,2,2,2,3) con Q=0.6548 y distribución de landmarks: HD:(0.36, 0.43,0.21)-HR:(0.58,0.42)-TMP:(0.43,0.57)-DV:(0.46,0.54)-VV:(0.60,0.40)-O <sub>3</sub> :(0.31,0.22,0.47). Máscara:(7,8,12,17,18). RMS <sub>prue</sub> =0.0164. Método <b>EXPERTOS+AG2</b> .....	218
49	Predicción del mejor modelo FIR para <b>Mes Agosto</b> . Problema concentraciones de ozono en México. Función objetivo MSE <sub>entr</sub> . Granularidad (3,2,2,2,2,3) con Q=0.5772 y distribución de landmarks: HD:(0.36, 0.37,0.27)-HR:(0.50,0.50)-TMP:(0.57,0.43)-DV:(0.60,0.40)-VV:(0.52,0.48)-O <sub>3</sub> :(0.32,0.33,0.35). Máscara:(5,12,13,14,18). RMS <sub>prue</sub> =0.0164. Método <b>EXPERTOS+AG2</b> .....	221
50	Frontera del Pareto para [Max Q(x), Max T(x)].....	226
51	Predicción del mejor modelo FIR para el mes Enero. Problema concentraciones de ozono en México. Granularidad (4,4,2,3,3,2) con Q=0.7636 y distribución de landmarks: HD:(0.25,0.25,0.25,0.25)-HR:(0.25, 0.25,0.25,0.25)-TMP:(0.25,0.75)-DV:(0.40,0.30,0.30)-VV:(0.33,0.33,0.34)-O <sub>3</sub> :(0.84,0.16). Máscara: (13,16,17,18). RMS <sub>prue</sub> =0.0164. Método <b>AGMO3</b> ...	249

## Lista de Acrónimos

ACDEs	Algoritmos Clustering Difusos Evolutivos
AEMOs	Algoritmos Evolutivos Multiobjetivo
AEs	Algoritmos Evolutivos
AGMOs	Algoritmos Genéticos Multiobjetivos
AGs	Algoritmos Genéticos
AIR	Aprendizaje Iterativo de Reglas
AMCM	Area Metropolitana de la Ciudad de México
AOC	Area total ocupada por la ciudad
AOE	Area ocupada por edificios
ASF	Agente de Sintonización neuronal por Refuerzo
ATM	Asynchronous Transfer Mode
BC	Base de Conocimiento
BD	Base de Datos
BRD	Base de Reglas Difusas
BRL	Base de Reglas Lingüísticas
°C	Grados Centígrados
C <sub>6</sub> H <sub>6</sub>	Benceno
CE	Computación Evolutiva
CM	Contractibilidad del Miocardio
CMT	Costo de mantenimiento de la línea de media tensión
CO	Monóxido de carbono
CO <sub>2</sub>	Dióxido de carbono
CPU	Central Processing Unit
DV	Dirección del viento
ECM	Error cuadrático medio
EES	Estrategias de evolución
EFP	Equal Frequence Partition
EMOO	Optimización Multiobjetivo Evolutiva
EN	Redes Neuronales tipo Elman Networks
ESC	Energía suministrada a la ciudad
FCM	Fuzzy C-Means
FE	Factores o funciones de escala
FIR	Razonamiento Inductivo Difuso
FP	Funciones de Pertenencia
GSPP	General Systems Problem Solver
H <sup>+</sup>	Hidrogeniones
H <sub>2</sub> S	Acido Sulfhídrico
HC	Hidrocarburos
HD	Hora del día

HNO <sub>3</sub>	Acido nítrico
HR	Humedad Relativa
H <sub>2</sub> SO <sub>4</sub>	Acido Sulfúrico
IA	Inteligencia Artificial
IMECA	Indice Metropolitano de la Calidad de Aire
°K	Grados Kelvin
Kms	Kilómetros
Kms <sup>2</sup>	Kilómetros cuadrados
LSTM	Long Short Term Memory
LTC	Suma de la longitud de todas las calles de la ciudad
m/s	Metros por segundo
MEN	Redes Neuronales tipo Modified Elman Networks
micro-GA	micro-Genetic Algorithm
MLP	Redes Neuronales Perceptrónicas multicapa
MOGA	Multiobjective Optimization Genetic Algorithm
MOMGA	Multiobjective Optimization with Messy GA
MOP	Optimización Multiobjetivo
MSE	Error cuadrático medio normalizado en porcentaje (Mean Square Error)
MSE <sub>entr</sub>	Error MSE con respecto al conjunto de datos de entrenamiento
MSE <sub>prue</sub>	Error MSE con respecto al conjunto de datos de prueba
MWh	Megavatios-hora
NO <sub>2</sub>	Bióxido de nitrógeno
NO <sub>3</sub> <sup>-</sup>	Nitratos
NO <sub>x</sub>	Oxidos de nitrógeno
NPGA	Niched Pareto GA
NSGA	Non-dominated Sorting GA
NUB	Cobertura de nubes o nivel de nubosidad
O <sub>3</sub>	Ozono
PAES	Pareto Archived Evolution Strategy
PD	Técnica de Proporción de Datos
PE	Programación Evolutiva
PESA	Pareto Envelope-based Selection Algorithm
PG	Programación Genética
PIC	Sistema de péndulo invertido sobre un carro
PM <sub>10</sub>	Partículas menores a 10 micrómetros
PM <sub>2.5</sub>	Partículas menores a 2.5 micrómetros
ppb	Partes por mil millones o por billones
ppm	Partes por mil millón
PSC	Presión Seno Carotídeo
PST	Partículas suspendidas totales
PTA	Pesetas
Q	Calidad de la máscara

RC	Resistencia Coronaria
RCA	Ritmo Cardíaco
RMS	Raíz del error medio cuadrático (Root Mean Square Error)
RNA	Redes Neuronales Artificiales
RNN	Redes Neuronales Recurrentes
RNs	Redes Neuronales
RP	Resistencia Periférica
RS	Recocido Simulado
RW-GA	Random Weights GA
SBRDDs	Sistemas Basados en Reglas Difusas Descriptivos
SBRDEs	Sistemas Basados en Reglas Difusas Evolutivos
SBRDs	Sistemas Basados en Reglas Difusas
SCH	Sistema Cardiovascular Humano
SDEs	Sistemas Difusos Evolutivos
SDs	Sistemas Difusos
SGA	Simple Genetic Algorithm
SGBRDs	Sistemas Genéticos Basados en Reglas Difusas
SH	Sistema Hemodinámico
SIMAT	Sistema Monitoreo Atmosférico
SISO	Single Input Single Output
SNC	Sistema Nervioso Central
SNDEs	Sistemas Neurodifusos Evolutivos
SNDs	Sistemas Neurodifusos
SO <sub>4</sub> <sup>=</sup>	Sulfatos
SO <sub>x</sub>	Oxidos de azufre
SPEA	Strength Pareto Evolutionary Algorithm
TDNN	Redes Neuronales Time Delay
TMP	Temperatura ambiente
TV	Tono Venoso
UVb	Radiación solar UV-B
VEGA	Vector Evaluated Genetic Algorithm
VOC	Compuestos volátiles orgánicos
VOW-GA	Variable Objective Weighting GA
VV	Velocidad del viento



# Introducción

## Planteamiento

La *Teoría de Conjuntos Difusos* y la *Lógica Difusa* [KY95, Zad65, Zim96, DP97] son una gran herramienta para manejar situaciones de incertidumbre inherentes a sistemas complejos. Cualquier sistema dinámico o estático que hace uso de estas herramientas, se considera un *Sistema Difuso* [DP97].

Los Sistemas Difusos han demostrado su habilidad para resolver diferentes tipos de problemas tales como control [Zha95, DHR93, SYZL05, SGB05, PRG+04], modelado [Ped96, SY93, TS85, Jin00, HD97a] o clasificación [MB05, Kun00, CYP96, Leo99, Jin03]. Asimismo, han sido empleados en un amplio rango de aplicaciones, por ejemplo, procesamiento de señales e imágenes [CYP96, DP97, ST99, SISM01], evaluación de riesgos [LH04, Leo99, Jin03], recuperación de información [CHL01, Miy89], procesos industriales [Jin03, Leo99, HS95, BBC+95, DO01], etc. En muchos de los casos, la clave del éxito era la incorporación del conocimiento experto humano. Sin embargo, muchas de las investigaciones realizadas en el campo durante los años 90 no contienen capacidad de aprendizaje y adaptación. En la última década, ha existido un gran interés en incluir capacidad de aprendizaje a los sistemas difusos. Esto se ha logrado por medio del desarrollo de técnicas híbridas que incluyen los sistemas difusos junto con técnicas complementarias como lo son las redes neuronales, los algoritmos evolutivos o los métodos probabilísticos.

Existe consenso entre los investigadores en que se pueden obtener sistemas más inteligentes por medio de la hibridación de metodologías de *Soft Computing*<sup>1</sup> [Bon97, Ova04], haciendo de este modo que las debilidades de unos sistemas se compensen con las bondades de otros. Los *Sistemas Neurodifusos* (SNDs) y los *Sistemas Difusos Evolutivos* (SDEs) constituyen la más notoria representatividad de sistemas híbridos dentro de *Soft Computing*. Estos sistemas mezclan los métodos de razonamiento aproximado de los sistemas difusos con las capacidades de aprendizaje de las *Redes Neuronales* (RNs) y los *Algoritmos Evolutivos* (AEs), respectivamente. Hay numerosas investigaciones en ambos temas pero los más usados han sido los SNDs [JSM97, NKK97, Ful99]. Sin embargo, se han obtenido resultados muy prometedores en los últimos años para los SDEs [CHL96, CHL97a, CHL97b, HV96, Ped97, SSZ97, Tet95, CHHM01, CPC04, CCdJH05, ACP06].

---

<sup>1</sup> El término traducido en castellano es *Computación Flexible*.

Un Sistema Difuso Evolutivo [CH95, CHHM01, HM97] es básicamente un sistema difuso robustecido por un proceso de aprendizaje basado en un AE [Bác96], en particular los *Algoritmos Genéticos* (AGs) [Gol89, Hol75, Mic96], los cuales están considerados actualmente como la técnica de búsqueda global más conocida y empleada. Este tipo de algoritmos presentan la capacidad de explorar y explotar espacios de búsqueda complejos, lo que les permite obtener soluciones muy próximas a la óptima en dichos espacios. Además, la codificación genética que emplean les permite incorporar conocimiento a priori de una forma muy sencilla y aprovecharlo para guiar la búsqueda.

En esta memoria, se proponen SDEs que mejoran la técnica de modelado y simulación denominada *Razonamiento Inductivo Difuso (FIR<sup>2</sup>)*. Se persigue aprovechar las potencialidades de los AGs para aprender los parámetros de discretización de la metodología FIR, es decir, el número de clases por variable (granularidad) y las funciones de pertenencia (landmarks) que definen su semántica. Debido al hecho que es una metodología basada en lógica difusa, la eficiencia en el modelado y predicción de FIR está influenciada de forma directa por estos parámetros de discretización.

FIR surge o emerge del enfoque del *Resolvidor de Problemas de Sistemas Generales (GSPS<sup>3</sup>)* propuesto por Klir [Kli85] y es una herramienta para el análisis de sistemas generales que permite estudiar los modos de comportamiento de los sistemas dinámicos. Realiza dos tareas principales, la primera es identificar las relaciones causales entre las variables del sistema y construir el modelo cualitativo del sistema observado; y la segunda es predecir el comportamiento futuro del sistema a partir de las observaciones pasadas y del modelo previamente identificado. Para cumplir con estas tareas, la metodología FIR cuenta con cuatro funciones básicas llamadas: *fusificación, modelado cualitativo, simulación cualitativa y defusificación*.

El proceso de *fusificación* se encarga de convertir los valores cuantitativos del sistema (datos sin procesar) en sus valores cualitativos equivalentes (datos discretizados). En el proceso de codificación difusa un valor cuantitativo se convierte en una *tripleta cualitativa*, donde el primer elemento de la tripleta es el valor de la clase, el segundo elemento el valor de pertenencia difusa y el último elemento corresponde al valor de lado. El valor de lado permite conservar en la tripleta cualitativa el conocimiento completo del valor cuantitativo original, determinando con mayor precisión donde se encuentra el valor cualitativo, esto es, a la derecha, al

---

<sup>2</sup> El término castellano Razonamiento Inductivo Difuso es una traducción del término inglés *Fuzzy Inductive Reasoning*.

<sup>3</sup> En Inglés *General Systems Problem Solver*.

centro o a la izquierda del máximo de la función de pertenencia. De esta forma no se pierde información durante el proceso de fusificación.

El proceso de *modelado cualitativo* es capaz de obtener las relaciones cualitativas entre las variables que componen el sistema, construyendo la base de reglas patrón que guiará el proceso de simulación cualitativa.

El proceso de *simulación cualitativa* realiza la predicción del comportamiento del sistema. El motor de inferencia de FIR es una especialización del método de los *k-vecinos más cercanos*, comúnmente usado en el campo de reconocimiento de patrones [Das91, DHS01, MMV03, SDI05]. La *defusificación* es simplemente el proceso inverso al de fusificación.

En los últimos años la metodología FIR ha sido aplicada a diferentes tipos de problemas (control, biomedicina, ecología, etc.), obteniendo generalmente buenos resultados [NCL96, MC94, NMG01, etc.]. Cuando se tiene acceso a la opinión de expertos del problema en estudio es recomendable que los parámetros de discretización sean determinados directamente por ellos. Sin embargo, éste no es frecuentemente el caso. En la mayoría de las ocasiones se dispone de un conjunto de datos del sistema a analizar y modelar, sin mayor información adicional. Es entonces, el investigador quien asume la tarea de establecer los parámetros del proceso de discretización. Existen trabajos en los que se prueba que estos parámetros tienen una influencia determinante en el comportamiento de los *Sistemas Basados en Reglas Difusas* (SBRDs) [CHV00, Vil00]. Por lo tanto, una discretización inadecuada conlleva muy frecuentemente a la obtención de un mal modelo del sistema, el cual no representa adecuadamente la dinámica del problema estudiado. Sin embargo, cuando no se dispone de información adicional ni del acceso a expertos, una opción ampliamente usada, es definir valores por defecto para cada uno de los parámetros, de manera que siempre se usen los mismos valores de estos parámetros para llevar a cabo la discretización.

En FIR el valor por defecto del parámetro número de clases para cada variable del sistema es tres y se emplea el método EFP<sup>4</sup> para obtener las funciones de pertenencia. Este método distribuye uniformemente las clases (etiquetas o términos lingüísticos) a lo largo del universo de discurso de la variable.

Pese a que los resultados obtenidos por FIR en diferentes aplicaciones usando valores por defecto de estos parámetros fueron buenos (desde el punto de vista de la precisión en la capacidad de predicción de los modelos obtenidos) y en muchas ocasiones mejores que los logrados por otras metodologías, la experiencia ha demostrado que en algunas de estas aplicaciones, especialmente las referidas como

---

<sup>4</sup> En Inglés *Equal Frequency Partition*.

*soft sciences*, como son las biomédicas y las ecológicas, la determinación de los parámetros necesarios en el paso de discretización es relevante para la identificación de un buen modelo que capture el comportamiento del sistema de manera precisa. En este sentido, en muchas de las aplicaciones estudiadas fue necesario seleccionar modelos FIR subóptimos con un poder de predicción mayor al obtenido por el modelo óptimo e incluso se comprobó como pequeños cambios en los valores de los landmarks producían consecuencias importantes en la capacidad de predicción de los modelos identificados, tanto óptimos como subóptimos. Por todo ello, la determinación automática de parámetros adecuados de discretización en la metodología FIR surge como una alternativa de gran interés y utilidad al uso de valores heurísticos y/o por defecto. Más aún, automatizar la selección de los valores adecuados para estos parámetros permite el uso de la metodología FIR a usuarios no expertos en modelado de sistemas ni en lógica difusa garantizándoles el mejor rendimiento de esta metodología.

## Objetivos de la tesis

El objetivo principal de esta tesis doctoral es desarrollar métodos evolutivos de aprendizaje automático de particiones difusas para la metodología FIR.

Para llevar a cabo este objetivo general se desprenden los siguientes objetivos específicos:

1. *Analizar los diferentes métodos de aprendizaje de particiones difusas*, con el fin de profundizar en el tema de estudio.
2. *Diseñar e implementar en el entorno de FIR un método de aprendizaje automático de particiones difusas usando algoritmos genéticos*, considerando para ello: a) El aprendizaje de la granularidad con las funciones de pertenencia uniformes, b) El ajuste de las funciones de pertenencia con un número fijo de clases, y c) El aprendizaje conjunto de la granularidad y de las funciones de pertenencia que definen su semántica.
3. *Probar los métodos de aprendizaje de particiones difusas propuestos sobre un conjunto de benchmarks*, que permitan analizar los resultados obtenidos en función de los resultados anteriores y en comparación con los resultados obtenidos con otras metodologías para los mismos conjuntos de prueba.
4. *Aplicar la herramienta obtenida a un caso real complejo: Predicción de concentraciones de ozono en la ciudad de México*, a objeto de estudiar el

comportamiento de los métodos de aprendizaje desarrollados en esta memoria y evaluar su eficacia a través de la comparación de los resultados obtenidos con estudios previos realizados con el mismo problema.

## **Metodología aplicada**

Para cumplir con los objetivos planteados se procedió a la elaboración de un plan de trabajo. Esta tesis doctoral se llevó a cabo en (9) etapas:

*1.- Revisión bibliográfica y documental:* En esta etapa se realizó una extensa búsqueda bibliográfica relacionada con métodos que permitían realizar particiones difusas y sus algoritmos de aprendizaje, con el fin de establecer un marco conceptual y observar las aportaciones realizadas sobre el tema en estudio.

*2.- Análisis de las diferentes técnicas de aprendizaje de particiones difusas:* Esta actividad contempló un estudio de diferentes técnicas de aprendizaje de particiones difusas con el fin discernir la aplicabilidad y funcionalidad de cada una de ellas.

*3.- Elaboración de una propuesta:* En esta etapa se determinaron las posibilidades de aplicación de los algoritmos genéticos al aprendizaje automático de particiones difusas. Asimismo, implicó el diseño y desarrollo de cada una de las alternativas propuestas, y la selección de la más adecuada. Considerando para ello: a) El aprendizaje del número de etiquetas con las funciones de pertenencia uniformes, b) El ajuste de funciones de pertenencia con un número fijo de etiquetas, y c) El aprendizaje conjunto del número de etiquetas y de la semántica de las mismas.

*4.- Implementación de las mejores alternativas:* Esta actividad comprendió la incorporación de los métodos de aprendizaje automático de particiones difusas seleccionados en el contexto de la metodología FIR a objeto de generar una herramienta que mejore la habilidad de predicción de FIR. Considerando para ello: a) El aprendizaje del número de etiquetas con las funciones de pertenencia uniformes, b) El ajuste de funciones de pertenencia con un número fijo de etiquetas, y c) El aprendizaje conjunto del número de etiquetas y de la semántica de las mismas.

*5.- Aplicación práctica:* Este paso consistió en el uso de la herramienta desarrollada sobre un conjunto de benchmarks así como también su aplicación a un problema real complejo con miras a realizar un análisis comparativo de los resultados obtenidos con otros sistemas de modelado.

6.- *Presentación y discusión de resultados*: Su objetivo fue mostrar los resultados de la investigación con el fin de realizar posteriormente su análisis, discusión e interpretación.

7.- *Elaboración de conclusiones*: Esta etapa consistió en presentar las conclusiones que se desprendieron del análisis de los resultados, considerando los aspectos más relevantes de éste y en función de los objetivos de la investigación.

8.- *Elaboración de recomendaciones*: En esta actividad se pretendió mostrar las recomendaciones que puedan surgir del proceso investigativo, con el objeto de sugerir ante cualquier limitación suscitada, a fin de que otros investigadores interesados en profundizar en el tema o replicar la investigación, pudieran realizar estudios más completos. Incluyó además, sugerencias, interrogantes o aspectos de interés develados por los resultados que pudieran ser ampliados o investigados en profundidad en otros trabajos o extensiones futuras de la tesis.

9.- *Redacción de la memoria de la tesis*: Finalmente, se procedió a la elaboración de la memoria de la tesis doctoral. En la misma se plasmaron los antecedentes del tema, las técnicas utilizadas, los métodos propuestos, los resultados y las conclusiones obtenidas, y las extensiones futuras de la tesis.

## Resumen

En la presente memoria se aborda un estudio acerca de la influencia de los parámetros de discretización, la granularidad y las funciones de pertenencia asociadas, en la eficiencia del modelado y predicción de la *Metodología de Razonamiento Inductivo Difuso* (FIR). Para ello se presentan tres métodos de aprendizaje automático de las particiones difusas:

- *Aprendizaje de la granularidad*: Este método trata de obtener modelos FIR con buen comportamiento aprendiendo un nivel de granularidad adecuado para cada una de las variables de un sistema manteniendo particiones difusas uniformes.
- *Aprendizaje de las funciones de pertenencia*: Este método persigue ajustar localmente las funciones de pertenencia dado un número de clases para cada variable.
- *Aprendizaje en conjunto de la granularidad y funciones de pertenencia*: Este método realiza un aprendizaje global de la granularidad y de las funciones de pertenencia asociadas a cada una de las variables de un sistema.

Los métodos propuestos han sido implementados en la herramienta de programación Matlab y sirven tanto para entornos Windows como para Linux.

La memoria está organizada en 8 capítulos y una sección de “Comentarios Finales”. A continuación se describen brevemente los contenidos de cada uno de ellos:

En el Capítulo I se presenta la herramienta bajo estudio en esta investigación, la *Metodología de Razonamiento Inductivo Difuso* (FIR). Se introducen nociones básicas, se establece su definición y se describe detalladamente su composición y funcionamiento.

El Capítulo II está dedicado a la descripción de los marcos de trabajo y teorías relacionadas con los *Sistemas Difusos Evolutivos* (SDEs). En un principio se hace un estudio acerca de los antecedentes y estado actual del tema. Luego, se trata sobre el modelado e identificación de sistemas. Se señalan algunas técnicas de *Soft Computing* utilizadas para llevar a cabo esa tarea, enfatizando aquellas que utilizan la lógica difusa. Seguidamente, se introducen las nociones básicas sobre composición y funcionamiento de los *Sistemas Basados en Reglas Difusas Descriptivos* (SBRDDs) y los *Sistemas Neurodifusos* (SNDs). No se incluyen los principios básicos de la Lógica Difusa, que pueden ser consultados en [KY95, Zim96]. Por último, se describe el tipo de *Algoritmos Evolutivos* (AEs) implementados en esta memoria, los *Algoritmos Genéticos* (AGs), considerados actualmente la técnica de búsqueda global más conocida y empleada.

En el Capítulo III se presenta un método de aprendizaje de la granularidad de las variables para el *Razonamiento Inductivo Difuso* (FIR). Se describen sus componentes principales y se presentan sus ventajas e inconvenientes.

En el Capítulo IV se introduce un método de aprendizaje de las funciones de pertenencia de cada variable para *Razonamiento Inductivo Difuso* (FIR) dado un nivel óptimo de granularidad, realizándose un ajuste local. Se describen sus aspectos fundamentales, así como se exhiben las ventajas e inconvenientes de su utilización.

En el Capítulo V se presenta un método de aprendizaje en conjunto de la granularidad y las funciones de pertenencia para *Razonamiento Inductivo Difuso* (FIR), permitiéndose realizar un ajuste global sobre las particiones difusas. Se especifican sus elementos principales y se señalan las ventajas e inconvenientes del mismo.

En el Capítulo VI se estudia el comportamiento de los métodos propuestos en esta memoria sobre una serie de aplicaciones y se comparan con los resultados

obtenidos en estudios previos. Las aplicaciones seleccionadas son: el sistema nervioso central humano, los costos óptimos de la línea de media tensión en núcleos urbanos y la contaminación por ozono en Austria.

En el Capítulo VII se describe una aplicación compleja y de gran interés en la actualidad, la predicción de concentraciones de ozono, en este caso en la Ciudad de México. Se presentan las características del problema, las variables que intervienen y los estudios previos realizados. Se aborda, posteriormente, su resolución mediante los distintos métodos propuestos en esta memoria.

El Capítulo VIII contiene una aportación adicional, un método multiobjetivo de aprendizaje para *Razonamiento Inductivo Difuso* (FIR). Se introduce el tema, se describen los aspectos principales que lo conforman y se presentan las ventajas e inconvenientes de esta herramienta. Posteriormente se presentan los resultados obtenidos con este método en el problema de la concentración de ozono en la ciudad de México, comparando éstos con los presentados en el Capítulo VII.

Posteriormente, se incluye un apartado de “Comentarios Finales”, que resume los resultados obtenidos en esta memoria y presenta algunas conclusiones sobre éstos. Finalmente se presenta la bibliografía que recoge las contribuciones más destacadas en la materia estudiada.

# **CAPITULO I:**

---

## **Razonamiento Inductivo Difuso**

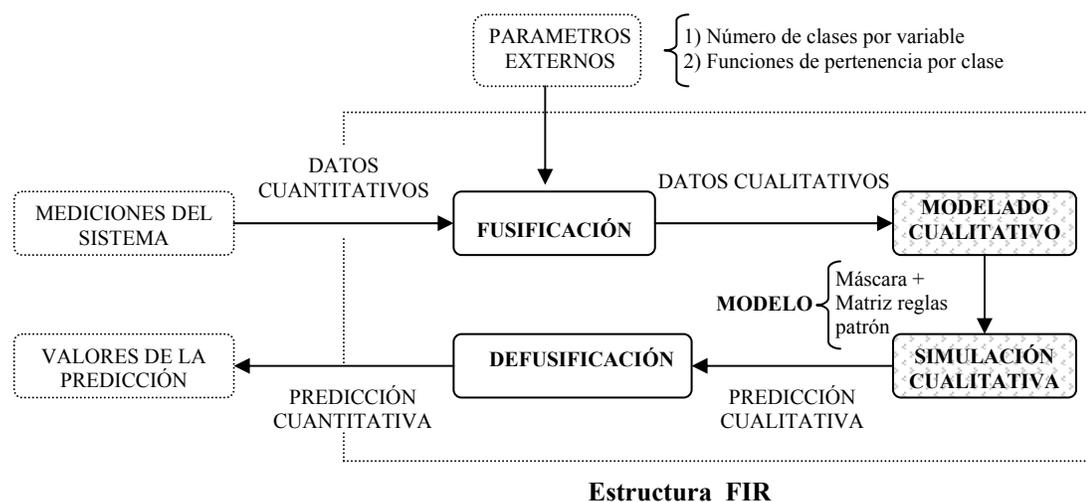


# Capítulo I: Razonamiento Inductivo Difuso

En este capítulo se presenta la entidad bajo estudio en este trabajo de investigación, la metodología de *Razonamiento Inductivo Difuso* (FIR). A continuación se dan unas nociones básicas acerca de su origen e implementación, se define su significado, y por último, se describe detalladamente su composición y funcionamiento.

El Razonamiento Inductivo Difuso (FIR) surge del enfoque del *Resolvedor de Problemas de Sistemas Generales (GSPS)* propuesto por Klir [Kli85]. Consiste en una metodología de modelado y simulación cualitativa basada en la observación del comportamiento del sistema más que en el conocimiento de su estructura interna. Actualmente FIR está implementada como una toolbox de Matlab [HL02] y recientemente se ha desarrollado la plataforma *Visual-FIR* que permite trabajar con FIR a través de un entorno amigable y de muy fácil manejo [ENC04]. Esta herramienta está pensada para el uso generalizado de FIR. Cabe señalar que *Visual-FIR* está actualmente en fase de experimentación y prueba.

Como ya se ha señalado, FIR realiza dos tareas principales. La primera es identificar las relaciones causales y temporales entre las variables del sistema para construir el modelo cualitativo del sistema observado. La segunda es predecir el comportamiento futuro del sistema a partir de las observaciones pasadas y del modelo previamente identificado. Para cumplir con estas tareas, la metodología FIR cuenta con cuatro funciones básicas llamadas: *fusificación*, *modelado cualitativo*, *simulación cualitativa* y *defusificación*.



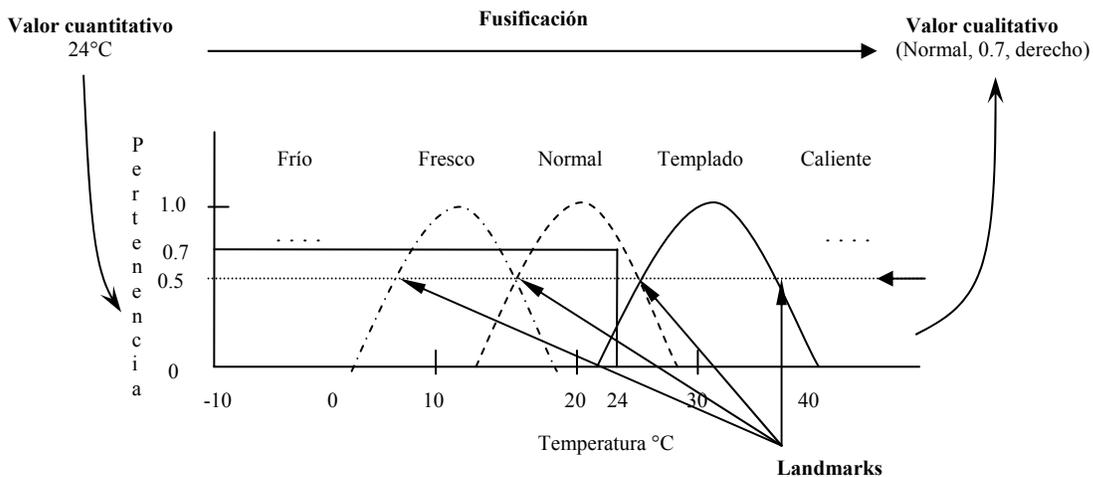
**Figura 1.** Representación esquemática de la metodología FIR

La *figura 1*, muestra las cuatro funciones de la metodología FIR de forma esquemática, en donde se somborean las dos tareas principales. Es decir, la identificación de modelos cualitativos y el uso de éstos en la simulación cualitativa, con el propósito de predecir el comportamiento futuro del sistema en estudio.

Los cuatro procesos que comprende la metodología *FIR* se describen con mayor detalle a continuación.

## 1.1 Fusificación

FIR se alimenta de datos reales registrados del sistema. El proceso de *fusificación* es el encargado de convertir los valores cuantitativos del sistema (datos sin procesar) en sus valores cualitativos equivalentes (datos discretizados). En el proceso de codificación difusa un valor cuantitativo se convierte en una tripleta cualitativa, donde el primer elemento de la tripleta es el valor de la clase, el segundo elemento el valor de pertenencia difusa y el último elemento corresponde al valor de lado. El valor de clase corresponde al valor discretizado de la variable (etiqueta o valor lingüístico). El valor de pertenencia (*membership*) permite conocer en qué grado pertenece ese valor a la clase. El valor de lado permite conservar en la tripleta cualitativa el conocimiento completo del valor cuantitativo original, determinando con mayor precisión donde se encuentra el valor cualitativo, esto es, a la derecha, al centro o a la izquierda del máximo de la función de pertenencia. De esta forma no se pierde información durante el proceso de fusificación.



**Figura 2.** Proceso de fusificación de la metodología FIR

Para ilustrar este proceso consideremos que se tiene un valor cuantitativo de 24°C que indica la temperatura del medio ambiente. La discretización de esta variable se hace en cinco clases cuyos valores cualitativos son: frío, fresco, normal, templado y caliente. Cada una de estas etiquetas lingüísticas se representa por medio de una función de pertenencia, como se muestra en la *figura 2*. Así, 24 °C corresponde a la clase “*normal*”, con un valor de pertenencia a esta clase de 0.7 y un valor de lado de “*derecho*”, puesto que se encuentra en el lado derecho respecto al valor máximo de la función de pertenencia. En FIR, el punto donde dos clases adyacentes se cortan con el valor de pertenencia de 0.5 es denominado landmark. Cabe señalar que los landmarks están determinando la amplitud de las funciones de pertenencia correspondientes. Es decir, definen de alguna manera la forma (estrecha o ancha) de estas funciones.

Una vez definidos el número de clases y las marcas entre clases (landmarks) para cada variable se lleva a cabo el proceso de *fusificación* de FIR sobre los datos registrados del sistema. Al finalizar este proceso se obtiene una tripleta de matrices con el mismo número de datos, la primera contiene los valores de clase, la segunda los valores de pertenencia y la tercera los valores de lado.

### 1.1.1 Propiedades que garantizan la interpretabilidad de una partición difusa

Suponiendo que el dominio  $U$  de la variable  $x$  se particiona en un conjunto de  $E$  etiquetas definidas con funciones de pertenencia triangulares:  $A_1, A_2, \dots, A_E$ , se definen las siguientes propiedades:

- *Completitud*, que garantiza que ningún punto del dominio  $U$  queda fuera de la partición difusa:  $\forall x \in U, \exists j$  tal que  $\mu_{A_j}(x) \geq \delta > 0$
- *Ordenación semántica*, que garantiza el orden de los términos lingüísticos de la partición:  $\forall i < j, a_i < a_j$ , siendo  $a_i$  y  $a_j$  los puntos modales de las etiquetas  $A_i$  y  $A_j$ .
- *Dominancia*, que impide que los puntos del dominio  $U$  tengan un alto grado de pertenencia  $\gamma$  a más de dos términos lingüísticos:

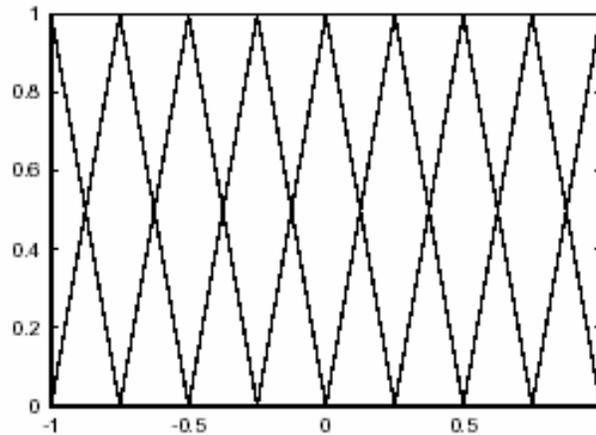
$$\forall x \in U, \text{ si } \mu_{A_j}(x) > \gamma > 0, \text{ y } \mu_{A_i}(x) > \gamma > 0 \quad (1)$$

$$\text{entonces } \forall k \in \{1, \dots, E\}, k \neq j \neq i, \mu_{A_k}(x) \leq \gamma$$

Relacionado con estas propiedades, se dice que una partición difusa es fuerte si cumple la siguiente condición:

$$\forall x \in U, \sum_{i=1}^E \mu_{A_i}(x) = 1 \quad (2)$$

Se puede decir que una partición difusa fuerte (Véase *figura 3*), también llamada partición difusa de Ruspini, posee el máximo grado de interpretabilidad. Para el caso de una partición difusa definida con funciones de pertenencia triangulares, dicha partición será fuerte si el punto de corte de todas las etiquetas se encuentra a grado 0.5. Evidentemente, una partición difusa uniforme sobre el universo de discurso de una variable es una partición fuerte.



*Figura 3. Ejemplo de una partición difusa fuerte.*

## 1.2 Modelado cualitativo

El comportamiento cualitativo está compuesto por las tres matrices obtenidas en el proceso de fusificación, en donde cada columna representa una de las variables observadas y los renglones son los registros realizados en cada intervalo de tiempo. El proceso de modelado cualitativo consiste básicamente en realizar una búsqueda sobre estas matrices para encontrar las relaciones causales y temporales más deterministas entre las variables. Estas relaciones constituyen la estructura del modelo FIR (que llamaremos *máscara óptima*). La máscara es usada para extraer del conjunto de datos de entrenamiento las reglas cualitativas basadas en patrones (o *reglas patrón*). La máscara óptima y el conjunto de reglas patrón derivado de ella constituyen el modelo FIR que representa el comportamiento del sistema. Vamos a describir de forma más detallada todo este proceso.

Como acabamos de comentar, el primer paso es obtener la estructura o máscara que mejor represente el sistema dinámico en estudio. Para ello, se evalúan todas las posibles máscaras y se concluye cual de ellas tiene el más alto poder de predicción por medio de una medida de reducción de entropía, denominada calidad de la

máscara,  $Q$ . Una máscara describe una relación dinámica entre las variables cualitativas. La máscara con el máximo valor de  $Q$ , es la máscara óptima.

Cada fila de una máscara representa un instante de tiempo en función del intervalo de muestreo. Al número de filas se le conoce como profundidad de la máscara. La máscara se representa en forma matricial como se muestra la ecuación 3 y está compuesta por:

- Los antecedentes, conocidos como *m-entradas*, que indican aquellas relaciones causales que permiten determinar el valor del consecuente, y se denotan con valores negativos.
- El consecuente, conocido como *m-salida*, que corresponde a la variable de salida que deseamos predecir y se denota con un valor positivo.
- Los componentes nulos que indican qué relaciones no son relevantes para predecir el valor del consecuente.

$$\begin{array}{c}
 t^x \\
 t - 2\delta t \\
 t - \delta t \\
 t
 \end{array}
 \begin{array}{cccc}
 u_1 & u_2 & u_3 & y_1 \\
 \left( \begin{array}{cccc}
 0 & 0 & -1 & 0 \\
 0 & -2 & 0 & 0 \\
 -3 & 0 & 0 & +1
 \end{array} \right)
 \end{array}
 \quad (3)$$

La máscara de la *ecuación 3* contiene cuatro m-entradas. En notación de posición, ésta puede ser escrita como (3,6,9,12), indicando que la primera m-entrada está situada en la tercera posición de la máscara si numeramos ésta de izquierda a derecha y de arriba abajo, que la segunda m-entrada está situada en la posición 6 de la máscara, etc. La secuencia en que las m-entradas y la m-salida son numeradas no tiene ningún significado especial. Los términos m-entrada (entrada de la máscara) y m-salida (salida de la máscara) son utilizados para evitar confusión con las entradas y las salidas del sistema. El número de elementos diferentes de cero en la máscara, indica el nivel de *complejidad* de la misma. Aquí la máscara tiene un nivel de complejidad de cuatro. En este ejemplo, la primera m-entrada ( $i_1$ ) corresponde a la variable de entrada del sistema  $u_3$  muestreada dos intervalos de tiempo anteriores al actual,  $u_3(t-2\delta t)$ , la segunda m-entrada ( $i_2$ ) corresponde a la entrada del sistema  $u_2$  muestreada en el instante de tiempo anterior al actual,  $u_2(t-\delta t)$ , etc.

Para iniciar el proceso de búsqueda se introduce el concepto de máscara candidata con la finalidad de encontrar entre todo el conjunto de máscaras posibles (de cada complejidad hasta llegar a la máxima), la que mejor represente al sistema. A partir de la matriz de máscara candidata, el proceso de búsqueda de la máscara óptima procede a evaluar todas las posibles combinaciones a través de un mecanismo

de búsqueda exhaustiva de complejidad exponencial [CNMda92] o mediante estrategias de búsqueda genética o de árboles de complejidad polinomial [JN97].

$$\begin{matrix} t^x & u_1 & u_2 & u_3 & y_1 \\ t - 2\delta t & -1 & -1 & -1 & -1 \\ t - \delta t & -1 & -1 & -1 & -1 \\ t & -1 & -1 & -1 & +1 \end{matrix} \quad (4)$$

La máscara candidata contiene elementos “-1” donde tiene una m-entrada potencial, un elemento “+1” donde tiene la m-salida, y se utiliza el “0” para denotar relaciones prohibidas o sin significado físico. Una posible máscara candidata para un sistema de cuatro variables podría ser aquella en la que se permita que FIR estudie todas las posibles relaciones causales entre las variables del sistema en los diferentes instantes de tiempo y la salida en el instante  $t$ , como se muestra en la *ecuación 4*.

La optimalidad de la máscara se evalúa con respecto a la maximización de su poder de predicción. Se usa la medida de entropía de Shannon [Sha48] para determinar la incertidumbre asociada al pronóstico de un estado de salida particular dado cualquier estado de entrada legal. La entropía de Shannon relativa a un estado de entrada se calcula con la ecuación:

$$H_i = -\sum_{\forall o} p(o|i) \cdot \log_2 p(o|i) \quad (5)$$

donde  $p(o|i)$  es la probabilidad condicional de que ocurra un cierto estado  $o$  en la m-salida, dado que ha ocurrido el estado  $i$  en el patrón de m-entradas. El término probabilidad se utiliza en un sentido más bien estadístico que en uno estrictamente probabilístico. La entropía total de la máscara se calcula como:

$$H_m = \sum_{\forall i} p(i) \cdot H_i \quad (6)$$

donde  $p(i)$  es la probabilidad de que ocurra ese estado de entrada. La entropía más alta posible  $H_{max}$  se obtiene cuando todas las probabilidades son iguales y una entropía cero corresponde a las relaciones totalmente deterministas.

El término reducción de entropía normalizada  $H_r$  se define como:

$$H_r = 1.0 - \frac{H_m}{H_{max}} \quad (7)$$

$H_r \in [0,1] \subset \mathbb{R}$ , donde valores altos indican normalmente mayor poder de predicción. Las máscaras con valores de reducción de entropía altos generan predicciones con menor incertidumbre.

Por otro lado, es importante tener en cuenta la predictibilidad de la máscara que se está evaluando. La metodología FIR considera que desde el punto de vista estadístico cada estado se debe observar por lo menos cinco veces [LK91]. Es por ello que se utiliza el ratio de observación  $O_r$ , como un factor adicional para determinar la calidad total de la máscara [LC90].

$$O_r = \frac{5 \cdot n_{5x} + 4 \cdot n_{4x} + 3 \cdot n_{3x} + 2 \cdot n_{2x} + n_{1x}}{5 \cdot n_{leg}} \quad (8)$$

$n_{leg}$  = número de estados legales de m-entradas;

$n_{1x}$  = número de estados de m-entradas observados solamente una vez;

$n_{2x}$  = número de estados de m-entradas observados dos veces;

$n_{3x}$  = número de estados de m-entradas observados tres veces;

$n_{4x}$  = número de estados legales de m-entradas observados cuatro veces;

$n_{5x}$  = número de estados de m-entradas observados cinco o más veces.

Si todo estado legal de m-entradas ha sido observado al menos cinco veces,  $O_r$  es igual a 1.0 y no tiene ninguna influencia en la calidad de la máscara. Si no se ha observado ningún estado de m-entradas (es decir, no hay datos disponibles),  $O_r$  es igual a 0.0.

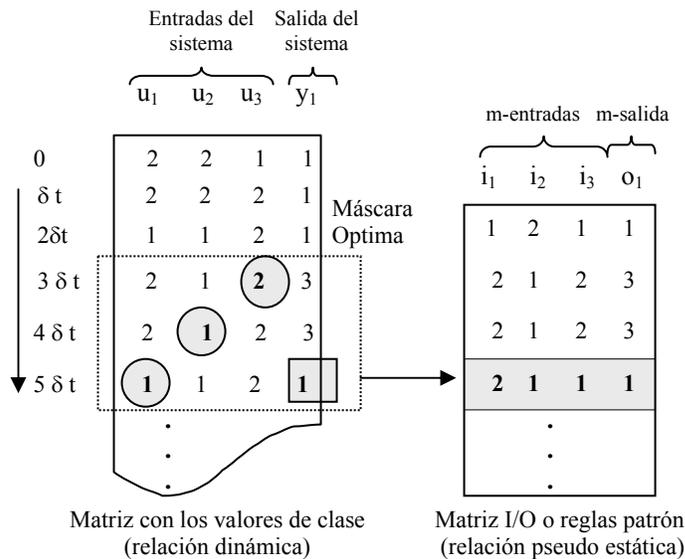
La calidad total de una máscara,  $Q$ , se define como el producto de su medida de reducción de entropía,  $H_r$ , y su ratio de observación,  $O_r$ :

$$Q_m = H_r \cdot O_r \quad (9)$$

Por lo tanto, la máscara óptima es la que tiene el valor más grande de  $Q$ . Una vez identificada la máscara óptima, ésta se utiliza para obtener la base de reglas patrón por medio del proceso de conversión de patrones dinámicos a patrones pseudo-estáticos (Véase *figura 4*).

La *figura 4* muestra el proceso de conversión de relaciones dinámicas en relaciones pseudo-estáticas utilizando la máscara de la *ecuación 3*. En la parte izquierda se muestra un fragmento de la matriz con los valores de clase, la primera de las tres matrices que forman el conjunto de datos cualitativos. En el ejemplo, se muestran las variables de entrada del sistema:  $u_1$ ,  $u_2$ , y  $u_3$ , las cuales se discretizaron en dos clases, mientras que la variable de salida  $y_1$ , se discretizó en tres clases. El rectángulo punteado simboliza la máscara que va desplazándose hacia abajo, a través de la matriz de valores de clase. Los círculos de la máscara denotan las posiciones de las m-entradas, mientras que el cuadrado indica la posición de la m-salida. Los valores de clase se leen a través de los ‘hoyos’ de la máscara y se colocan en forma consecutiva en la matriz de entrada/salida (I/O) o de reglas patrón. Como se puede apreciar en el lado derecho de la *figura 4*, cada renglón representa una posición de la máscara a través del recorrido que hace por la matriz de valores de clase. Estos

valores se almacenan en la matriz I/O y se alinean con el último renglón de la máscara. Cada renglón de esta matriz representa un estado cualitativo pseudo-estático o regla cualitativa basada en patrones. Por ejemplo, la regla patrón sombreada de la *figura 4* se lee de la siguiente forma: “si la primera m-entrada,  $i_1$ , tiene un valor de 2 (correspondiendo a *alto*); la segunda m-entrada,  $i_2$ , tiene un valor de 1 (correspondiendo a *bajo*); y la tercera m-entrada,  $i_3$ , tiene un valor de 1 (correspondiendo a *bajo*); entonces la m-salida,  $o_1$ , asume el valor de 1 (correspondiendo a *bajo*)”.



**Figura 4.** Proceso de obtención de la matriz I/O (base de reglas patrón).

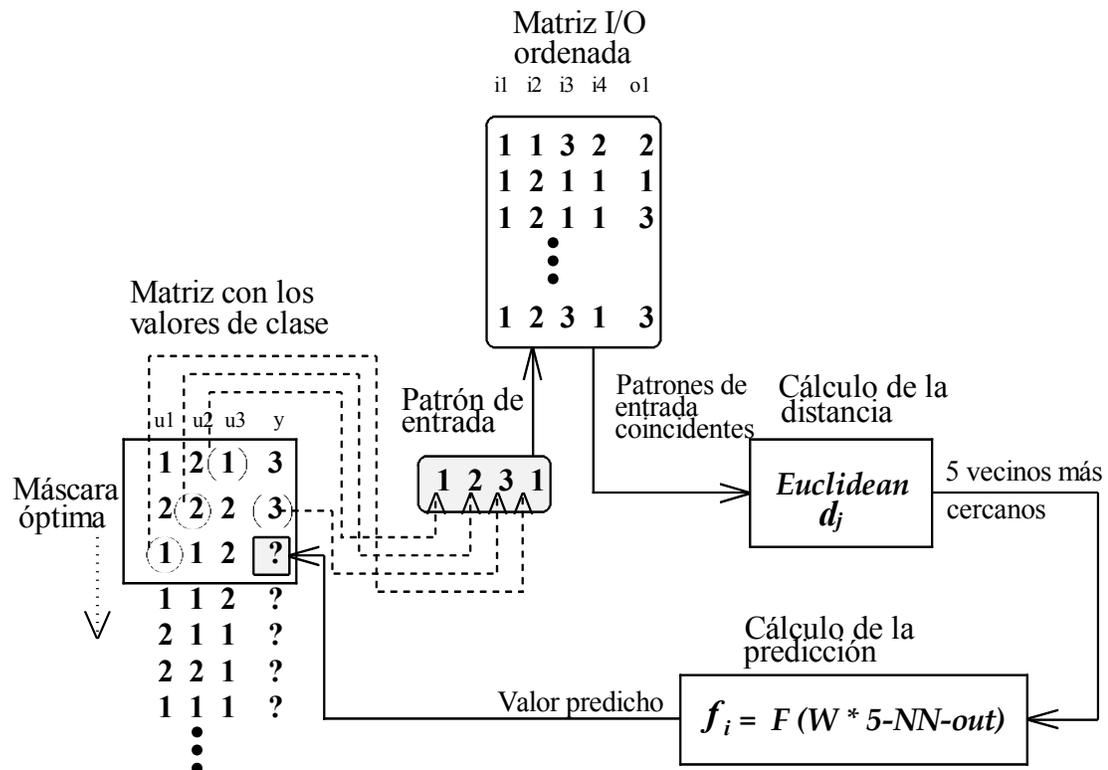
Como ya se mencionó anteriormente, la máscara junto con la base de reglas patrón configuran el modelo FIR del sistema en estudio.

### 1.3 Simulación cualitativa

A partir de la máscara óptima y de la base de reglas basada en patrones generada, se realiza la predicción de los valores futuros de *clase*, *pertenencia* y *lado*. La simulación cualitativa tiene lugar utilizando el motor de inferencia que está en el módulo de simulación cualitativa de FIR. El motor de inferencia utiliza el algoritmo de los 5 vecinos más cercanos (5NN) para producir el valor de predicción más consistente posible con el estado actual del sistema. La *figura 5* muestra de manera

esquemática este proceso.

La máscara es desplazada por encima de la matriz cualitativa de clases. Se extraen los valores correspondientes a las  $m$ -entradas (agujeros de la máscara) formando un patrón de entrada estático. Se buscan en la base de reglas patrón (matriz I/O en la *figura 5*) todos aquellos patrones que coincidan con el patrón de entrada actual respecto a los valores de clase. Para todos ellos se calcula su distancia respecto al mismo y se conservan aquellos 5 que estén más cercanos. Los valores de clase, pertenencia y lado de la predicción se obtendrán como una combinación ponderada de las salidas de los 5 vecinos más cercanos seleccionados. La tripleta cualitativa predicha se copia en las matrices cualitativas de clase, pertenencia y lado, y luego se procede a desplazar la máscara una posición hacia abajo (que corresponde a un intervalo de muestreo) para predecir el siguiente valor.



**Figura 5.** Esquema del proceso de simulación cualitativa de FIR.

Para realizar la comparación entre el patrón de entrada actual y todos aquellos patrones de la matriz I/O (Véase *figura 5*) que conciden con el patrón actual en los

valores de clase, se utiliza una función de normalización sobre cada uno de los elementos,  $p_i$ , del nuevo patrón de entrada. Existen diferentes funciones de normalización implementadas en FIR [LC90], pero la más utilizada es la siguiente:

$$p_i = Side_i \cdot B \cdot \sqrt{\ln Memb_i} + 0.5 \quad (10)$$

donde  $B = (4 \cdot \ln 0.5)^{-1/2}$ . Para las clases del extremo izquierdo la función se calcula mediante la ecuación:

$$p_i = C \cdot \sqrt{\ln Memb_i} \quad (11)$$

y para la clase del extremo derecho:

$$p_i = 1 - C \cdot \sqrt{\ln Memb_i} \quad (12)$$

donde  $C = (\ln 0.5)^{-1/2}$ . Independientemente del rango de las señales originales, las señales  $p_i$  están dentro del rango 0.0 y 1.0.

Los valores  $p_i$  correspondientes a las diferentes variables de un estado de entrada se concatenan para formar un vector:

$$\mathbf{p} = [p_1, p_2, \dots, p_N] \quad (13)$$

Asumiendo que el estado tiene  $N$  m-entradas. Llamamos al vector  $\mathbf{p}$  la imagen normalizada del estado de entrada original. El análisis prosigue calculando imágenes normalizadas para cada estado de entrada encontrado en la base de reglas patrón que contiene las mismas clases. A estos vectores los llamaremos  $\mathbf{p}_k$ . Cada vector  $\mathbf{p}_k$  es diferente puesto que solo coinciden los valores de clase del estado de entrada, pero no los valores de pertenencia ni los de lado. Finalmente, se calcula la norma 2 ( $\mathcal{L}_2$ ) de las diferencias entre el vector  $\mathbf{p}$  que representa la imagen normalizada del nuevo patrón de entrada y los vectores  $\mathbf{p}_k$  que corresponden a las imágenes normalizadas de las reglas patrón con el mismo estado de entrada:

$$d_k = \|\mathbf{p} - \mathbf{p}_k\|_2 = \sqrt{\sum_{i=1}^N (p_i - p_{ki})^2} \quad (14)$$

Se conservan las 5 reglas patrón con menor  $d_k$ , que se llamarán los cinco vecinos más cercanos al nuevo patrón de entrada del cual se requiere predecir su salida (o consecuente). La contribución de cada vecino a la estimación de la predicción del nuevo estado de salida se realiza en función a su proximidad. Esto se expresa dando un peso a cada vecino en función de la distancia. Si ninguna de las 5 distancias es igual a cero, los pesos absolutos se calculan utilizando la ecuación:

$$w_{abs_k} = \frac{(d_{\max}^2 - d_k^2)}{d_{\max} \cdot d_k} \quad (15)$$

donde  $k$  representa el vecino  $k$ -ésimo y  $d_i \leq d_j, i < j$ ;  $d_{\max} = d_5$ . En el caso que uno de los vecinos tenga una distancia igual a cero, el peso de ese vecino se hace 1.0 y el peso de los demás vecinos será 0.0. A partir de los valores absolutos de los pesos, se procede a calcular los pesos relativos que serán números entre el 0.0 y el 1.0, y su suma será siempre 1.0.

$$w_{rel_k} = \frac{w_{abs_k}}{\sum_{\forall k} w_{abs_k}} \quad (16)$$

La predicción del estado de salida en el espacio normalizado se calcula como una suma ponderada de los estados de salida de los cinco vecinos más cercanos. De esta manera, el estado cualitativo de salida puede obtenerse usando la siguiente fórmula:

$$State_{out_{new}} = \sum_{\forall k} w_{rel_k} \cdot (Class_{out_k} + p_{out_k}) \quad (17)$$

donde  $Class_{out_k}$  y  $p_{out_k}$  son la clase y la imagen normalizada de los valores de salida de los 5 vecinos más cercanos, respectivamente. El valor de predicción de clase se obtendrá mediante la ecuación siguiente:

$$Class_{out_{new}} = IFIX(State_{out_{new}}) \quad (18)$$

donde  $IFIX$  es una función que devuelve la parte entera del parámetro. El valor de predicción de pertenencia y de lado se obtendrán utilizando el inverso de la función de normalización anterior a partir de:

$$p_{out_{new}} = State_{out_{new}} - Class_{out_{new}} \quad (19)$$

## 1.4 Defusificación

Finalmente, la predicción cualitativa puede ser convertida a datos cuantitativos mediante el proceso de defusificación. Este es el proceso inverso al de fusificación, ya explicado anteriormente.

Después de obtener el valor cuantitativo de la predicción sólo hace falta compararlo con los datos originales, lo que nos permite validar el modelo cualitativo.

Para calcular el error de las predicciones generadas utilizamos normalmente el error cuadrático medio normalizado por la varianza, en porcentaje:

$$MSE = \frac{E[(y(t) - \hat{y}(t))^2]}{y_{var}} \cdot 100\% \quad (20)$$

donde  $y_{var}$  es la varianza definida como:

$$y_{var} = E[y^2(t)] - \{E[y(t)]\}^2 \quad (21)$$

donde  $y$  es el valor real y  $\hat{y}$  es el valor de predicción.

Sin embargo, existen diferentes fórmulas de error ampliamente usadas en la literatura con el mismo fin y que también aparecerán en las aplicaciones que incluimos en esta tesis. Para mayores detalles de la metodología FIR, referirse a [CNMdA96, NCV98, Neb94].

## **CAPITULO II:**

---

### **Antecedentes y Estado de Arte**



## Capítulo II: Antecedentes y Estado del Arte

Este capítulo describe los marcos de trabajo y teorías relacionadas con los *Sistemas Difusos Evolutivos* (SDEs). Se comienza con un estudio acerca de los antecedentes y estado actual del tema. Seguidamente, se aborda el modelado e identificación de sistemas. Se señalan algunas técnicas de *Soft Computing* que son usadas para llevar a cabo esa tarea. A continuación se establecen las nociones básicas sobre composición y funcionamiento de los *Sistemas Basados en Reglas Difusas Descriptivos* (SBRDDs) y los *Sistemas Neurodifusos* (SNDs). Los principios básicos de la Lógica Difusa no son incluidos, pero pueden ser consultados en [KY95, Zim96, DP97]. Para finalizar, se describe el tipo de *Algoritmos Evolutivos* (AEs) implementados en esta memoria, los *Algoritmos Genéticos* (AGs), considerados en este momento la técnica de búsqueda global más conocida y empleada.

### 2.1 Trabajos Previos

Se ha señalado que los *Sistemas Neurodifusos* (SNDs) y los *Sistemas Difusos Evolutivos* (SDEs) representan los productos más evidentes de la hibridación de sistemas dentro de *Soft Computing*. Existen numerosas investigaciones en ambos tópicos pero los SNDs han sido los más empleados en un gran número de aplicaciones, en particular en el área industrial [Jin03, Leo99, DO01]. Sin embargo, se han obtenido resultados muy prometedores en los últimos años para los SDEs [Jin04, CHHM01, Ped97].

Las primeras publicaciones sobre SDEs se realizaron en 1991 y entre ellas están [Kar91a, PK91, Thr91, VR91]. En [CGH+04] se presenta un excelente resumen de las investigaciones llevadas a cabo en los últimos 10 años en este campo. Como describen los autores en ese estudio, los SDEs más conocidos y estudiados son los *Sistemas Basados en Reglas Difusas Evolutivos* (SBRDEs) [CHHM01, CH95, HM97, GH97], quienes emplean AEs para aprender o ajustar diferentes componentes de un *Sistema Basado en Reglas Difusas* (SBRD). Sin embargo, también se han desarrollado con exitosos resultados otros tipos de SDEs tales como *Sistemas Neurodifusos Evolutivos* (SNDEs) [ADB02, RFK02, CLL00, Rus98, LX06] y *Algoritmos Clustering Difusos Evolutivos* (ACDEs) [NMP97, HBBB94, Van95, YKSS95, KA05].

Esta memoria está focalizada en los SDEs y se profundizará en los SBRDEs por ser los sistemas más cercanos a la investigación realizada en esta tesis doctoral.

Aunque los AEs no son algoritmos de aprendizaje, ofrecen una metodología de búsqueda potente e independiente del dominio que puede ser aplicada a gran cantidad de tareas de aprendizaje. De hecho, los AEs han sido muy usados en problemas de aprendizaje automático de reglas de producción a partir de conjuntos de ejemplos [Gre94]. Esto es debido, a que los individuos se pueden tratar al mismo tiempo como datos que son manipulados por el AE y como código ejecutable que lleva a cabo cierta tarea [Mic96].

Existen distintos tipos de SBRDEs según la tarea de diseño que lleven a cabo sea adaptar las definiciones de los factores o funciones de escala y/o de las funciones de pertenencia de los conjuntos difusos de las reglas (la Base de Datos, en el caso de los sistemas descriptivos), la obtención de las propias reglas (la Base de Reglas Lingüísticas, en el caso de los sistemas descriptivos) o ambas a la vez (la Base de Conocimiento completa en el caso de los sistemas descriptivos o la Base de Reglas Difusas en el caso de los aproximativos). Ejemplos de SBRDEs para factores o funciones de escala son [GGP98, Mag97, MV97, Hof01], funciones de pertenencia [Gür99, HLV95, CCdJH05, Hof01, ACP06], bases de reglas [INM99, GP01, Hof01, CPC04, dJHNS04] y bases de conocimiento [Mag97, MV97, HP97, HD97b, CPC04, CL05, ACP06].

Para el caso de la *Base de Datos* (BD) de un *Sistema Basado en Reglas Difusas Descriptivo* (SBRDD), los AEs se han empleado para adaptar las definiciones de los factores de escala y de las funciones de pertenencia, considerando definiciones parametrizadas de los mismos, además de para adaptar los contextos. Estos métodos parten de una *Base de Reglas Lingüísticas* (BRL) ya derivada que no van a modificar, además lo que se codifica en cada individuo del AE es la definición completa de la BD. Las funciones de escala transforman los dominios de las variables lingüísticas en los dominios de definición de las funciones de pertenencia difusas. Desde el punto de vista lingüístico, las funciones de escala pueden interpretarse como una especie de información de contexto, mientras que las funciones de pertenencia describen la semántica relativa de las particiones difusas (independiente del contexto). La unión de los factores de escala y las funciones de pertenencia generan la semántica completa de las variables lingüísticas [Vil00].

La forma más común de las funciones de pertenencia en los SBRDEs es tipo funciones triangulares (isósceles o asimétricas), trapezoidales o gaussianas. En [GGP98] se introduce el uso de funciones de transformación contextuales para ajustar funciones de pertenencia. Un buen ajuste de las funciones de pertenencia es crítico cuando se evalúa la efectividad de los sistemas difusos en problemas de control, modelado o clasificación. La adaptación lineal de contextos es simple y rápida, pero las funciones de pertenencia obtenidas se distribuyen uniformemente. En cambio la adaptación no lineal de contextos es costosa computacionalmente, pero las funciones

de pertenencia pueden estrecharse o ensancharse para representar mejor conceptos en entornos reales con alta sensibilidad en las clases extremas o en las clases intermedias. En [GGP98] se emplea un AG para encontrar una transformación no lineal dando la base de las funciones de pertenencia y un conjunto de datos disponibles de la aplicación estudiada. Por otro lado, el método propuesto por Magdalena en [Mag97] utiliza un AG para adaptar la definición de la totalidad de la *Base de Conocimiento* (BC) de un controlador difuso, es decir, se usa para aprender tanto la base de reglas como la ganancia y la sensibilidad del sistema estudiado. En lo que respecta al aprendizaje de la BD, el AG adapta el universo de discurso de las variables y la definición de las funciones de pertenencia de las etiquetas, esto último por medio de una función de escala no lineal que produce áreas de mayor sensibilidad de las etiquetas en el centro o en los extremos de la partición difusa. En esta investigación se utilizó un problema clásico de control: *El péndulo invertido sobre un carro* (PIC)<sup>5</sup>, que en la literatura se ha establecido como un *benchmark* para la evaluación de aprendizajes en sistemas de control. El método propuesto por este autor se amplía en [MV97] añadiendo un nuevo parámetro a la función de escala no lineal, que permite que el área de mayor sensibilidad de la partición difusa esté solamente en uno de los extremos.

En [Hof01] se describen dos aplicaciones de SDEs: una estrategia evolutiva (EEs) que ajusta las funciones de escala y de pertenencia en un PIC y un AG que aprende las reglas difusas de control para un comportamiento obstáculo-evasión de un robot móvil. En la primera aplicación, el proceso evolutivo empleado permitió reducir el costo funcional que mide lo rápido que el controlador estabiliza el sistema de péndulo invertido, y en la segunda, la herramienta evolutiva fue capaz de transferir al robot el comportamiento real adaptado en la simulación, demostrando la misma eficiencia en ambientes desconocidos previamente.

Por otro lado, en [CCdJH05] se presenta un proceso de ajuste genético para encajar conjuntamente las representaciones simbólicas de las reglas difusas y el significado de las funciones de pertenencia involucradas. La eficiencia de esta propuesta radica, principalmente, en el enfoque de ajuste realizado a dos niveles diferentes de descripción de cada una de las reglas lingüísticas difusas, lográndose obtener un buen equilibrio entre interpretabilidad y precisión. Por su parte, en [Gür99, HLV95] se describen métodos basados en AGs para alterar las formas de los conjuntos difusos cambiando sus localizaciones máximas. En estas investigaciones se utiliza el sistema de péndulo invertido para realizar los experimentos y se asume que la base de reglas y los conjuntos difusos están ya definidos. Otros ejemplos de SBRDEs que ajustan la BD son [BKC96, Kar91b, DD01].

---

<sup>5</sup> El objetivo de tal sistema es el control de las fuerzas de traslación aplicadas a un carro, colocado en el centro de una pista finita, mientras se equilibra una varilla que está montada o descansa sobre el carro y que gira libremente alrededor de un pivote fijo.

Hay que señalar que los AEs se han utilizado fundamentalmente para el aprendizaje de la BRL y existen tres enfoques de aplicación. De ellos, dos son clásicos y han sido muy usados a lo largo de los últimos años, mientras que el tercero ha aparecido recientemente tratando de combinar las mejores características y de resolver los distintos problemas de los dos anteriores. Los dos primeros enfoques, Michigan y Pittsburgh, propuestos respectivamente en [Boo82, HR78] y [Smi80], se diferencian principalmente en el esquema de representación. Mientras que en el enfoque Michigan cada cromosoma codifica una única regla y la solución final al problema está formada por la última población obtenida por el AE al completo, en el enfoque Pittsburgh, los cromosomas corresponden a bases de reglas completas, por lo que la solución al problema de aprendizaje está compuesta finalmente por el mejor cromosoma de la última población. En cambio, en el tercero de los enfoques, conocido como *Aprendizaje Iterativo de Reglas* (AIR) [CH97, GH97, Ven92], cambia un poco la filosofía de trabajo. Al igual que en el enfoque Michigan, cada cromosoma codifica una regla individual, con lo que el proceso evolutivo se hace a nivel de regla y no de base de reglas, pero el único individuo considerado para formar parte de la solución final es el mejor adaptado de la última población.

Investigaciones tales como [INM99, GP01] permiten la generación automática de reglas difusas a través de métodos genéticos. En [INM99] las reglas difusas se obtienen usando AGs para problemas multidimensionales de clasificación de patrones y se usan funciones de pertenencia fijas, por consiguiente, no se aplica ningún mecanismo de ajuste a las mismas. Este enfoque obtuvo mejor desempeño con respecto a otros métodos de clasificación tales como técnicas no difusas de *machine learning* y redes neuronales. Por otro lado, en [GP01] se analizan diferentes estrategias de búsqueda (AGs, Recocido Simulado<sup>6</sup> y Ascensión de Colinas<sup>7</sup>) para encontrar las mejores reglas difusas que describen los sistemas bajo estudio<sup>8</sup> [MA98]. Los resultados experimentales permitieron concluir que el AG propuesto presentó buen desempeño en comparación con otros mecanismos de búsqueda estudiados en este trabajo, ya que tuvo una alta capacidad para detectar las variables irrelevantes y obtener reglas simples. Sin embargo, otros algoritmos de búsqueda como el Recocido Simulado obtuvieron una capacidad predictiva tan buena como el AG.

---

<sup>6</sup> El término castellano Recocido Simulado (RS) es una traducción del término inglés *Simulated Annealing*. El RS deriva de la analogía entre la mecánica estadística de partículas de un compuesto (líquido o sólido) y la búsqueda de soluciones en problemas complejos de optimización combinatoria [VLA88, Aar89].

<sup>7</sup> En inglés *Hill-climbing*. Estos métodos permiten obtener valores óptimos locales y dependen fuertemente del punto inicial. Para ello, utilizan una metodología de mejora en forma iterativa. Durante cada iteración, se aplica una función de evaluación sobre un punto y alguno de su vecindad. El mejor posicionado es el elegido para continuar [Pap00; RN03, ch. 4].

<sup>8</sup> *UCI Repository of Machine Learning Databases and Domain Theories*. Disponible vía *ftp* anónimo desde <ftp.ics.uci.edu/pub/machine-learning-databases>.

Por su parte, en [dJHNS04] se propone el uso de un algoritmo Adaboost optimizado por un AG que también se usa para el aprendizaje de reglas difusas. Adaboost [FS96] es comúnmente utilizado para combinar clasificadores lineales y consiste en un algoritmo incremental (*boosting algorithm*<sup>9</sup>), el cual repetidamente invoca a un algoritmo de aprendizaje que identifica la regla difusa que mejor coincide con el conjunto de datos de entrenamiento actual. En este estudio, se plantea la aplicación del nuevo algoritmo Adaboost para aprender un clasificador basado en reglas difusas, y experimentalmente se muestran las ventajas de este enfoque sobre otros métodos de clasificación difusos y no difusos.

Por último, se pueden mencionar varias investigaciones en relación al aprendizaje de la *Base de Conocimiento* (BC), por ejemplo en [HP97, HD97b, CL05] se presentan tres perspectivas genéticas. En [HP97] se describe un método de aprendizaje el cual diseña un controlador difuso por medio de un AG de manera automática. Este trabajo propone el uso de un nuevo esquema de codificación compacto y flexible para la base de reglas difusa, en imitación a los AGs desordenados (*messy GAs* [GKD89]). Esta codificación no depende de la posición de los genes dentro del cromosoma y permite reducir la complejidad y el tamaño de la base de reglas, a través del cual el AG es capaz de resolver la tarea de diseño incluso para controladores difusos con muchas variables de entrada. El comportamiento del control adaptado en este estudio se prueba en experimentos sobre un robot móvil. Por otro lado en [HD97b] se introduce un AG en cascada con la idea de dividir el proceso de diseño de un sistema difuso en la optimización de la estructura y de los parámetros, respectivamente. Este algoritmo se prueba en el diseño de un controlador difuso, en este caso un PIC. Los experimentos muestran que el algoritmo propuesto obtiene considerablemente mejores resultados que un AG convencional. En [CL05] se emplea un algoritmo evolutivo iterativo de dos niveles para construir un controlador difuso por medio de AGs. El nivel superior se encarga de la composición de las reglas difusas utilizando las funciones de pertenencia ajustadas en el nivel inferior, mientras este nivel se encarga de determinar la forma de las funciones de pertenencia usando las reglas difusas aprendidas en el nivel superior. Asimismo, los autores proponen un esquema nuevo de codificación para vencer el problema de las limitaciones que se presentan durante el ajuste de las funciones de pertenencia. El SDE resultante se comparó con otros SDEs, redes neuronales y SNDs por medio de ejemplos de comportamientos tanto teóricos como de campo observados en vehículos. Se concluyó que el SDE propuesto daba mejores resultados que las demás aproximaciones empleadas.

Del mismo modo, en [ACP06] se expone un esquema novedoso de

---

<sup>9</sup> Los algoritmos incrementales (*boosting algorithms*) son técnicas de modelado aditivo estadístico que combinan diferentes clasificadores de baja calidad con el fin de obtener un clasificador compuesto que funcione mejor que cualquiera de sus componentes.

administración de búfer para memoria compartida en *switches ATM*<sup>10</sup> basado en computación evolutiva. El esquema de control presentado se basa en un sistema difuso y AGs. El SD se usa para implementar un esquema de control adaptativo para la memoria compartida vía un conjunto de reglas difusas sencillas, las cuales representan las relaciones entre las variables del sistema. El SD óptimo se obtiene usando una metodología sistemática basada en AGs para seleccionar y ajustar las funciones de pertenencia. Los AGs se utilizan para derivar la mejor BC para un *switch* dado. La novedad de este enfoque estriba en el uso de AGs como un esfuerzo aislado para ajustar el SD *offline* para un *switch* determinado, mientras que el control en tiempo real de la memoria compartida se realiza por el SD derivado. Este enfoque metodológico permite obtener los parámetros del SD para cada tamaño de *switch*, ofreciendo un alto grado de escalabilidad en el sistema de control difuso resultante. El SDE propuesto obtuvo una eficiencia comparable al mecanismo ideal que usa políticas tipo *push-out*.

Por otro lado, en [CPC04] se presentan dos enfoques evolutivos diferentes con el propósito de analizar la eficiencia de ambos cuando se aplican a problemas de clasificación. En el primer enfoque, los conjuntos difusos son definidos previamente por un algoritmo de *clustering* y la base de reglas se crea y optimiza automáticamente usando AGs. En el segundo enfoque, la BD es el objeto de aprendizaje del AG en vez de la base de reglas. En este caso la base de reglas se genera por medio de un método auxiliar denominado Wang y Mendel [WM92] que considera la BD previamente derivada, obteniendo la definición completa de la BC por medio de la acción cooperativa de ambos métodos. Los resultados obtenidos por los enfoques propuestos son promisorios, ya que se obtiene un alto número de clasificaciones correctas para todos los conjuntos de prueba usados. También existen otras propuestas de SBRDEs que aprenden la BC en su totalidad [CH97] o mediante el aprendizaje evolutivo de la BD [Glo96, FJ96, IM96].

En la *Tabla 1* se presenta un resumen de las principales investigaciones consideradas en ésta sección sobre SBRDEs. En la primera columna se indica la referencia de la investigación. La segunda y tercera columnas indican las metodologías empleadas y la tarea de diseño que persiguen, respectivamente. El ítem “tarea de diseño” se refiere al objetivo que se busca: a) la definición de los factores o funciones de escala (FE) y de las funciones de pertenencia (FP) de los conjuntos difusos (BD en el caso de los sistemas descriptivos), b) la obtención de las propias reglas (BRL en el caso de los sistemas descriptivos), ó c) ambas a la vez, es decir, la Base de Conocimientos (BC) en SBRDDs o la Base de Reglas Difusas (BRD) en el caso de SBRDs aproximativos. La siguiente columna indica el tipo de problema tratado (control, predicción o clasificación), y la última columna, indica el campo de aplicación usado para probar esos métodos.

---

<sup>10</sup> En inglés *Asynchronous Transfer Mode*.

**Tabla 1.** Resumen de las principales referencias consideradas sobre SBRDEs.

REFERENCIA	METODOLOGIA EMPLEADA	TAREA DE DISEÑO	TIPO DE PROBLEMA	APLICACION
[GGP98]	AGs	FE	Clasificación	Funciones de prueba
[Mag97]	AGs	FE / BC	Control	PIC
[MV97]	AGs	FE / BC	Control	PIC / Bípodo antropomórfico
[HLV95]	AGs	FP	Control	Péndulo invertido
[ACP06]	AGs	FP	Control	Switches ATM
[Gür99]	AGs	FP	Control	PIC
[CCdJH05]	AGs	FP	Predicción / Clasificación	Estimación de costos de mantenimiento en redes de electricidad / Valoración de la calidad de arroz
[FJ96]	AGs	FP / BC	Control	Sistema de tratamiento de aguas residuales
[Glo96]	AGs	FP / BC	Predicción	Series temporales de prueba
[IM96]	AGs	FP / BC	Clasificación	Bases de datos del Repositorio UCI
[DD01]	AGs	BD=FE+FP	Predicción	Estimación de la velocidad del viento para generación de energía en aerogeneradores.
[BKC96]	AGs	BD=FE+FP	Control	Sistema ferroviario de carga.
[Kar91b]	AGs	BD=FE+FP	Control	Sistemas: a) Nivel de líquidos, b) Controlador de pH y c) Acoplamiento de satélites.
[Hof01]	EEs / AGs	FE + FP / BRL	Control	PIC / Comportamiento robot móvil.

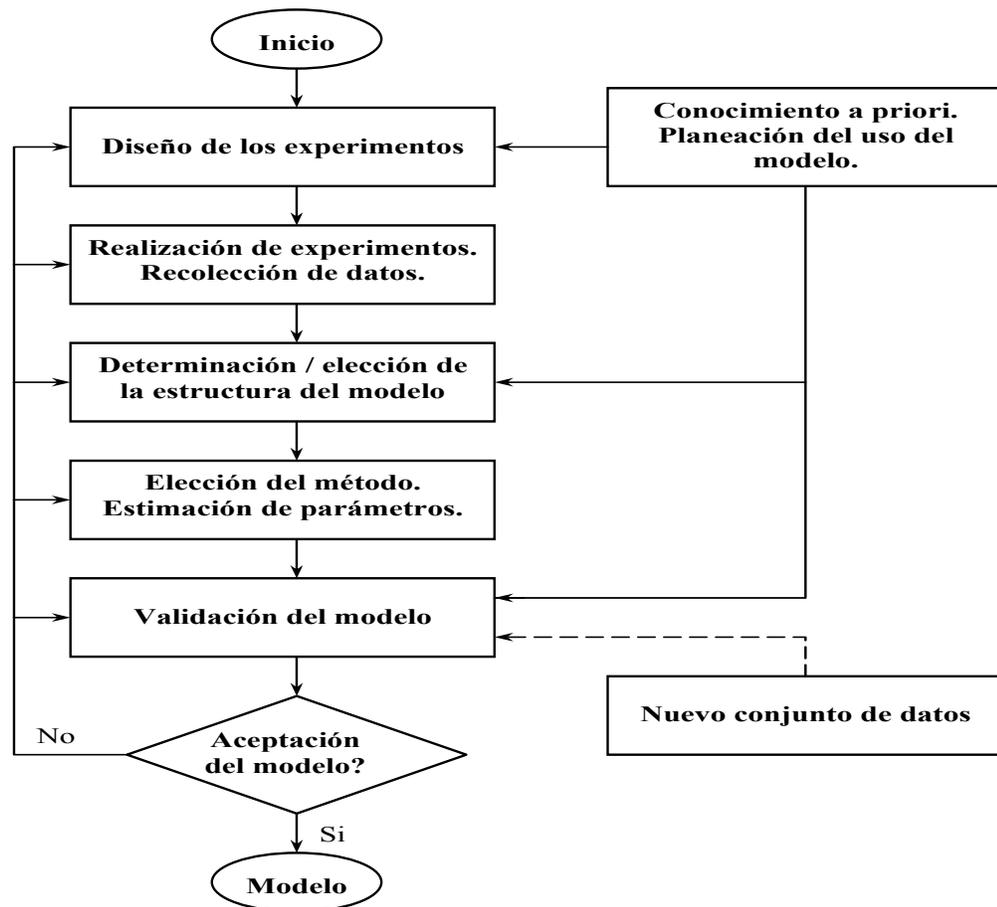
REFERENCIA	METODOLOGIA EMPLEADA	TAREA DE DISEÑO	TIPO DE PROBLEMA	APLICACION
[INM99]	AGs	BRL	Clasificación	Bases de datos del Repositorio UCI
[GP01]	AGs / RS / Ascensión de Colinas	BRL	Clasificación	Bases de datos del Repositorio UCI
[dJHNS04]	Algoritmo Adaboost + AGs	BRL / BRD	Clasificación	Bases de datos del Repositorio UCI
[CPC04]	AGs	BRL / BC	Clasificación	Bases de datos del Repositorio UCI y otras
[HP97]	AGs	BC	Control	Comportamiento robot móvil
[HD97b]	AGs	BC	Control	PIC
[CL05]	AGs	BC	Control	Comportamiento de vehículos
[CH97]	EEs / AGs	BC	Control	

Como se puede apreciar de la *Tabla 1*, los AEs mayoritariamente empleados son los AGs, lo cual demuestra que representan la técnica de búsqueda global más conocida y empleada. Asimismo, se concluye que existen pocos estudios dedicados a tareas de predicción, por lo cual nuestra investigación constituye una novedosa aportación en este sentido. El tipo de problema principalmente tratado es el control de sistemas de ahí que la mayor parte de las aplicaciones encontradas sean de ese ámbito.

## 2.2 Modelado e identificación de sistemas

El modelado e identificación de sistemas dinámicos tiene aplicación en diferentes áreas de la ingeniería y la ciencia. Es así como, para la aplicación de técnicas de control en la industria se requiere de una representación que caracterice al proceso físico, o sea, un modelo matemático del mismo [Joh93]. Esta representación se logra a través de distintas técnicas que pueden ser clasificadas de acuerdo a varios enfoques. Desde la perspectiva de las ciencias puras, la representación matemática se puede obtener de forma estrictamente teórica, basada en consideraciones físicas, leyes de Newton, balance de masa, balances de energía, etc, lo que denominan

comúnmente como *Modelado*; así como también en forma empírica, mediante experimentos sobre el sistema real, basándose en datos de entrada y salida del proceso, llamado *Identificación*. La *figura 6* muestra el procedimiento de identificación.



**Figura 6.** Procedimiento general para llevar a cabo una identificación de sistemas (Extraída de [Lju99]).

La principal ventaja del modelado e identificación de sistemas, es que permite predecir el comportamiento dinámico y estático de los mismos antes de construirlos y también se puede analizar el desempeño de uno ya existente, con la intención de estudiar su comportamiento a través del uso de modelos continuos, discretos o combinados. Además, sin exponer el proceso a daños, se puede determinar qué puede ocurrirle al sistema, simulando el modelo con una condición, entrada inusual o perturbación, cosa que facilita el diseño, la evaluación de alternativas, la predicción, la experimentación de estrategias, el contrastar hipótesis, el entrenamiento o la toma

de decisiones. Todo ello debería permitir obtener según los casos un mejor aprovechamiento de los recursos de los que se dispone, un incremento en los resultados esperados, una mejor organización, o una mayor formación y comprensión del sistema ó proceso bajo estudio, lo cual es esencial desde el punto de vista industrial u organizacional para lograr un adecuado nivel de competitividad en un entorno condicionado por la globalización de las actividades económicas.

### **2.2.1 Tipos de modelos matemáticos**

Según F. Cellier [Cel91] existen cuatro categorías de modelos matemáticos: a) modelos de tiempo continuo, b) modelos de tiempo discreto, c) modelos cualitativos y d) modelos de eventos discretos. En los primeros, las ecuaciones describen su comportamiento en cada instante de tiempo, mientras que en los segundos sólo en muestras discretas.

La tercera clase son los modelos cualitativos, en donde las variables dependientes son discretizadas. Los modelos cualitativos son usualmente codificados usando representación de estados finitos. En otras palabras estos modelos permiten describir situaciones en el mundo: objetos y procesos, en términos de relaciones temporales, espaciales y causales.

La última categoría se refiere al conjunto de modelos de eventos discretos. Paradójicamente, ambos ejes (tiempo y estado) de estos modelos son usualmente continuos, pero los modelos de eventos discretos difieren de los modelos de tiempo continuo en el hecho que, en un lapso finito de tiempo, sólo puede ocurrir un número finito de cambios de estado.

### **2.2.2 Técnicas de Soft Computing usadas para la identificación de sistemas**

Para obtener modelos existen diferentes técnicas de identificación de sistemas; algunas de ellas requieren medición completa del estado, pero su aplicabilidad se ve limitada ya que esto no siempre es posible, debido a la gran cantidad de no linealidades y perturbaciones presentes en los procesos o sistemas dinámicos. Una gran cantidad de los métodos clásicos no conducen a resultados aceptables, por lo tanto, puede ser algunas veces difícil encontrar modelos matemáticos representativos. Para obviar este problema es necesario hacer uso de observadores de estado [AW95, FPW94, IS96, Mos95], lo que conlleva a un incremento en la dinámica del sistema. Esto ha llevado a explorar otras técnicas que ayuden a eliminar estos inconvenientes. Las *Redes Neuronales* (RNs), los *Sistemas Basados en Reglas Difusas* (SBRDs), el *Razonamiento Inductivo Difuso* (FIR), los *Algoritmos Evolutivos* (AEs) o la

hibridación de los mismos, tales como los *Sistemas Neurodifusos* (SNDs) o los *Sistemas Basados en Reglas Difusas Evolutivos* (SBRDEs) son los métodos dentro del área de *Soft Computing* que han despertado mayor interés en los últimos años para tratar este problema. En particular, estos enfoques resultan muy ventajosos cuando se trata de la identificación de sistemas no lineales (como es el caso de todos los procesos industriales), donde la mayoría de los métodos clásicos no conducen a resultados aceptables.

Algunas de estas técnicas como los SBRDs, los SBRDEs, los SNDs y FIR hacen uso de la lógica difusa [KY95, Zim96, DP97] para llevar a cabo esta tarea. En las siguientes secciones se introducirán por un lado, los aspectos principales de la naturaleza y funcionamiento de los SBRDs, profundizando en los SBRD Descriptivos (SBRDDs) por su aplicabilidad práctica, y por el otro, nociones básicas de los SNDs. La metodología FIR fue ya presentada en el Capítulo I. Al final del presente capítulo se describen los aspectos fundamentales de los algoritmos genéticos que servirán de base para introducir los métodos propuestos en esta investigación.

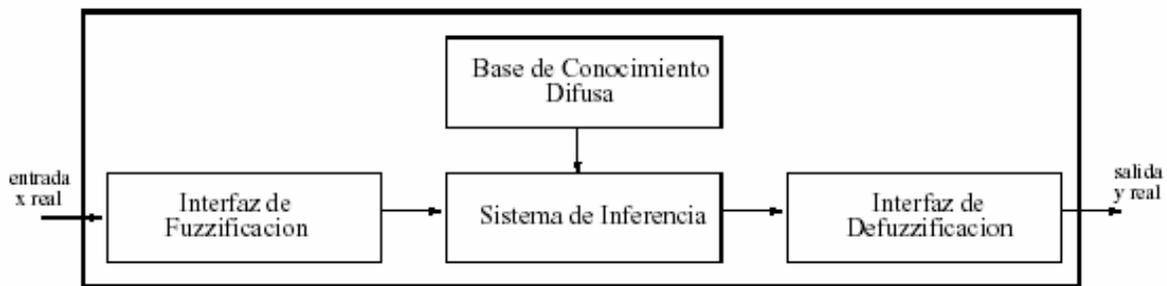
### **2.3 Sistemas Basados en Reglas Difusas Descriptivos**

En general, un *Sistema Basado en Reglas Difusas* (SBRD) es un sistema basado en reglas en el que la Lógica Difusa puede ser empleada tanto como una herramienta para representar distintas formas de conocimiento sobre el problema a resolver como para modelar las interacciones y relaciones existentes entre las variables del mismo. La principal aplicación de estos sistemas inteligentes ha sido el modelado difuso de sistemas [BD95, Ped96, SY93]. Esto se debe a que existe una amplia gama de sistemas reales para los cuales, debido a su nivel de complejidad o imprecisión, no resulta posible obtener buenos resultados con las herramientas clásicas de modelado. Los SBRDs se han aplicado con éxito a una serie de problemas reales a lo largo de los últimos años [BD95, GT95, Hir93, Ped96, Wan94]. Existen distintos tipos de SBRDs dependiendo de la naturaleza de las entradas y salidas que manejan, así como la estructura de reglas difusas con la que trabajan, pudiendo distinguirse entre SBRDs puros, descriptivos y aproximativos de tipo Mandani, y de tipo TSK. Para más detalles de los SBRDs, referirse a [Cor97].

En la actualidad los SBRDs descriptivos (SBRDDs) han sido los más usados porque presentan una serie de cualidades muy interesantes. Pueden ser empleados en aplicaciones reales de ingeniería, puesto que manejan datos de entrada-salida reales. Asimismo, proporcionan un marco natural para incluir conocimiento experto en forma de reglas lingüísticas y permiten combinarlo de forma muy sencilla con reglas obtenidas a partir de conjuntos de datos que reflejen el comportamiento del sistema.

Por último, presentan un alto grado de libertad a la hora de elegir las Interfaces de Fuzzificación y Defuzzificación, así como el Sistema de Inferencia, de modo que permite diseñar el sistema más adecuado para un problema concreto.

Los SBRDDs fueron propuestos por Mamdani [Mam74], quien fue capaz de plasmar las ideas preliminares de Zadeh [Zad73] en el primer SBRD concreto en una aplicación de control. Este tipo de Sistemas Difusos se conocen también con el nombre de SBRDs con Fuzzificador y Defuzzificador o más comúnmente por el de Controladores Difusos<sup>11</sup>.



**Figura 7.** Estructura básica de un Sistema Basado en Reglas Difusas de tipo Mamdani

La figura 7 muestra la estructura general de un SBRDD. La *Interfaz de Fuzzificación* establece una aplicación de valores precisos, correspondientes al dominio  $U$  de las variables de entrada al sistema, en conjuntos difusos definidos sobre el mismo universo de discurso. La *Interfaz de Defuzzificación*, en cambio, realiza la operación inversa estableciendo una aplicación de conjuntos difusos definidos en el dominio  $V$  de las variables de salida, en valores precisos definidos en el mismo universo. La *Base de Conocimiento* (BC) almacena el conocimiento disponible sobre el problema en forma de reglas. Por último, el *Sistema de Inferencia* se encarga de llevar a cabo el proceso de inferencia sobre las entradas haciendo uso de la información almacenada en la BC.

A continuación, se analizará en detalle cada uno de los componentes de un SBRDD.

### 2.3.1 Base de Conocimiento

La BC es la parte esencial de los SBRDDs debido a que los tres componentes restantes del sistema se ocupan de interpretar las reglas contenidas en ella y de

<sup>11</sup> Nombre acuñado por Mamdani en sus primeros trabajos [MA75], ya que su aplicación principal ha sido históricamente el control de sistemas.

posibilitar su utilización en problemas concretos. Está compuesta por la *Base de Reglas Lingüísticas* (BRL) y por la *Base de Datos* (BD):

- La BRL está formada por un conjunto de reglas lingüísticas de tipo “SI-ENTONCES” que en el caso de los SBRDDs con múltiples entradas y una única salida, presentan la siguiente estructura:

$$R1 : SI \ x_1 \text{ es } A_{11} \text{ y } \dots \text{ y } x_m \text{ es } A_{1m} \text{ ENTONCES } Y \text{ es } B_1 \\ ADEMÁS$$

$$R2 : SI \ x_1 \text{ es } A_{21} \text{ y } \dots \text{ y } x_m \text{ es } A_{2m} \text{ ENTONCES } Y \text{ es } B_2 \\ ADEMÁS$$

...

ADEMÁS

$$Rn : SI \ x_1 \text{ es } A_{n1} \text{ y } \dots \text{ y } x_m \text{ es } A_{nm} \text{ ENTONCES } Y \text{ es } B_n$$

donde  $x_i$  e  $Y$  son variables lingüísticas de entrada y salida respectivamente, y los  $A_{ij}$  y  $B_i$  son etiquetas lingüísticas asociadas con conjuntos difusos que determinan su semántica en cada una de las reglas. La BRL está compuesta por una serie de reglas de este tipo unidas por el operador ADEMÁS, lo que indica que todas las reglas pueden dispararse ante una entrada concreta.

- La BD contiene la definición de los conjuntos difusos asociados a los términos lingüísticos empleados en las reglas de la BRL. Además, almacena los valores de los factores de escala que efectúan las transformaciones necesarias para trasladar los universos de discurso en los que están definidos dichos conjuntos a aquellos en que se definen las variables de entrada y salida del sistema.

En los SBRD aproximativos, sin embargo, este último elemento no es necesario puesto que las variables que se emplean en la reglas no son lingüísticas sino difusas. La razón es obvia puesto que, al tomar cada variable como valor de un conjunto difuso distinto para cada regla y no etiqueta lingüística, no tiene sentido establecer una relación entre etiquetas y conjuntos difusos. En este tipo de sistemas, la BC queda reducida a una Base de Reglas Difusas (BRD) compuesta por un conjunto de reglas difusas aproximativas de tipo Mandami conectadas entre sí mediante el operador *ADEMÁS*.

### 2.3.2 Interfaz de Fuzzificación

La *Interfaz de Fuzzificación* permite al SBRDD trabajar con entradas y salidas reales. Tiene como tarea establecer una correspondencia entre cada valor preciso del espacio de entrada y un conjunto difuso definido en el universo de discurso de dicha entrada. Es por ello que trabaja del siguiente modo:

$$A' = F(x_0) \quad (22)$$

donde  $x_0$  es un valor preciso de entrada al SBRDD definido en el universo de discurso  $U$ ,  $A'$  es un conjunto difuso definido sobre el mismo dominio y  $F$  es un operador de fuzzificación.

Principalmente, existen dos posibilidades para la elección de  $F$ :

1. Fuzzificación puntual: Es la más empleada por su sencillez.  $A'$  se construye como un conjunto difuso puntual con soporte  $x_0$ , es decir, con la siguiente función de pertenencia:

$$A'(x) = \begin{cases} 1, & \text{si } x = x_0 \\ 0, & \text{en otro caso} \end{cases} \quad (23)$$

2. Fuzzificación no puntual o aproximada: En este caso,  $A'(x_0) = 1$  y el grado de pertenencia de los valores restantes de  $U$  va disminuyendo según se alejan éstos de  $x_0$ . Este segundo tipo de operador de fuzzificación permite el empleo de distintos tipos de funciones de pertenencia. Por ejemplo, en el caso de una función de pertenencia triangular, se puede emplear el siguiente:

$$A'(x) = \begin{cases} 1 - \frac{|x - x_0|}{\sigma}, & \text{si } |x - x_0| \leq \sigma \\ 0, & \text{en otro caso} \end{cases} \quad (24)$$

### 2.3.3 Sistema de Inferencia

El *Sistema de Inferencia* es el componente encargado de llevar a cabo el proceso de inferencia difuso. Es así como se hace uso de principios de la Lógica Difusa para establecer una aplicación entre conjuntos difusos definidos en  $U = U_1 \times U_2 \times \dots \times U_n$  y conjuntos difusos definidos en  $V$  (donde  $U_1, \dots, U_n$  y  $V$  son los dominios en los que están definidas las variables de entrada  $x_1, \dots, x_n$  y la de salida  $Y$ , respectivamente). Este proceso de inferencia está basado en la aplicación del Modus Ponens Generalizado, extensión del Modus Ponens de la Lógica Clásica, propuesto por Zadeh [Zad73]. Dicho mecanismo difuso de inferencia se pone en práctica mediante la Regla Composicional de Inferencia, que en su expresión más sencilla, toma la siguiente forma:

$$\mu_{B'}(y) = I(\mu_A(x), \mu_B(y)) \quad (25)$$

al ser aplicada sobre una regla de tipo *SI  $x_1$  es  $A_1$  y... y  $x_m$  es  $A_m$  ENTONCES  $Y$  es  $B$* , donde  $\mu_A(x) = T(\mu_{A_1}(x), \dots, \mu_{A_n}(x))$ ,  $T$  es un operador de conjunción difuso e  $I$  es un operador de implicación difuso.

De aquí puede deducirse claramente el hecho de que el proceso de inferencia difusa se aplica a nivel de reglas individuales. De esta manera, una vez aplicada la inferencia sobre las  $n$  reglas  $R_i$  que componen la BRL, se obtienen  $n$  conjuntos difusos  $B'_i$  que representan las acciones difusas que ha deducido el SBRDD a partir de las entradas recibidas.

### 2.3.4 Interfaz de Defuzzificación

Se ha comentado en la sección anterior que si la BRL se compone de  $n$  reglas, el proceso de inferencia devuelve  $n$  conjuntos difusos. Puesto que el sistema debe proporcionar una salida precisa, la *Interfaz de Defuzzificación* debe asumir la tarea de agregar la información aportada por cada uno de los conjuntos difusos individuales y transformarla en un valor preciso. Para efectuar esta agregación existen dos formas de trabajo diferentes [BD95, CHP97, Wan94]:

1. *Modo A-FATI: agregar primero, defuzzificar después.* En este caso, la interfaz de Defuzzificación lleva a cabo las siguientes tareas:

- a) Agrega los conjuntos difusos individuales inferidos ( $B'_i$ ) para obtener un conjunto difuso final ( $B'$ ). Para ello, emplea un *operador de agregación difuso*  $G$  que, modelando el *operador ADEMÁS*, relaciona las reglas de la base:

$$\mu_{B'}(y) = G\{\mu_{B'_1}(y), \dots, \mu_{B'_n}(y)\} \quad (26)$$

- b) Mediante un *método de defuzzificación*  $D$ , transforma el conjunto difuso  $B'$  obtenido en un valor preciso  $y_0$ , que será proporcionado como salida global del sistema:

$$y_0 = D(\mu_{B'}(y)) \quad (27)$$

2. *Modo B-FITA: defuzzificar primero, agregar después.* Este segundo modo de trabajo considera individualmente la contribución de cada conjunto difuso inferido y el valor preciso final se obtiene mediante una operación (una media, una suma ponderada o la selección de uno de ellos, entre otras) sobre un valor preciso característico de cada uno de los conjuntos difusos individuales. De este modo, se evita el cálculo del conjunto difuso final  $B'$ , hecho que ahorra una gran cantidad de tiempo computacional. Este modo de operación supone una aproximación distinta al concepto representado por el *operador ADEMÁS*.

El *Modo A-FATI* fue propuesto inicialmente y fue empleado por Mamdani en su primera aproximación al control difuso [Mam74]. En los últimos años, la modalidad

*B-FITA* ha sido muy utilizada [CHP97, DHR93, SY93], sobre todo en sistemas de tiempo real, donde se requieren tiempos de respuesta rápidos.

Cuando se trabaja en *Modo A-FATI* los métodos de defuzzificación más habituales son: el *Centro de Gravedad*, el *Centro de Sumas* (aproximación al Centro de Gravedad computacionalmente más rápida de obtener) y la *Media de los Máximos* [DHR93]. Los operadores de agregación considerados suelen ser el máximo y el mínimo.

En caso de emplear el *Modo B-FITA*, los operadores de agregación más usados son la *media*, la *media ponderada* o la selección de algún *valor característico* de los conjuntos difusos en función del grado de importancia de la regla que los ha generado en el proceso de inferencia [CHP97]. Como métodos para extraer valores representativos se suelen utilizar el *Centro de Gravedad* y el *Punto de Máximo Criterio*; y como grados de importancia de la regla, el *área* y la *altura del conjunto difuso* inferido o el *grado de emparejamiento de los antecedentes* de la misma entrada al sistema. El operador más empleado dentro de este grupo es la *media ponderada por el grado de emparejamiento*, que se suele combinar con el *Centro de Gravedad* como valor característico del conjunto difuso [CHP97, HT93, SY93].

## 2.4 Sistemas Neurodifusos

El avance de los sistemas de *Inteligencia Artificial* y la búsqueda de soluciones cada vez más completas para los problemas que pueden ser resueltos por estos métodos, han llevado a la unión de diferentes técnicas de *Soft Computing* [Bon97, Ova04], haciendo que las debilidades de unos sistemas se compensen con las bondades de otros. Esta última, resulta ser una fuerte motivación para realizar la unión de las técnicas de *Redes Neuronales Artificiales* (RNA) y *Lógica Difusa*, las cuales han propiciado el nacimiento en 1990 de los *Sistemas Híbridos Neurodifusos* o simplemente *Sistemas Neurodifusos* (SNDs). En la actualidad se habla de SNDs donde una red neuronal captura el conocimiento y el sistema difuso lo expresa en reglas que pueden ser entendidas por los seres humanos [NKK97]. En la *Tabla 2* se pueden observar las similitudes y diferencias entre estas dos técnicas.

**Tabla 2.** Comparación entre las técnicas de RNA y Lógica Difusa.

ATRIBUTO	RNA	Lógica Difusa
Paralelismo	Bueno	Bueno
Dominio	Continuo	Continuo

<b>ATRIBUTO</b>	<b>RNA</b>	<b>Lógica Difusa</b>
Linealidad	No lineal	No lineal
Base de Conocimiento	Distribuida en pesos	Distribuida en Base de Reglas
Método de Solución	Matemática mediante propagación de la información de entrada	Razonamiento sobre la Base de Reglas
Representación del conocimiento	Numérica	Lingüística
Conocimiento explotado	Datos empíricos	Conocimiento humano previo
Capacidad de aprendizaje	Alta	No tiene - Es una forma lógica de tratamiento matemático.

Redefiniendo las motivaciones para crear los SNDs, se pueden esbozar las siguientes:

- Los problemas a resolver con estas técnicas inteligentes tienen generalmente un componente de datos empíricos y otro de conocimiento previo, que con los sistemas neurodifusos pueden ser explotados íntegramente mezclando los métodos cuantitativos y cualitativos de las RNA y la lógica difusa respectivamente [BBH95].
- Habiendo encontrado que la representación del conocimiento como un sistema difuso se acerca a la forma de razonamiento humana, es natural intentar adicionar a estos sistemas capacidades de aprendizaje, de forma que se emule mejor tal comportamiento o que se mejoren los resultados obtenidos con el sistema difuso sólo.

Definidas las motivaciones para la realización de los SNDs, a continuación se revisará la taxonomía de estos sistemas, clasificándolos según la interacción de sus dos componentes (lógica difusa y RNA), posteriormente se estudiarán las ventajas y limitaciones de los mismos.

#### **2.4.1 Taxonomía de los SNDs**

La división de los SNDs que se adopta aquí, fue propuesta por Nauck [Nau95, NKK97]. La razón para usar tal taxonomía es que la mayoría de las aproximaciones neurodifusas implementadas hasta hoy, pueden ser enmarcadas dentro de estas

divisiones que se basan en las diferentes combinaciones de las RNA y los sistemas difusos:

- A. Modelos Concurrentes: Cuando la RNA y el sistema difuso trabajan juntos. Sin embargo, ninguno de ellos determina las características o parámetros del otro.
- B. Modelos Cooperativos: Cuando la RNA se usa para determinar los parámetros de funcionamiento del sistema difuso. En estos modelos, se distinguen dos fases: la de entrenamiento y la de funcionamiento. En la primera, la RNA interactúa con el sistema difuso determinando los parámetros del mismo, mientras en la segunda, la RNA desaparece dejando sólo el sistema difuso. Ejemplos de estos sistemas son: el ASR (Agente de Sintonización neuronal por Refuerzo) [GG97], la determinación de parámetros de control de un sistema difuso, etc.
- C. Modelos Híbridos: En esta aproximación, el sistema difuso y la RNA, trabajan juntos en una arquitectura homogénea que puede ser interpretada como una RNA con parámetros difusos o como un sistema difuso con parámetros o funcionamiento distribuidos. Algunos ejemplos son ANFIS [Jan93a] y NEFCON [NK93].

#### 2.4.2 Ventajas de los SNDs

La estructura neurodifusa permite el uso de métodos cualitativos y cuantitativos en la construcción de modelos, en la etapa de aprendizaje, en la de funcionamiento y en la realimentación de conocimiento al usuario, permitiendo así, todo el conocimiento disponible. Los métodos cuantitativos permiten realizar un aprendizaje a partir de datos disponibles del sistema y de la forma matemática de funcionamiento, mientras los métodos cualitativos reflejan el sentido lingüístico que tiene la base de conocimiento de los sistemas difusos y que genera una interacción bilateral sistema-hombre, permitiendo la introducción de conocimiento previo al sistema y la realimentación del estado final del modelo en forma de reglas expresadas lingüísticamente desde el sistema al hombre.

Las otras ventajas, son ya ampliamente conocidas y que vienen heredadas de los sistemas difusos y de las RNA [NKK97], tales como:

- *Autoaprendizaje*: La principal cualidad de los SNDs la constituye su adaptabilidad dinámica, esto es su capacidad para variar de comportamiento en situaciones cambiantes. Estos sistemas van modificando su comportamiento para ajustarlo a los datos que se le van proporcionando, lo

cual les brinda ciertos atributos de flexibilidad y adaptabilidad que son de gran interés en cualquier entorno.

- *Generalización:* Los SNDs tienen capacidades de aproximación a comportamientos no lineales (lo que implica la posibilidad de usarlos para diseñar predictores de comportamiento futuro o para la identificación de sistemas físicos), así como capacidades para la creación de grupos, los que les confiere la potencialidad de ser usados en el diseño de controladores y el diagnóstico y detección de fallas.
- *Introducción de conocimiento previo:* Los SNDs son un tipo de sistemas especialmente capacitados para la incorporación de conocimiento experto y tratamiento de incertidumbre. Las reglas de las que dispone el motor de inferencia de estos sistemas pueden ser formuladas por expertos, o bien aprendidas por el propio sistema.
- *Interpretación lingüística:* Los SNDs permiten trabajar con conocimiento representado en forma lingüística de muy clara interpretación.

### 2.4.3 Limitaciones de los SNDs

A pesar de las múltiples ventajas, los SNDs poseen diversas limitaciones:

- *Maldición de la dimensionalidad:* Los sistemas difusos presentan el problema del crecimiento exponencial del tamaño del modelo respecto al número de variables que poseen, haciendo impráctico su uso para sistemas de más de cuatro dimensiones. Este crecimiento se puede representar en dos formas diferentes:

A. Tamaño del modelo: El número de conjuntos multivariados de la entrada crece exponencialmente con el número de entradas:

$$p = \prod_{i=1}^n p_i \quad (28)$$

donde  $p$  es el número de conjuntos multivariados,  $n$  es el número de entradas y  $p_i$  es el número de conjuntos difusos para la entrada  $i$ .

Como resultado de este crecimiento exponencial, el costo de cálculo y de implementación crece en igual forma y la interpretabilidad se limita cada vez más.

B. Tamaño de las muestras: Si  $N_i$  es el tamaño de muestra apropiado para un sistema univariado, entonces  $N$ , representado en la ecuación 29, es el tamaño de muestra requerido para un sistema de  $n$  variables. En este caso, se

presentan los mismos resultados que en el caso anterior, y además el espacio muestral puede verse tan poblado que puede representar pobremente al sistema.

$$N = N_1^n \quad (29)$$

- *Calidad de los datos:* Existen problemas asociados con la cantidad, distribución y ruido de los datos. Si los datos están mal distribuidos, el aprendizaje puede no ser muy bueno, presentando especialmente problemas de generalización si faltan muchos datos en cierta región. Este inconveniente es herencia de los sistemas neuronales y puede ser superado haciendo un tratamiento previo de interpolación de los datos o buscando una representación adecuada de los mismos.
- *Conocimiento previo:* En los SNDs es deseable tener el conocimiento previo y confiable de un experto en el sistema, pues de otro modo, se debe realizar un modelado por "fuerza bruta" a través de los datos numéricos empíricos y pueden aparecer problemas al no disponer de una descripción cualitativa del sistema, a menos que se apliquen técnicas de minería de datos a la base de conocimiento del SND.

## 2.5 Algoritmos Genéticos

La *Inteligencia Artificial* (IA)<sup>12</sup> no sólo consiste en idear algoritmos y estructuras de datos para solucionar problemas. También trata acerca de la inteligencia humana, y por extensión, sobre la vida. Dentro de la IA, la vida artificial ofrece algunos mecanismos de resolución de problemas muy eficientes y originales. Además, toma muy en serio aspectos muy filosóficos que incluso están inspirados en la teoría de la evolución de las especies de Darwin<sup>13</sup> [Dar59] y en los descubrimientos realizados por Mendel en el campo de la genética, tratándose de imitar en lo posible a la *Naturaleza*. Este enfoque ha dado lugar a lo que en la actualidad se denomina *Computación Evolutiva* (CE).

En un sentido general, se puede indicar que la CE se refiere a una rama de la IA que trata acerca del conjunto de técnicas que abordan problemas complejos de búsqueda y aprendizaje basados en la emulación de la teoría de evolución biológica. Los distintos modelos computacionales que se han propuesto dentro de ésta filosofía

---

<sup>12</sup> Dar una definición de Inteligencia Artificial no es tan fácil como se podría esperar, pero se puede decir que es una ciencia, disciplina o inter-disciplina científica relacionada con el diseño de sistemas que exhiben las características que asociamos con la inteligencia en la conducta humana. Según [CBM95, p.1] su nombre se debe a J. McCarthy.

<sup>13</sup> Independientemente de Darwin, A. R. Wallace llegó a las mismas conclusiones.

suelen recibir el nombre genérico de *Algoritmos Evolutivos* (AEs). Existen cuatro tipos de AEs bien definidos que han servido como base a la mayoría del trabajo desarrollado en el área: los *Algoritmos Genéticos* (AGs), propuestos inicialmente por Holland [Hol62, Hol75] y han sido posteriormente estudiados en profundidad por otros autores [Gol89, Mic96, Gol02]; las *Estrategias Evolutivas* (EEs), desarrolladas por Rechenberg [Rec65, Rec73] y Schwefel [Sch65, Sch77, Sch95a, BS95]; la *Programación Evolutiva* (PE), en un principio introducida por Fogel [Fog62, Fog63, Fog64, FOW66] y luego extendida por su hijo [Fog88, Fog91, Fog92, FA92, Fog95, Fog98]; y la *Programación Genética* (PG) propuesta por Koza [Koz90, Koz92, Koz94] y profundizada por otros autores [Tac94, OO95, ACT99, HN01, SCC01, CLH02]. Entre las referencias principales sobre AEs están [Bäc96, BS96, dFa197, DD03]. En [BS93] se presenta un excelente resumen y estado del arte acerca de AEs, constituyendo un buen punto de partida para documentarse en el tema<sup>14</sup>.

Un AE se basa en mantener una población de posibles soluciones del problema a resolver, llevar a cabo una serie de alteraciones sobre las mismas y efectuar una selección para determinar cuáles permanecen en generaciones futuras y cuáles son eliminadas. Aunque todos los modelos existentes siguen esta estructura general, existen algunas diferencias en cuanto al modo de ponerla en práctica. Los AGs se basan en operadores que tratan de modelar los operadores genéticos existentes en la naturaleza, como el cruce y la mutación en un punto, los cuales son aplicados a los individuos que codifican las posibles soluciones. En cambio, las EEs [BHS91] y la PE aplican transformaciones basadas en mutaciones efectuadas sobre los padres para obtener los hijos, lo que permite mantener la línea general de comportamiento del individuo en su descendencia. Eliminar el cruce permite que en programación evolutiva la representación de los individuos pueda ser bastante más compleja y flexible. En las estrategias evolutivas la representación de cada individuo de la población consta de 2 tipos de variables: las variables objeto y las variables estratégicas. Las variables objeto son los posibles valores que hacen que la función objetivo alcance el óptimo global y las variables estratégicas son los parámetros mediante los que se gobierna el proceso evolutivo. Finalmente, la PG codifica las soluciones al problema en forma de programas, habitualmente codificados en una estructura de árbol, y adapta dichas estructuras empleando operadores muy específicos.

Cada individuo de la población recibe un valor de una medida de adaptación o fitness que representa su grado de adecuación al entorno. La selección hace uso de estos valores y se centra en los individuos que presentan mayor valor en la medida.

---

<sup>14</sup> J. Heitkötter y D. Beasley han elaborado también una guía completa sobre AEs, titulada “The Hitch-Hiker’s Guide to Evolutionary Computación: A list of Frequently Asked Questions (FAQ). USENET: comp.ai.genetic”, con muchos enlaces a artículos, libros, software y principales grupos de trabajo alrededor del mundo. Disponible vía <ftp://rtfm.mit.edu/pub/usenet/news.answers/ai-faq/genetic>.

Los operadores de recombinación y/o mutación alteran la composición de dichos individuos, guiando heurísticamente la búsqueda a través del espacio. Aunque simples desde un punto de vista biológico, este tipo de algoritmos son suficientemente complejos para proporcionar mecanismos de búsqueda adaptativos muy robustos. Los mismos procedimientos pueden ser aplicados a problemas de distintos tipos, sin necesidad de hacer muchos cambios [Fog95].

Esta sección se dedicará a describir la técnica de búsqueda que se utilizará en esta memoria: los *Algoritmos Genéticos* (AGs), los cuales están considerados actualmente como la técnica de búsqueda global más conocida y empleada. Los AGs son algoritmos de búsqueda de propósito general que se basan en principios inspirados en la genética de las poblaciones naturales para llevar a cabo un proceso evolutivo sobre soluciones de problemas.

Los AGs pertenecen a los llamados algoritmos probabilísticos, aunque son diferentes a los algoritmos basados al azar. En este sentido, los AGs son más potentes que los métodos de búsqueda local debido a que combinan elementos de este tipo de búsqueda con otros de búsqueda estocástica. Otra propiedad importante de estos métodos de búsqueda basados en genética es que mantienen una población de soluciones potenciales, en lugar de un único punto en el espacio de búsqueda. Este hecho permite que los AGs realicen búsquedas en múltiples direcciones. Es así como han demostrado ser, tanto desde un punto de vista teórico como práctico, una herramienta óptima para proporcionar una búsqueda robusta en espacios complejos, ofreciendo un enfoque válido para solucionar problemas que requieran una búsqueda eficiente y eficaz.

Este tipo de AE se ha aplicado con mucho éxito en problemas de búsqueda y optimización. Como se ha mencionado, la razón de gran parte de este éxito se debe a su habilidad para explotar la información que van acumulando sobre el espacio de búsqueda que manejan, desconocido inicialmente, lo que les permite redirigir posteriormente la búsqueda hacia subespacios útiles. La capacidad de adaptación que presentan los AGs es su característica principal, especialmente en espacios de búsqueda grandes, complejos y con poca información disponible, en los que las técnicas clásicas de búsqueda (enumerativas, heurísticas, ...) [AL97, Ree93, Ree96, VLA88, PS82] no presentan buenos resultados.

La idea básica de los AGs consiste en mantener una población de individuos que codifican soluciones del problema. Dichos individuos emplean una representación genética para codificar los valores de las características parciales que definen las distintas soluciones. Debido a ello, cada individuo recibe el nombre de *cromosoma* y cada una de sus componentes el de *gen*.

La *figura 8*, en la que  $P(t)$  denota la población en la generación  $t$ , muestra la estructura general de un AG básico.

```

Procedimiento Algoritmo Genético
EMPEZAR
  t = 0;
  inicializar P(t);
  evaluar P(t);
MIENTRAS NO (condicion de parada) HACER
  EMPEZAR
    t = t + 1;
    seleccionar Po(t) a partir de P(t - 1);
    cruzar y mutar Po(t);
    P(t) = Po(t);
    evaluar P(t);
  FIN
FIN

```

**Figura 8.** Estructura básica de un Algoritmo Genético.

Los cromosomas se generan inicialmente a partir de la información disponible sobre el problema, o bien de un modo aleatorio cuando no se dispone de esta información, y la población se hace evolucionar a lo largo del tiempo mediante un proceso de competición y alteración controlada que emula los procesos genéticos que tienen lugar en la naturaleza. A lo largo de sucesivas iteraciones, denominadas generaciones, los cromosomas se ordenan con respecto a su grado de adaptación al problema, es decir, con respecto a lo bien que resuelven dicho problema y, tomando como base estas evaluaciones, se construye una nueva población mediante un proceso de selección y una serie de operadores genéticos tales como el cruce y la mutación. Como en todos los AEs, es necesario diseñar una función de adaptación para cada problema que se desee resolver. Dado un cromosoma de la población, esta función devuelve un único valor numérico que se supone proporcional al grado de bondad de la solución que dicho cromosoma codifica. Esta función es la encargada de guiar al AG por el espacio de búsqueda. Por esta razón, debe estar bien diseñada para que sea capaz, no sólo de distinguir de un modo claro los individuos bien adaptados de los que no lo están, sino también de ordenar éstos en función de su capacidad para resolver el problema.

A continuación, se detallan brevemente los aspectos principales relacionados con los AGs: la representación de las soluciones, el mecanismo de selección y los operadores de cruce y mutación.

### 2.5.1 Representación de las Soluciones

El esquema de representación o codificación es un factor clave en la aplicación de los AGs, ya que estos manipulan directamente una representación codificada del problema y, por lo tanto, el esquema escogido puede limitar de una forma muy severa la vista desde la cual el AG afronta el problema. Existen distintos esquemas generales de codificación entre los que destacan los siguientes:

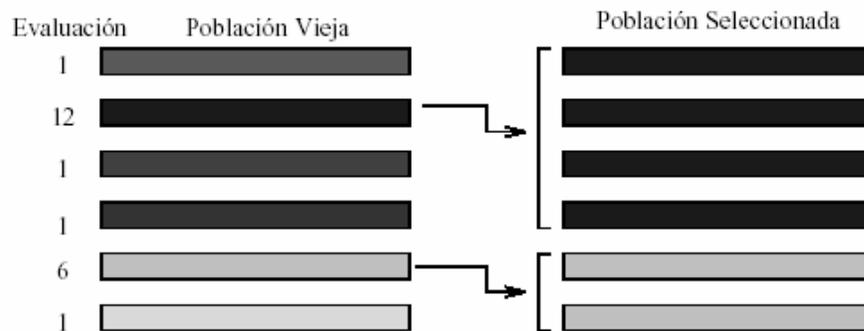
1. La *codificación binaria*: Es la más antigua de todas las existentes [Hol75, Gol89]. Se basa en la representación de los cromosomas como cadenas de bits de modo que, dependiendo del problema, cada gen del cromosoma puede estar formado por una subcadena de varios bits.
2. La *codificación real*: La codificación binaria presenta una serie de inconvenientes importantes cuando se trabaja con problemas que incluyen variables definidas sobre dominios continuos: excesiva longitud de los cromosomas, falta de precisión, etc. En los últimos años, se ha estudiado ampliamente la codificación real [HLV98, Wri91], más adecuada para este tipo de problemas. En este esquema de representación, cada variable del problema se asocia a un único gen que toma un valor real dentro del intervalo especificado, por lo que no existen diferencias entre el genotipo (la codificación empleada) y el fenotipo (la propia solución codificada). Es así como los valores de los genes están forzados a permanecer en el intervalo establecido para las variables a las cuales representan, de modo que los operadores genéticos tendrán que preservar este requerimiento. Gracias a esta propiedad se solucionan los problemas comentados.
3. La *codificación basada en orden*: Este esquema está diseñado específicamente para problemas de optimización combinatoria en los que las soluciones son permutaciones de un conjunto de elementos determinado [Gol89, Mic96]. Como ejemplos de este tipo de problemas podemos citar los conocidos problemas del viajante de comercio y del coloreado de grafos.

Además de estos esquemas generales de representación, se pueden emplear muchos otros particulares al tipo de problema que se esté resolviendo, siempre que sea necesario.

### 2.5.2 El Mecanismo de Selección

El mecanismo de selección es el encargado de seleccionar la población intermedia de individuos la cual, una vez aplicados los operadores de cruce y mutación, formará la nueva población del AG en la siguiente generación. De este

modo, si notamos por  $P$  la población actual formada por  $n$  cromosomas,  $C_1, \dots, C_n$ , el mecanismo de selección se encarga de obtener una población intermedia  $P_o$ , formada por copias de los cromosomas de  $P$  (Véase *figura 9*). El número de veces que se copia cada cromosoma depende de su adecuación, por lo que generalmente aquellos que presentan un valor mayor en la función de adaptación suelen tener más oportunidades para contribuir con copias a la formación de  $P_o$ .



**Figura 9.** Ejemplo de aplicación del mecanismo de selección.

Existen diferentes formas de poner en práctica la selección [BH91]. Por ejemplo, puede establecerse un paralelismo entre la población y una ruleta, en la que cada cromosoma está representado por un sector de la misma cuyo tamaño es proporcional a la adaptación de dicho cromosoma. Los cromosomas se seleccionan girando la ruleta tantas veces como individuos tengamos que seleccionar para formar la población intermedia. Este mecanismo de selección es uno de los más conocidos y se denomina *muestreo estocástico con reemplazamiento*. Uno de los más eficientes es el *muestreo universal estocástico*, propuesto por Baker en [Bak87], en el cual el número de copias de cada individuo en la población intermedia está acotado inferior y superiormente por un número de copias esperado calculado en función de su adaptación.

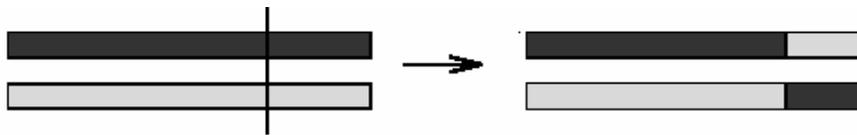
El mecanismo de selección puede ser complementado por el *modelo de selección elitista* introducido en principio por [DeJ75], basado en mantener un número determinado de los individuos mejor adaptados de la población anterior en la nueva población (la obtenida después de llevar a cabo el proceso de selección y de aplicar los operadores de cruce y mutación) [Gol89, Mic96]. Por último, algunos trabajos sobre comparaciones técnicas entre diferentes métodos de selección pueden consultarse en [Bak85, Whi89, GD91, dMT93, Han94].

### 2.5.3 El Operador de Cruce

Este primer operador constituye un mecanismo para compartir información entre cromosomas. Es considerado el principal operador genético y juega un papel central en los AGs. De hecho puede considerarse como una de las características principales que definen y diferencian a los AGs de otros algoritmos basados en la evolución natural, hasta el punto que se puede decir que un método de optimización evolutivo no es un AG si no se tiene operador de cruce, y sin embargo, puede serlo perfectamente sin operador de mutación, según descubrió Holland. Es por ello que es uno de los componentes que más se ha tenido en cuenta para mejorar el comportamiento de los AGs.

Este operador generalmente combina las características de dos cromosomas padre para obtener dos descendientes, con la posibilidad de que los cromosomas hijo, obtenidos mediante la recombinación de sus padres, estén mejor adaptados que éstos. No suele ser aplicado a todas las parejas de cromosomas de la población intermedia sino que se lleva a cabo una selección aleatoria en función de una determinada probabilidad de aplicación, la *probabilidad de cruce*,  $P_c$ .

El operador de cruce juega un papel fundamental en el AG. Su tarea es la de *explotar el espacio de búsqueda* refinando las soluciones obtenidas hasta el momento mediante la combinación de las buenas características que presenten. Como ya se ha comentado, tanto la definición del operador de cruce como la del de mutación dependen directamente del tipo de representación empleada. Por ejemplo, trabajando con el esquema de codificación binario, se suele emplear el clásico *cruce simple* en un punto [Mic96] basado en seleccionar aleatoriamente un punto de cruce e intercambiar el código genético de los dos cromosomas padre a partir de dicho punto para formar los dos hijos (Véase *figura 10*), el cruce multipunto [Mit97] que procede como el anterior pero trabajando sobre dos o más puntos de cruce, o el cruce uniforme [Sys89, SD91] en donde los valores de cada gen en los hijos son determinados por la elección de un valor uniforme aleatorio de los valores de estos genes en los padres. Otros tipos de operadores de cruce para este tipo de esquema de codificación se describen en [ECS89].



**Figura 10.** Ejemplo de aplicación del operador de cruce simple en un punto.

También se pueden emplear los operadores de cruce simple o multipunto

cuando se trabaja con el esquema de codificación real, aunque existe una serie de operadores diseñados para su uso específico con esta representación [HLV98]. Entre éstos, se puede destacar una familia de operadores que manejan técnicas basadas en Lógica Difusa para mejorar el comportamiento del operador de cruce [HLV97]. Como ejemplo de estos operadores, se introduce el operador de *cruce max-min-aritmético*. Dados dos cromosomas de la población  $P(t)$ ,  $C_v^t = (c_1, \dots, c_k, \dots, c_H)$  y  $C_w^t = (c'_1, \dots, c'_k, \dots, c'_H)$ , que van a ser cruzados, este operador genera los cuatro descendientes siguientes:

$$\begin{aligned}
 C_1^{t+1} &= aC_w^t + (1-a)C_v^t \\
 C_2^{t+1} &= aC_v^t + (1-a)C_w^t \\
 C_3^{t+1} &\text{ con } C_{3k}^{t+1} = \min \{c_k, c'_k\} \\
 C_4^{t+1} &\text{ con } C_{4k}^{t+1} = \max \{c_k, c'_k\}
 \end{aligned} \tag{30}$$

y escoge los dos mejor adaptados para formar parte de la nueva población. El parámetro  $a$  empleado en los dos primeros puede definirse como constante a lo largo de toda la ejecución del AG o variable dependiendo de la edad de la población.

Otra propuesta de cruce en el esquema de codificación real es el operador de *cruce aritmético* [Mic96], el cual pertenece al grupo de operadores de cruce basados en agregación (OCA) y entre los cuales están: el Geométrico [MNM96], lineal BGA [Sch94, HLV98, OHM01], heurístico [VMC95] y dinámico heurístico [HLV96]. El operador de cruce aritmético genera dos descendientes como una media ponderada del valor de los padres. Un valor real, “ $u$ ”, en el rango  $[0..1]$  se selecciona de manera aleatoria y se usa para generar los nuevos hijos por medio de ecuación 31. En la literatura se indica que si “ $u$ ” es una constante se denomina cruce aritmético uniforme, de lo contrario si varía a lo largo del tiempo, es decir, de sucesivas iteraciones pasaría a llamarse cruce aritmético no uniforme.

$$\begin{aligned}
 C_t &= u \cdot PADRE + (1-u) \cdot MADRE \\
 C'_t &= (1-u) \cdot PADRE + u \cdot MADRE
 \end{aligned} \tag{31}$$

Este operador de cruce es considerado el más adecuado para las representaciones de cromosomas en donde los descendientes deban cumplir también con alguna regla o principio general, por ejemplo que la sumatoria de sus valores de 1. Es por ello que a partir de dos padres “válidos”, se obtienen dos hijos igualmente “válidos” (siempre que se utilice el mismo valor de “ $u$ ” para generar todos los genes descendientes de esa pareja de padres).

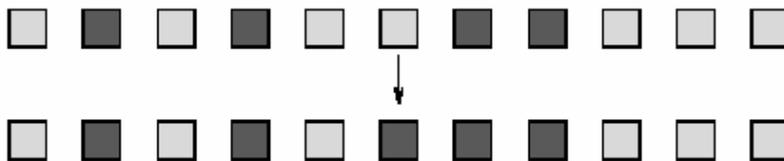
En el caso de la codificación real, los operadores de cruce influyen decisivamente sobre el nivel de diversidad en la población, y por ello, son factores determinantes para evitar el problema de convergencia prematura [BD01]. Estudios en donde los operadores tienen múltiples descendientes (más de dos) a partir de dos padres, han funcionado mejor globalmente frente a los operadores con solo dos descendientes, alcanzando un buen equilibrio entre la exploración y la explotación. Algunas referencias relevantes son [Wri91, HLV97, HLV96, ELG97].

### 2.5.4 El Operador de Mutación

Este operador altera arbitrariamente uno o más genes del cromosoma seleccionado con el propósito de aumentar la diversidad de la población. Todos los genes de los cromosomas existentes están sujetos a la posibilidad de mutar de acuerdo a una *probabilidad de mutación*,  $P_m$ .

En este caso, la propiedad de búsqueda asociada al operador de mutación es la *exploración*, ya que la alteración aleatoria de una de las componentes del código genético de un individuo suele conllevar el salto a otra zona del espacio de búsqueda que puede resultar más prometedora.

El operador de mutación clásicamente empleado en los AGs con codificación binaria se basa en cambiar el valor del bit seleccionado para mutar por su complementario en el alfabeto binario [Hol75, Gol89], tal y como se muestra en la *figura 11*.



**Figura 11.** Ejemplo de aplicación del operador de mutación.

Este operador puede trasladarse al campo de los AGs con codificación real, de forma que el nuevo valor del gen mutado se escoja aleatoriamente dentro del intervalo de definición asociado [Thr91]. Al igual que en el caso del operador de cruce, existen distintos operadores de mutación específicos para trabajar con esta codificación [HLV98, Mic96]. De entre ellos se destacan el operador de *mutación no uniforme* propuesto por Michalewicz en esta última referencia, el cual ha demostrado buen comportamiento en numerosas aplicaciones basadas en codificación real. A

continuación, se describe su modo de trabajo.

Dado un cromosoma de la población  $P(t)$ ,  $C_v^t = (c_1, \dots, c_k, \dots, c_H)$ , y uno de sus genes,  $c_k$ ,  $k \in 1, \dots, H$ , definido en  $[c_{ki}, c_{kd}]$ , seleccionado para ser mutado, el cromosoma obtenido tras la mutación presenta la forma  $C_v^{t+1} = (c_1, \dots, c'_k, \dots, c_H)$ , con

$$c'_k = \begin{cases} c_k + \Delta(t, c_{kd} - c_k), & \text{si } p = 0 \\ c_k - \Delta(t, c_k - c_{ki}), & \text{si } p = 1 \end{cases} \quad (32)$$

donde  $p$  es un número aleatorio generado en  $\{0,1\}$  y la función  $\Delta(t, y)$  devuelve un valor en el intervalo  $[0, y]$ , de modo que la probabilidad de que  $\Delta(t, y)$  sea cercana a 0 aumenta cuando lo hace el contador de generaciones  $t$ :

$$\Delta(t, y) = y(1 - r^{(1-\frac{t}{T})^b}) \quad (33)$$

donde, a su vez,  $r$  es un número aleatorio generado en  $[0, 1]$ ,  $T$  es el número de generaciones durante las que se ejecutará el AG y  $b$  es un parámetro escogido por el usuario que determina el grado de dependencia existente con respecto al número de generaciones. Esta propiedad da lugar a que el operador lleve a cabo una búsqueda uniforme en el espacio cuando  $t$  es pequeña, es decir, en las primeras iteraciones, y una mucho más localizada en generaciones posteriores.

Otras propuestas de mutación indicadas en la literatura para el esquema de codificación real son: la Random [Mic96], Real Number Creep [Dav91b], Mühlentien [MSV93], la Modal Discreta y la Modal Continua propuestas por [VA94].

### 2.5.5 Características principales de los Algoritmos Genéticos

La aplicación de un AG a un problema [BSEK06] conlleva el análisis y diseño de los cinco componentes siguientes:

1. Una representación genética de las posibles soluciones al problema.
2. Un mecanismo de creación de la población inicial de soluciones.
3. Una función de adaptación que asocie un valor de adecuación a cada cromosoma.
4. Una serie de operadores genéticos que alteren la composición genética de los descendientes durante el proceso de reproducción, así como un mecanismo de selección de los padres que intervienen en dicho proceso.
5. Una serie de valores para los parámetros del AG: tamaño de la población,

probabilidades de aplicación de los operadores genéticos, etc.

Los AGs han sido aplicados con éxito en una gran cantidad de áreas distintas. Algunas de las razones de este éxito son las siguientes:

1. Pueden resolver problemas muy complejos de una forma rápida y sencilla.
2. Pueden ser enlazados fácilmente con simulaciones y modelos existentes.
3. Presentan gran facilidad para ser hibridizados con otras técnicas.

En resumen, los AGs son muy robustos. Destacan por su buen comportamiento en problemas difíciles en los que el espacio de búsqueda es grande, discontinuo, complejo y poco conocido. Aunque no garantizan la obtención de la solución óptima al problema, suelen proporcionar soluciones bastante aceptables en un tiempo razonable.

Todas las razones anteriores han dado lugar a que su aplicación haya crecido enormemente a lo largo de los últimos años en campos tan diversos como la optimización de funciones numéricas y la optimización combinatoria [Gol89, Mic96], la Investigación Operativa [Nis93], las Redes Neuronales [WS92], los Sistemas Difusos [HV96, Ped97], la Robótica [Dav91a], la Vida Artificial [BB91, For93], el Aprendizaje Automático [DSG93, Jan93b], etc.

## **CAPITULO III:**

---

### **Un Método de Aprendizaje de la Granularidad para Razonamiento Inductivo**



## Capítulo III: Un Método de Aprendizaje de la Granularidad para Razonamiento Inductivo

La granularidad de las particiones difusas constituye uno de los parámetros necesarios para el proceso de fusificación de la *Metodología de Razonamiento Inductivo Difuso* (FIR). Se ha señalado varias veces la importancia de este elemento en el proceso de modelado y predicción de FIR, además de comentar la dificultad que supone una búsqueda exhaustiva del número de clases (etiquetas, términos lingüísticos) más adecuado por variable. El objetivo de este capítulo es diseñar un método automático de búsqueda de la granularidad de las variables que componen el sistema en un tiempo razonable.

En la primera sección, se analiza la problemática planteada y se establecen sus posibles vías de solución, asimismo se justifica la selección del método más idóneo para abordarla. Luego, se describen los componentes del método propuesto para el aprendizaje automático de la granularidad. Se finaliza este capítulo analizando las ventajas e inconvenientes de este método.

### 3.1 Algoritmo genético para encontrar de forma automática la granularidad: AG1

Dado un problema concreto, se persigue encontrar el nivel de granularidad óptimo para cada variable manteniendo particiones difusas uniformes. Por tanto, se plantea un problema con  $8^N$  posibles soluciones, siendo  $N$  el número de variables del problema y 8 el cardinal del conjunto de valores posibles para las clases (es decir, el conjunto  $\{2, \dots, 9\}$ ). Para esta situación no es posible utilizar técnicas clásicas como un algoritmo de *greedy*. Esto se debe a que este tipo de algoritmos generan soluciones a los problemas de un modo constructivo, tomando en cada paso la mejor decisión posible. Para ello, es necesario disponer de una preferencia heurística sobre las posibles decisiones para cada variable del problema. Es fácil ver que en la problemática aquí planteada, no se dispone de dicha información heurística.

El Recocido Simulado (RS), el Hill-Climbing y los AGs son métodos de optimización de propósito general que comparten muchas similitudes. Estos métodos requieren poco conocimiento del problema a ser optimizado y comienzan inicialmente con unas soluciones dentro del espacio de búsqueda que luego van cambiando incrementalmente por la aplicación de operadores. No obstante poseen

ciertas características diferenciables entre ellos. El RS y el Hill-Climbing son técnicas de búsqueda local que utilizan una metodología de mejora en forma interactiva, que permite escoger soluciones peores a la actual en algunos casos pero con unos parámetros iniciales adecuados pueden encontrar la solución óptima. Por consiguiente, estos métodos de búsqueda obtienen valores óptimos locales y dependen fuertemente del punto inicial.

Ante la dimensión de la problemática planteada, parece adecuada la utilización de algún algoritmo de búsqueda local como el Recocido Simulado (RS). Sin embargo, se decide trabajar con los AGs por los siguientes motivos:

- Se cree que los AGs van mejor para estudios de problemas de optimización complejos, es decir, que posean más de tres variables y en el área industrial o empresarial son muy frecuentes los sistemas con éstas características. En este caso, la rapidez en encontrar una buena solución resulta importante porque de lo contrario se tendrían pocas ventajas incluso frente a una búsqueda exhaustiva. No obstante, el paralelismo implícito de los AGs les permite encontrar con éxito resultados óptimos o razonables en un corto período de tiempo, tras muestrear sólo regiones pequeñas de un vasto paisaje adaptativo [For93, p.877]. Investigaciones previas tales como [GP01,Sch99] sugieren la escalabilidad de los AGs.
- Las características de la problemática suscitada se adaptan muy bien a la forma de trabajar de los AGs, ya que el espacio de soluciones es finito y es factible realizar una codificación de las soluciones sin excesiva dificultad.
- Los AGs corresponden a los algoritmos probabilísticos, sin embargo son diferentes a los algoritmos basados en el azar. Es así como los AGs son más robustos que los métodos de búsqueda local, ya que combinan elementos de este tipo de búsqueda con otros de búsqueda estocástica. Asimismo, estos métodos basados en genética tienen la facultad de mantener una población de soluciones potenciales, en lugar de un único punto en el espacio de búsqueda. Este hecho les permite que realicen búsquedas en múltiples direcciones, siendo por ello intrínsecamente paralelos. Sin embargo, la ventaja del paralelismo va más allá de esto, un AG que evalúe explícitamente un número pequeño de individuos está evaluando implícitamente un grupo de individuos mucho más grande, esto se conoce como teorema del esquema y es la ventaja principal de los AGs sobre otros métodos de resolución de problemas [Mit97, p. 28-29; Gol89, p. 20].
- En relación a la distancia entre el punto que es la mejor solución obtenida en un momento dado y los puntos próximos que puedan mejorar esa solución en el espacio de búsqueda, los AGs son catalogados como de rango largo. Esta característica les permite evitar fácilmente caer, dentro del espacio de

búsqueda, en una zona que constituya un óptimo local. Por consiguiente, este hecho mejora la capacidad del algoritmo de detectar regiones donde existe una alta posibilidad de encontrar un óptimo global incluso en paisajes adaptativos muy escabrosos y complejos (aquéllos en los que la función objetivo es discontinua, ruidosa, cambia con el tiempo, o tiene muchos óptimos locales). Debe indicarse que, en la realidad, a menudo no hay manera de decir si una cierta solución a un problema es el óptimo global o sólo un óptimo local muy alto. Sin embargo, aunque un AG no devuelva siempre una solución perfecta y demostrable a un problema, casi siempre puede devolver al menos una muy buena solución.

- Finalmente, una de las cualidades de los AGs que, a primera vista, puede parecer una desventaja, resulta ser una de sus ventajas: no saben nada de los problemas que deben resolver. En lugar de utilizar información específica conocida a priori para guiar cada paso y realizar cambios con un ojo puesto en la mejora como hacen los diseñadores humanos, son “relojeros ciegos” [Daw96] ya que realizan cambios aleatorios en sus soluciones candidatas y luego utilizan la función objetivo para determinar si esos cambios producen una mejora. La virtud de esta técnica es que permite a los AGs comenzar con una mente abierta, por así decirlo. Como sus decisiones están basadas en la aleatoriedad, todos los caminos posibles de búsqueda están abiertos teóricamente. En contraste, cualquier estrategia de resolución de problemas que dependa de un conocimiento previo, debe inevitablemente comenzar descartando muchos caminos a priori, perdiendo así cualquier solución novedosa que pueda existir [KFAK99, p. 547]. Los AGs al carecer de ideas preconcebidas basadas en creencias establecidas sobre “como deben hacerse las cosas” o sobre lo que “de ninguna manera podría funcionar”, no tienen ese conflicto. De manera similar, cualquier técnica que dependa del conocimiento previo fracasará cuando no esté disponible el mismo, pero de nuevo los AGs no se ven afectados negativamente por la ignorancia [Gol89, p.23].

Es por todo ello que en esta investigación nos hemos decantado por estos algoritmos de búsqueda para diseñar y desarrollar métodos automáticos de aprendizaje para los diferentes parámetros del proceso de fuzziificación de FIR. Concretamente en este capítulo se propone un método de aprendizaje de particiones difusas en el contexto de la metodología FIR usando *Algoritmos Genéticos* (AGs) con el propósito de buscar un buen nivel de granularidad para todas las variables. Los cromosomas de dicho AG codificarán el número de clases (granularidad) de cada variable. Para evaluar dichos cromosomas, se proponen dos funciones objetivo diferentes: a) la calidad de la máscara óptima y b) el error de predicción MSE (Véase ecuación 35) de una parte del conjunto de datos de entrenamiento. La distribución de las clases o términos lingüísticos a lo largo de los dominios de las variables será

uniforme, es decir, se usará el método EFP para obtener las funciones de pertenencia. Las secciones siguientes presentan el método propuesto describiendo los componentes principales del AG.

## 3.2 Aspectos principales

Los operadores genéticos principales utilizados son los de cruce y mutación. Se emplea elitismo para no perder la mejor solución encontrada. Dado que el espacio de búsqueda que se maneja es finito, el AG finalizará después de un número estipulado de evaluaciones de cromosomas que no debería superar el de posibles soluciones del problema en estudio. En caso de que el algoritmo parase en el máximo número de soluciones posible, no se obtendría ninguna ventaja frente a una búsqueda exhaustiva y ni siquiera estaría garantizado que se hubiera encontrado la solución óptima. La salida del AG será la granularidad que ha obtenido el mejor valor en la función objetivo utilizada.

### 3.2.1 Codificación de las soluciones

El componente que se pretende aprender con este AG es el número de clases (etiquetas, términos lingüísticos) de cada variable. Para un sistema con  $N$  variables (incluyendo las de entrada y las de salida), el número de términos lingüísticos de cada variable se codifica en un vector entero de longitud  $N$ . En esta memoria, los posibles valores considerados son los que están dentro del conjunto  $\{2, \dots, 9\}$ . Si denotamos por  $E_i$  el número de clases (etiquetas) de la variable  $i$ , podemos representar un cromosoma ( $C$ ) del siguiente modo:

$$C = (E_1, E_2, \dots, E_N) \quad (34)$$

Lógicamente, cada vez que cambie el número de clases de una variable por la acción de los operadores genéticos, se tendrá que construir la partición difusa con la distribución uniforme correspondiente usando el método EFP.

### 3.2.2 Población inicial

La población inicial estará compuesta por dos bloques diferentes. No se admiten cromosomas repetidos. El proceso de generación se describe a continuación:

- En el primer bloque, cada cromosoma tiene el mismo número de clases para todas las  $N$  variables del problema.

- En el segundo bloque, dentro de cada cromosoma puede haber distintos valores para el número de clases de las variables (escogidos aleatoriamente).

### 3.2.3 Evaluación de los cromosomas

Para evaluar un cromosoma se siguen los siguientes pasos:

- 1) Se decodifica la información del cromosoma, construyendo su correspondiente partición difusa en la estructura de datos que utiliza FIR para representar dichas particiones.
- 2) Se ejecuta el proceso de identificación de modelado cualitativo de la metodología FIR con los datos de entrenamiento, usando la partición construida en el paso anterior. Por lo tanto, se obtiene la máscara asociada a esa partición con la medida más alta de calidad.
- 3) Se calcula una función objetivo, de adaptación o evaluación. En esta memoria se proponen dos funciones objetivos: a) la calidad de la máscara óptima ó b) el error de predicción MSE de parte del conjunto de datos de entrenamiento.

Como se ha señalado, en el proceso de identificación del modelo cualitativo de la metodología FIR, la máscara óptima (la mejor estructura del modelo) es identificada por medio de una medida de calidad,  $Q$  (Véase Sección 1.2). La calidad de una máscara es un valor entre 0 y 1, donde 1 indica la más alta calidad. Por lo tanto, la primera función objetivo propuesta es  $1-Q$ , debido al hecho de que la tarea del algoritmo es minimizar la función objetivo.

La segunda función objetivo se define como la predicción del error de una porción del conjunto de datos de entrenamiento. El error cuadrático medio normalizado en porcentaje (MSE), dado en la ecuación (35), se usa para este propósito.

$$MSE = \frac{E[(y(t) - \hat{y}(t))^2]}{y_{\text{var}}}.100\% \quad (35)$$

$\hat{y}(t)$  es la salida predicha,  $y(t)$  la salida del sistema y  $y_{\text{var}}$  denota la varianza de  $y(t)$ . La idea es usar parte del conjunto de datos de entrenamiento para identificar el modelo y el resto del conjunto de datos para evaluar la eficiencia (*performance*) de la predicción del modelo encontrado con la partición difusa evaluada. Cabe recordar que el modelo FIR está compuesto por la máscara óptima y la base de reglas patrón (*I/O matrix*). Esto implica que ambos han de ser generados en el proceso de evaluación de una determinada partición difusa cuando se usa esta función objetivo. El error de predicción de la porción del conjunto de datos de entrenamiento no usado en el

proceso de identificación del modelo es usado como la función objetivo para el AG. Por lo tanto, es necesario llamar a la función de predicción de la metodología FIR para obtener el costo del cromosoma evaluado. En este sentido, es evidente que el costo computacional de esta función objetivo será notablemente mayor al obtenido con la función que depende únicamente de la calidad de la máscara. Sin embargo, el poder de predicción del modelo óptimo resultante de la partición propuesta en este caso será muy posiblemente superior. El tamaño de la porción del conjunto de datos de entrenamiento usado para la evaluación de la función objetivo propuesto se define con respecto al tamaño del conjunto de datos de entrenamiento disponible. Concretamente se ha especificado que para conjuntos de datos de entrenamiento con un número de instancias superior a 2000 el 25% final se usa para la evaluación de cada cromosoma, si se disponen de entre 400 y 2000 instancias se usa el 20% final de éstas y si es menor o igual a 400 instancias se utilizaría sólo el 8% final.

### 3.2.4 Operadores genéticos

En lo referente a los operadores genéticos, se ha considerado la estructura del cromosoma (con un solo nivel de información) para obtener una definición adecuada que haga el mejor uso posible de la representación escogida. Estos operadores se describen a continuación:

#### 3.2.4.1 Selección

El cálculo de la probabilidad de selección sigue el *linear ranking* [Bak85]. En este proceso, los cromosomas son ordenados de acuerdo al valor de adecuación de la función objetivo y entonces la probabilidad de selección de cada cromosoma,  $p_s(C_i)$ , se calcula de acuerdo a su rango,  $rank(C_i)$  (con  $rank(C_{mejor}) = 1$ ) utilizando la siguiente función de asignación no incremental:

$$p_s(C_i) = \frac{1}{NC} \cdot \left( \eta_{max} - (\eta_{max} - \eta_{min}) \cdot \frac{rank(C_i) - 1}{NC - 1} \right) \quad (36)$$

donde  $NC$  es el tamaño de población y  $\eta_{min} \in [0,1]$  especifica el número esperado de copias para el peor cromosoma (el mejor de ellos tiene un  $\eta_{max} = 2 - \eta_{min}$  copias esperadas). En los experimentos  $\eta_{min} = 0.75$ .

El *linear ranking* se realiza conjuntamente con el Muestreo Universal Estocástico, propuesto por Baker [Bak87]. Este procedimiento garantiza que el número de copias de cualquier cromosoma sea acotado por el piso y el techo de su número esperado de copias. Adicionalmente, a este operador de reproducción se le incorpora una selección elitista.

### 3.2.4.2 Cruce

Se utiliza un operador que dependerá de las características de los dos cromosomas padres implicados en el cruce:

- *Cruce cuando los dos padres tienen la misma granularidad en cada una de las variables:* Si los dos cromosomas tienen los mismos valores (cada variable tiene el mismo número de clases en ambos padres), se mantienen los valores en ambos descendientes.
- *Cruce cuando los dos padres tienen distinta granularidad en alguna variable:* En este segundo caso, parece interesante utilizar la información contenida en ambos padres para explorar el espacio de búsqueda y tratar de localizar nuevas zonas prometedoras. El funcionamiento del operador de cruce es sencillo, se selecciona aleatoriamente un punto de corte dentro del cromosoma y se cruzan las dos partes en ambos cromosomas de acuerdo al operador clásico de cruce.

A continuación, se muestra una representación gráfica del funcionamiento de este operador de cruce. Sean:

$$\begin{aligned} C_t &= (E_1, \dots, E_p, E_{p+1}, \dots, E_N) \\ C'_t &= (E'_1, \dots, E'_p, E'_{p+1}, \dots, E'_N) \end{aligned} \quad (37)$$

dos individuos que se van a cruzar en el punto  $p$ , los dos descendientes resultantes son:

$$\begin{aligned} C_t &= (E_1, \dots, E_p, E'_{p+1}, \dots, E'_N) \\ C'_t &= (E'_1, \dots, E'_p, E_{p+1}, \dots, E_N) \end{aligned} \quad (38)$$

### 3.2.4.3 Mutación

Debido a la naturaleza de los valores almacenados en el cromosoma, el operador de mutación seleccionado es similar a uno propuesto por Thrift en [Thr91], el cual indica que al gen del cromosoma seleccionado se le aumenta o disminuye en una unidad la granularidad asociada (la decisión es tomada de manera aleatoria). Cuando el valor que se pretende cambiar es el mínimo (2) ó el máximo (9), se realiza el único cambio posible, esto es, incrementar o decrementar en uno la granularidad, respectivamente.

### 3.3 Comentarios

No es recomendable utilizar el algoritmo propuesto en problemas de optimización sencillos, es decir, con menos de 4 variables. Aunque el **AG1** puede implementarse en este caso, una búsqueda exhaustiva ó un algoritmo de RS irían más rápidos. Sin embargo, en el sector industrial y empresarial se requiere usualmente obtener modelos de sistemas complejos en donde intervienen un número elevado de variables, por lo cual la efectividad del **AG1** quedaría demostrada al permitir obtener un buen nivel de granularidad sin disponer del conocimiento de algún experto.

En los sistemas difusos, el número de etiquetas (clases) por variable influye decisivamente en el comportamiento final del sistema difuso resultante. Como se ha mencionado anteriormente, en investigaciones previas se ha probado como la granularidad tiene una influencia determinante sobre el número máximo de reglas y en el comportamiento de los *Sistemas Basados en Reglas Difusas* [CHV00, Vil00]. La metodología FIR no es la excepción debido a que también está basada en lógica difusa y la determinación de los parámetros necesarios en el paso de discretización llega a ser relevante para la identificación de un buen modelo que capture el comportamiento del sistema de manera precisa. Por tanto, no parece lógico que una decisión que afecta directamente a la precisión e interpretabilidad de los sistemas difusos se deje poco más o menos que “al azar” (al uso de valores por defecto). Frecuentemente los valores por defecto definen la misma granularidad para todas las variables. En FIR, el número de términos lingüísticos utilizado por defecto es de 3 para todas las variables que intervienen en el sistema bajo estudio.

Que todas las variables tengan el mismo número de términos sólo se puede explicar por la comodidad que supone, ya que no todas las variables tienen la misma influencia sobre el problema, ni los datos disponibles suelen estar distribuidos de la misma forma sobre sus universos de discurso. En este sentido, la determinación automática de buenos parámetros de fusificación en la metodología FIR es una alternativa de gran interés y utilidad al uso de valores heurísticos y/o por defecto. En el Capítulo VI se presentan varios casos de estudio en donde se aplica el **AG1**, se compara con otros métodos y se observa su funcionalidad.

Por consiguiente, el método presentado en este capítulo permite plantear la posibilidad de diseñar otros métodos para obtener buenas particiones difusas para la metodología FIR, desde una doble perspectiva:

- Aprender las funciones de pertenencia de los conjuntos difusos, permitiendo hacer un *refinamiento* ó ajuste *local* dado un nivel de granularidad.
- Aprender la granularidad y las funciones de pertenencia “al mismo tiempo”, permitiendo un ajuste *global* sobre las particiones difusas.

## **CAPITULO IV:**

---

### **Un Método de Aprendizaje de las Funciones de Pertenencia para Razonamiento Inductivo**



## Capítulo IV:

# Un Método de Aprendizaje de las Funciones de Pertenencia para Razonamiento Inductivo

La psicología ha estudiado durante bastante tiempo cómo el contexto es capaz de influenciar nuestras percepciones. Un mismo estímulo en contextos diferentes origina diferentes percepciones. Se ha indicado en trabajos previos [GGP98, OP95] como un buen ajuste de las funciones de pertenencia es crítico cuando se evalúa la efectividad de los sistemas difusos en problemas de control, modelado o clasificación. La adaptación lineal de contextos es simple y rápida, pero las funciones de pertenencia obtenidas se distribuyen uniformemente. En cambio la adaptación no lineal de contextos es costosa computacionalmente, pero las funciones de pertenencia pueden estrecharse o ensancharse para representar mejor conceptos en entornos reales con alta sensibilidad en las clases extremas o en las clases intermedias. En la *Metodología de Razonamiento Inductivo Difuso* (FIR), las funciones de pertenencia constituyen el otro parámetro necesario para llevar a cabo el proceso de fusificación y una definición adecuada de las mismas resulta también muy importante.

En el capítulo anterior se propuso un método para el aprendizaje de la granularidad y se apuntó la posibilidad de extender esta idea para la definición de las funciones de pertenencia de los conjuntos difusos para FIR dado un nivel de granularidad. En este capítulo se aborda dicha propuesta desarrollando un método de aprendizaje que utiliza un AG que permite definir la función de pertenencia asociada a cada término lingüístico. En este caso, la granularidad podría obtenerse si se dispone de información experta acerca del problema que se esté estudiando o bien haciendo uso del AG1.

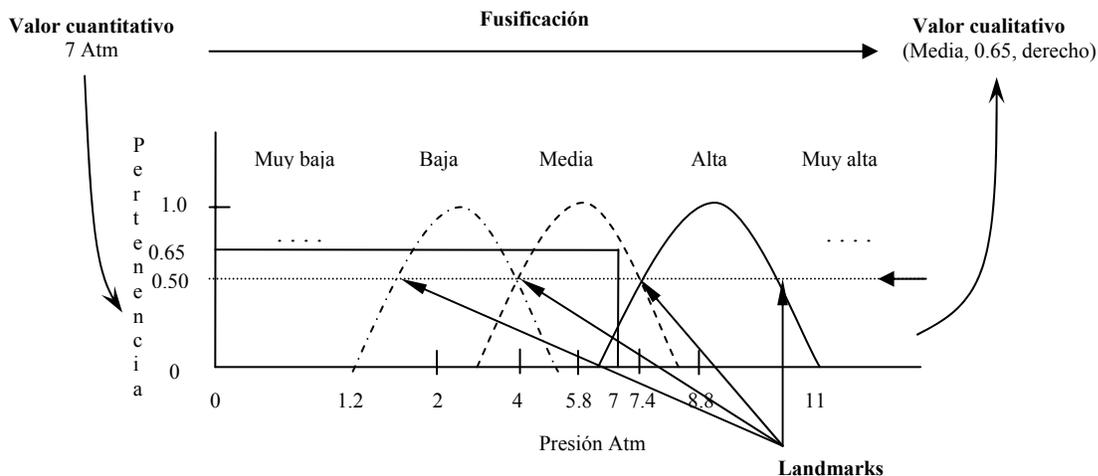
En primer lugar, se establece cual es la representación útil y eficaz para codificar las funciones de pertenencia en FIR. Seguidamente, se presentan los aspectos fundamentales del método propuesto para la determinación automática de la forma de las funciones de pertenencia dada la granularidad. Por último, se señalan las ventajas e inconvenientes del mismo.

## 4.1 Algoritmo genético para encontrar de manera automática la forma de las funciones de pertenencia con un número fijo de clases por variable: AG2

Los cromosomas del AG propuesto en este capítulo codifican las funciones de pertenencia que definen la semántica de las etiquetas o clases. Por consiguiente, es primordial definir una buena representación de estos cromosomas. Sin embargo para definir una codificación del cromosoma idónea, es necesario conocer profundamente el proceso de fusificación de la metodología FIR. En el Capítulo I de esta tesis doctoral se presenta éste proceso, aquí sólo recordamos los aspectos más relevantes.

Las formas más comunes de las funciones de pertenencia en FIR son la triangular y la gaussiana. Esta última es la considerada por defecto. La *figura 13* ilustra el proceso de fusificación por medio de un ejemplo. Como se mencionó anteriormente, un valor cuantitativo se convierte en una tripleta cualitativa, donde el primer elemento de la tripleta es el valor de la clase, el segundo elemento el valor de pertenencia difusa y el último elemento corresponde al valor del lado. El valor de lado, el cual es específico de FIR y no es usado comúnmente en lógica difusa, es el responsable de preservar en la tripleta cualitativa el conocimiento completo del valor cuantitativo original.

En la *figura 12* una presión de 7 atm podría ser fusificada en la clase "media", con un valor de pertenencia a esta clase de 0.65 y un valor de lado "derecho".



*Figura 12. Proceso de fusificación de la metodología FIR, variable presión*

Muchos enfoques de inferencia difusa preservan parte del conocimiento por medio de la asociación de los valores cuantitativo con múltiples reglas difusas formadas por tuplas de clases y valores de pertenencia. Estos representarían así la presión de 7 atm como “media” con un grado de pertenencia a esta clase de 0.65 y “alta” con un grado de pertenencia de 0.35. FIR es capaz de preservar la información completa por medio de reglas patrón formadas por tripletas cualitativas. Es por ello, que en la metodología FIR las colas de las funciones de pertenencia se descartan y solamente se usa la parte de las funciones de pertenencia en el rango [0.5...1]. El punto donde dos clases adyacentes se cortan con el valor de pertenencia de 0.5 es denominado landmark. Es así como, el componente a ser optimizado por este AG es la anchura de la función de pertenencia de cada clase, especificada por ambos landmarks. En el ejemplo de la *figura 12* la función de pertenencia de la clase “media” es definida por los landmarks {4,7.4}, siendo este par los valores de la presión que especifica los límites entre la clase “media” y sus clases adyacentes, “baja” y “alta”, respectivamente. Para mayores detalles de la metodología FIR, referirse a [CNMdA96, NCV98, Neb94].

Para evaluar dichos cromosomas se ejecutará una función de costo FIR. Se implementan dos funciones objetivo diferentes: a) la calidad de la máscara óptima y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento. Las secciones siguientes presentan el método propuesto describiendo los componentes principales del AG.

## 4.2 Aspectos principales

Las funciones de pertenencia consideradas serán gaussianas. Los operadores genéticos principales utilizados son los de cruce y mutación. Se emplea elitismo para no perder la mejor solución encontrada. El espacio de búsqueda se incrementa considerablemente en relación al **AG1**. Su finalización será después de un número estipulado de evaluaciones de cromosomas. La salida del AG será la distribución de landmarks que ha obtenido el mejor valor en la función de costo utilizada.

### 4.2.1 Codificación de las soluciones

La representación genética elegida toma en cuenta el número de elementos disponibles para cada variable. Una variable específica está representada por la proporción de elementos que contiene cada clase del total de datos disponibles para esa variable, codificada en el rango [0..1]. Un ejemplo de representación de un cromosoma para una única variable que tiene 4 clases podría ser (0.1,0.4,0.3,0.2),

significando que la función de pertenencia de la primera clase contiene el 10% de los elementos del total de datos disponibles para esta variable, mientras que la segunda, tercera y cuarta funciones de pertenencia contienen el 40%, 30% y 20% de los registros de datos disponibles para esa variable, respectivamente. La suma de las proporciones para cada variable debe ser 1. Es así como el componente que se pretende aprender con este AG es la distribución de los landmarks de cada variable. A esta representación la denominaremos *técnica de Proporción de Datos* (PD).

Por lo tanto, si denotamos  $D_{ij}$  como la proporción de datos para la clase  $i$  de la variable  $j$ , una representación del cromosoma ( $C$ ) para un sistema de  $N$  variables (incluyendo entradas y salida) con  $n$  clases por variable, se define del siguiente modo:

$$C = (D_{11} \dots D_{n1}, D_{12} \dots D_{n2}, \dots, D_{1N} \dots D_{nN}) \quad (39)$$

La mínima proporción ( $V_{MIN}$ ) se ha establecido en 0.05 y la máxima proporción ( $V_{MAX}$ ) es definida por  $V_{MAX} = 1 - V_{MIN} * (NETIQ - 1)$ , donde  $NETIQ$  es el número de clases de la variable. Lógicamente, cada vez que cambie la distribución de los landmarks de una variable por la acción de los operadores genéticos, se tendrá que construir la partición difusa correspondiente a esa nueva distribución. La ventaja primordial de esta representación es que ésta facilita la obtención de los landmarks, el cual se lleva a cabo de la siguiente manera:

- 1) Los valores de la trayectoria observados de cada variable son ordenados de manera ascendente.
- 2) El vector ordenado es dividido en segmentos (tantos segmentos como clases hayan sido determinadas para esa variable) que contienen los valores de proporción determinados por la solución sugerida por el **AG2**.
- 3) Finalmente, los landmarks son escogidos entre los valores extremos de segmentos adyacentes, por ejemplo, usando valores promedio de los datos adyacentes observados en diferentes segmentos.

#### 4.2.2 Población inicial

La población inicial estará compuesta por dos bloques diferentes, se considera un individuo distribuido por el método EFP (Equal Frequency Partition) y el resto de los cromosomas tienen valores escogidos aleatoriamente. No se admiten cromosomas repetidos.

### 4.2.3 Evaluación de los cromosomas

Para evaluar un cromosoma se llevan a cabo los siguientes pasos:

- 1) Se decodifica la información del cromosoma, construyendo su correspondiente partición difusa en la estructura de datos que utiliza FIR para representar dichas particiones.
- 2) Se ejecuta el proceso de identificación de modelado cualitativo de la metodología FIR con los datos de entrenamiento, usando la partición construida en el paso anterior. Por lo tanto, se obtiene la máscara asociada a esa partición difusa con la medida más alta de calidad.
- 3) Se calcula una función objetivo, de adaptación o evaluación para observar la capacidad de generalización del sistema difuso resultante. Se implementan las mismas funciones objetivos que se propusieron en el capítulo anterior: a) la calidad de la máscara óptima y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento. Hay que recordar que la calidad de una máscara  $Q$ , es un valor entre 0 y 1, donde 1 indica la más alta calidad. Por lo tanto, la primera función objetivo propuesta es  $1-Q$ , porque la tarea del algoritmo es minimizar la función objetivo. La segunda función objetivo corresponde a la predicción del error de una porción del conjunto de datos de entrenamiento. El error cuadrático medio normalizado en porcentaje (MSE), mostrado en la ecuación 35, se utiliza para este fin.

### 4.2.4 Operadores genéticos

Para la creación de los operadores genéticos se tomó en cuenta la estructura del cromosoma (con un solo nivel de información) para obtener una definición adecuada que haga el mejor uso posible de la representación elegida. Los operadores genéticos que se han establecido son:

#### 4.2.4.1 Selección

El mecanismo de selección empleado es el mismo descrito en el capítulo anterior, se utiliza el *linear ranking* conjuntamente con el Muestreo Universal Estocástico propuesto por [Bak87], incorporando además una selección elitista.

#### 4.2.4.2 Cruce

El operador de cruce aritmético no uniforme [Mic96] es considerado el más adecuado para la representación del cromosoma que se ha definido. El

funcionamiento de este operador se introdujo en la Sección 2.5.3.

#### 4.2.4.3 Mutación

Debido a la naturaleza de los valores almacenados en el cromosoma, el operador de mutación seleccionado está también basado en uno definido por Thrift en [Thr91]. En este caso, la proporción de datos asociada al gen del cromosoma seleccionado se le aumenta o disminuye (la decisión es tomada de manera aleatoria) por un factor entre el rango  $[V_{MIN}..MAX]$ , también aleatorio. Donde  $MAX = 0.50 - V_{MIN} * (NETIQ - 1)$ . Las otras proporciones de la misma variable son ajustadas para mantener la suma total a 1. Este ajuste es proporcional a los valores de las demás clases de dicha variable. Cuando el valor que se pretende cambiar más el factor sobrepasa los límites del rango  $[V_{MIN}..V_{MAX}]$  se realiza el único cambio posible, esto es, incrementar o decrementar por el factor, respectivamente.

### 4.3 Comentarios

El algoritmo propuesto en este capítulo permite ajustar localmente la definición de las funciones de pertenencia dado un número de clases para cada variable. Este método se puede aplicar a problemas en los que se disponga de información experta sobre la granularidad ó en su defecto pueda hacerse uso del **AG1** para obtenerla.

Actualmente, en la metodología FIR, una vez definida la forma que tendrían las etiquetas (triangulares o gaussianas), la elección usual es una distribución uniforme de las funciones de pertenencia a lo largo del universo de discurso de la variable (Método EFP). Cuando no se dispone de conocimiento experto, esta decisión es lógica. Sin embargo, la incorporación de mecanismos automáticos de ajuste permiten refinar la definición de las funciones de pertenencia, manteniendo siempre invariable la granularidad de las variables (número de etiquetas de las mismas). Se ha señalado que aunque la adaptación no lineal de contextos en sistemas difusos es costosa desde el punto de vista computacional, permite representar mejor conceptos en entornos reales con alta sensibilidad en las clases extremas o en las clases intermedias, por lo que resulta relevante la determinación automática de buenos parámetros de fusificación en la metodología FIR. En el Capítulo VI se presentan varios casos de estudio en donde se emplea el **AG2**. Los resultados obtenidos se comparan con otros métodos desarrollados en esta memoria y con otros obtenidos por diversas metodologías, demostrando su funcionalidad y eficiencia.

## **CAPITULO V:**

---

### **Un Método de Aprendizaje Conjunto de la Granularidad y Funciones de Pertenencia para Razonamiento Inductivo**



## **Capítulo V: Un Método de Aprendizaje Conjunto de la Granularidad y Funciones de Pertenencia para Razonamiento Inductivo**

En los capítulos anteriores se han presentado dos métodos que permiten aprender por separado cada uno de los parámetros de fusificación de la *Metodología de Razonamiento Inductivo Difuso* (FIR), ya que podía suceder que en algunos casos se necesitara aprender solamente el número de clases (etiquetas) por variable ó la definición de las funciones de pertenencia partiendo de un número óptimo de etiquetas. En este capítulo se aborda la propuesta que persigue el aprendizaje conjunto de dichos parámetros permitiendo un ajuste global sobre las particiones difusas.

Este capítulo está conformado por tres secciones. En la primera, se establecen las nociones básicas del nuevo método propuesto para el aprendizaje de ambos parámetros de fusificación al mismo tiempo. Posteriormente se describen sus componentes principales, y finalmente se presentan las ventajas e inconvenientes de la utilización de esta herramienta.

### **5.1 Algoritmo genético para encontrar de manera automática conjuntamente la granularidad y la forma de las funciones de pertenencia: AG3**

Los cromosomas del AG desarrollado codifican la granularidad de las particiones difusas de cada variable y las funciones de pertenencia que definen la semántica de las etiquetas. Para evaluar dichos cromosomas, se estudian dos funciones objetivo distintas: a) la calidad de la máscara óptima y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento.

Los apartados siguientes presentan el método propuesto describiendo los componentes principales del AG.

## 5.2 Aspectos principales

Las funciones de pertenencia de los conjuntos difusos que se han considerado son gaussianas. Los operadores genéticos principales usados son los de cruce y mutación. Se emplea elitismo para no perder la mejor solución encontrada a lo largo de la ejecución. En este caso, el espacio de búsqueda se incrementa notablemente con respecto al **AG1** y al **AG2**. El algoritmo finaliza después de un número estipulado de evaluaciones de cromosomas.

La salida del AG estará constituida tanto por la granularidad como por las funciones de pertenencia (distribución de landmarks) asociadas que hayan obtenido, en conjunto, el mejor valor en la función objetivo utilizada.

### 5.2.1 Codificación de las soluciones

El algoritmo propuesto tiene como objetivo encontrar simultáneamente un número de clases (etiquetas) para cada variable de un sistema y la función de pertenencia más adecuada para cada clase. Por lo tanto, cada cromosoma ( $C$ ) estará compuesto por dos partes:

- Número de clases ( $C_1$ ): La representación genética utilizada es la misma que ha sido descrita en la Sección 3.2.1.
- Funciones de pertenencia ( $C_2$ ): La representación genética utilizada es la misma que ha sido descrita en la Sección 4.2.1.

De este modo, una representación completa del cromosoma es:

$$C = C_1 C_2 \quad (40)$$

Si denotamos por  $E_j$  el número de clases de la variable  $j$ , la representación del número de clases para un sistema de  $N$  variables ( $C_1$ ), es definida por:

$$C_1 = (E_1, E_2, \dots, E_N) \quad (41)$$

También, si denotamos por  $D_{ij}$  la proporción de datos para la clase  $i$  de la variable  $j$ , y por  $C_{2j}$  la información de la proporción de datos para todas las clases de la variable  $j$ ,

$$C_{2j} = (D_{1j}, \dots, D_{E_{jj}}) \quad (42)$$

la representación de las funciones de pertenencia ( $C_2$ ) para un sistema de  $N$  variables (incluyendo entradas y salidas) se define como:

$$C_2 = (C_{21}, C_{22}, \dots, C_{2N}) \quad (43)$$

Lógicamente, cada vez que cambie el número de clases y/o la distribución de los landmarks de una variable por la acción de los operadores genéticos, se tendrá que construir la nueva partición difusa correspondiente.

### 5.2.2 Población inicial

La población inicial está compuesta por cuatro bloques. A excepción del primero el resto posee el mismo número de cromosomas. En ninguno de los casos se admiten cromosomas repetidos. El proceso de generación de la misma se describe a continuación:

- En el primer bloque cada cromosoma tiene el mismo número de etiquetas para todas las variables del problema y se consideran funciones de pertenencia distribuidas uniformemente (Método EFP) a lo largo del universo de discurso de la variable.
- En el segundo bloque, dentro de cada cromosoma puede haber distintos valores (escogidos aleatoriamente) para el número de etiquetas de las variables. Las funciones de pertenencia se distribuyen uniformemente (Método EFP) al igual que en el primer bloque considerando la granularidad correspondiente.
- En el tercer bloque se empieza igual que en el primero, cada cromosoma tiene el mismo número de etiquetas para todas las variables del problema. Las funciones de pertenencia no se distribuyen uniformemente, se generan aleatoriamente.
- En el último bloque se comienza igual que en el segundo, cada variable de cada cromosoma puede tener cualquier número de etiquetas. Las funciones de pertenencia se generan aleatoriamente igual que en el tercer bloque.

La generación de esta forma de la población inicial persigue lograr una diversidad apropiada. Aunque los AGs han probado ser robustos y capaces de obtener buenas soluciones en poblaciones que han sido generadas aleatoriamente (grupo cuatro), se puede obtener una mayor convergencia con la distribución propuesta.

### 5.2.3 Evaluación de los cromosomas

Para evaluar un cromosoma se llevan a cabo los siguientes pasos:

- 1) Se decodifica la información del cromosoma, construyendo su correspondiente partición difusa en la estructura de datos que utiliza FIR para representar dichas particiones.
- 2) Se ejecuta el proceso de identificación de modelado cualitativo de la metodología FIR con los datos de entrenamiento, usando la partición construida en el paso anterior. Por lo tanto, se obtiene la máscara asociada a esa partición difusa con la medida más alta de calidad.
- 3) Se calcula una función objetivo, de adaptación o evaluación para observar la capacidad de generalización del sistema difuso resultante. Se implementan las mismas funciones objetivos que se propusieron en los capítulos anteriores: a) la calidad de la máscara óptima y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento.

### 5.2.4 Operadores genéticos

En relación a los operadores genéticos, se ha considerado la especial estructura de los cromosomas (con dos niveles de información distintos, pero fuertemente relacionados) para obtener una definición adecuada que haga el mejor uso posible de la representación adoptada. Los operadores genéticos considerados se describen a continuación:

#### 5.2.4.1 Selección

Al igual que en los métodos anteriores **AG1** y **AG2**, el mecanismo de selección empleado es el *linear ranking* conjuntamente con el Muestreo Universal Estocástico propuesto por [Bak87], incorporando además una selección elitista.

#### 5.2.4.2 Cruce

Se utilizan dos operadores diferentes según las características de los dos cromosomas padres implicados en el cruce:

- *Cruce cuando los dos padres tienen la misma granularidad en cada una de las variables:* Si los dos cromosomas tienen los mismos valores en  $C_i$  (cada variable tienen el mismo número de etiquetas en ambos padres), se puede suponer que la búsqueda genética ha localizado una zona prometedora del

espacio que conviene explotar. Para ello, se utiliza el operador de cruce aritmético no uniforme [Mic96] sobre  $C_2$ , manteniendo el valor de  $C_1$  en ambos descendientes. El funcionamiento de este operador se estableció en la Sección 2.5.3.

- *Cruce cuando los padres tienen distinto número de etiquetas en alguna de las variables:* En este segundo caso, parece interesante utilizar la información contenida en ambos padres para explorar el espacio de búsqueda y tratar de localizar nuevas zonas prometedoras. El funcionamiento del operador de cruce es sencillo, se selecciona aleatoriamente un punto de corte dentro de  $C_1$ , y se cruzan las dos partes  $C_1$  y  $C_2$  de ambos cromosomas de acuerdo al operador clásico de cruce [Mic96].

A continuación, se muestra una representación gráfica del funcionamiento de este cruce. Sean:

$$\begin{aligned} C_t &= (E_1, \dots, E_p, E_{p+1}, \dots, E_N, C_{21}, \dots, C_{2p}, C_{2p+1}, \dots, C_{2N}) \\ C'_t &= (E'_1, \dots, E'_p, E'_{p+1}, \dots, E'_N, C'_{21}, \dots, C'_{2p}, C'_{2p+1}, \dots, C'_{2N}) \end{aligned} \quad (44)$$

dos individuos que se van a cruzar en el punto  $p$ , los dos descendientes resultantes son:

$$\begin{aligned} C_t &= (E_1, \dots, E_p, E'_{p+1}, \dots, E'_N, C_{21}, \dots, C_{2p}, C'_{2p+1}, \dots, C'_{2N}) \\ C'_t &= (E'_1, \dots, E'_p, E_{p+1}, \dots, E_N, C'_{21}, \dots, C'_{2p}, C_{2p+1}, \dots, C_{2N}) \end{aligned} \quad (45)$$

De esta forma, el proceso completo de recombinación permitirá al AG seguir una tasa adecuada de exploración y explotación del espacio de búsqueda. Nótese que  $C_{2i}$  “viaja” conjuntamente con su granularidad asociada  $E_i$ , asegurando de esta manera que se preservan los requisitos fundamentales de la representación escogida.

#### 5.2.4.3 Mutación

Debido a la naturaleza de los valores almacenados en las dos partes del cromosoma, se utilizarán dos operadores diferentes de mutación, uno para cada una de ellas:

- *Mutación en  $C_1$ :* El operador de mutación seleccionado es el propuesto por [Thr91]. Si el gen a mutar es de  $C_1$ , se aumenta o disminuye en una unidad la granularidad asociada (la decisión es tomada de manera aleatoria). Cuando el valor que se pretende cambiar es el mínimo (2) ó el máximo (9),

se realiza el único cambio posible, esto es, incrementar o decrementar en uno la granularidad, respectivamente. Por supuesto, un cambio en  $C_1$  origina una inconsistencia en  $C_2$ , donde se encuentran almacenados los conjuntos difusos asociados al número de clases (funciones de pertenencia) que tenía antes esa variable. Por tanto, cuando se cambia la granularidad de una variable en  $C_1$  se cambia su parte correspondiente en  $C_2$ , construyendo una partición difusa uniforme (Método EFP) con el nuevo número de clases del cromosoma.

- *Mutación en  $C_2$* : Debido a la naturaleza de los valores almacenados en esta parte del cromosoma, el operador de mutación seleccionado es similar a uno propuesto por Thrift en [Thr91] y ha sido descrito en la Sección 4.2.4.3.

#### 5.2.4.4 Reinicialización

Debido al espacio de búsqueda que manejamos, no es conveniente que la población se quede estancada en una zona determinada. En los primeros experimentos, se detectó que la población convergía demasiado pronto a un nivel de granularidad de las particiones concreto ( $C_1$ ) y no exploraba otras posibles zonas prometedoras. Este proceso se conoce como “deriva genética” [Gol89] y es habitual cuando se trabaja con espacios de búsqueda multimodales. Para evitar este problema, se ha diseñado un operador de reinicialización que se ejecuta cuando la diferencia entre la adaptación (el valor de la función objetivo) del mejor cromosoma y la adaptación media de la población es inferior al 10% del primero.

La forma habitual de proceder en la reinicialización es copiar directamente el mejor individuo y crear el resto de la nueva población a partir de éste, cambiando de forma aleatoria el 70% de los genes y manteniendo los valores del 30% restante [Esh91]. En nuestro caso, como no tiene el mismo efecto modificar un valor de  $C_1$  que de  $C_2$ , se selecciona sobre qué parte se va a actuar (ésta decisión es tomada aleatoriamente). Si es sobre  $C_2$ , se aplicará el esquema anterior, mientras que si es sobre  $C_1$ , se crearán nuevas particiones difusas uniformes en la parte de  $C_2$  que corresponde a las variables alteradas en  $C_1$ , es decir, que han recibido un nuevo número de etiquetas.

### 5.3 Comentarios

El algoritmo propuesto en este capítulo permite ajustar globalmente las particiones difusas a través del aprendizaje simultáneo del nivel de granularidad y la definición de las funciones de pertenencia asociadas a cada uno de los términos

lingüísticos. De esta manera, se tiene la posibilidad de generar mejores sistemas difusos, pero se trabaja con un espacio de búsqueda muy grande, que puede hacer muy difícil y lento el proceso de aprendizaje. Además, el tamaño de la población puede ser más grande al tener que considerar más diversidad para la población inicial.

Sin embargo, la perspectiva de optimizar los dos parámetros de manera conjunta debería facilitar la obtención de soluciones más precisas y de mayor interés para aquellas aplicaciones donde la precisión sea el aspecto más importante a tener en cuenta. Cabe señalar, también, que la determinación de la partición difusa más adecuada por medio del **AG3** se realiza como un proceso fuera de línea. Es decir, una vez determinados mediante el **AG3** los parámetros idóneos del proceso de discretización de FIR y una vez identificado el mejor modelo que representa el sistema, el proceso de predicción de éste puede realizarse en tiempo real.

En el próximo capítulo se presentan varios casos de estudio en donde se aplica el **AG3** desarrollado. Los resultados obtenidos se comparan tanto con los de los métodos **AG1** y **AG1 + AG2** como con otros obtenidos por diferentes metodologías en trabajos anteriores, observándose también su funcionalidad y rendimiento.



## **CAPITULO VI:**

---

### **Resultados Experimentales**



## Capítulo VI: Resultados Experimentales

Este capítulo está dedicado a estudiar el comportamiento de los distintos métodos evolutivos de aprendizaje desarrollados para la *Metodología de Razonamiento Inductivo Difuso* (FIR). Puesto que el fin de los paradigmas evolutivos de aprendizaje propuestos es poder ser aplicados en problemas reales de cualquier tipo en función de las necesidades de modelado que se presenten, las aplicaciones afrontadas en breve nos van a permitir aportar una visión más clara de las posibilidades de la metodología evolutiva de aprendizaje presentada en esta memoria.

Se han escogido tres problemas en donde el requerimiento principal del modelado será la precisión del modelo obtenido. En primer lugar, se presenta un problema sencillo (desde el punto de vista del número de variables involucradas) en el que el objetivo será obtener un modelo para cada uno de los controladores que conforman el Sistema Nervioso Central (SNC) Humano [NCV98].

La segunda aplicación seleccionada corresponde a un problema de estimación de los costos óptimos de mantenimiento de la línea de media tensión en núcleos urbanos de la provincia de Asturias [CHS98, CHS99].

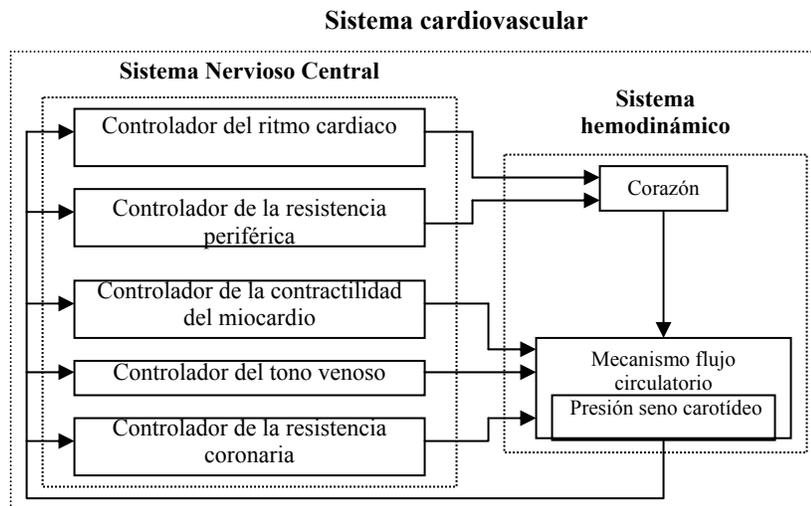
La última aplicación considerada se trata de un problema de predicción de las concentraciones locales máximas de ozono en la región oriental Austríaca, específicamente en Viena [WW99, GNRA+03].

La estructura de este capítulo incluye tres secciones dedicadas, respectivamente, a cada una de las aplicaciones comentadas. Cada sección está dividida en cuatro partes principales, una en la que se introduce el problema en cuestión, otra en donde se describen los trabajos previos llevados a cabo para solucionarlo, la tercera donde se muestra su resolución por medio de los métodos genéticos propuestos en esta memoria (**AG1**, **AG2** y **AG3**) y una última sección en la que se comparan los resultados con los obtenidos en estudios previos y se presentan algunas conclusiones.

## 6.1 Caso de estudio: Sistema nervioso central humano

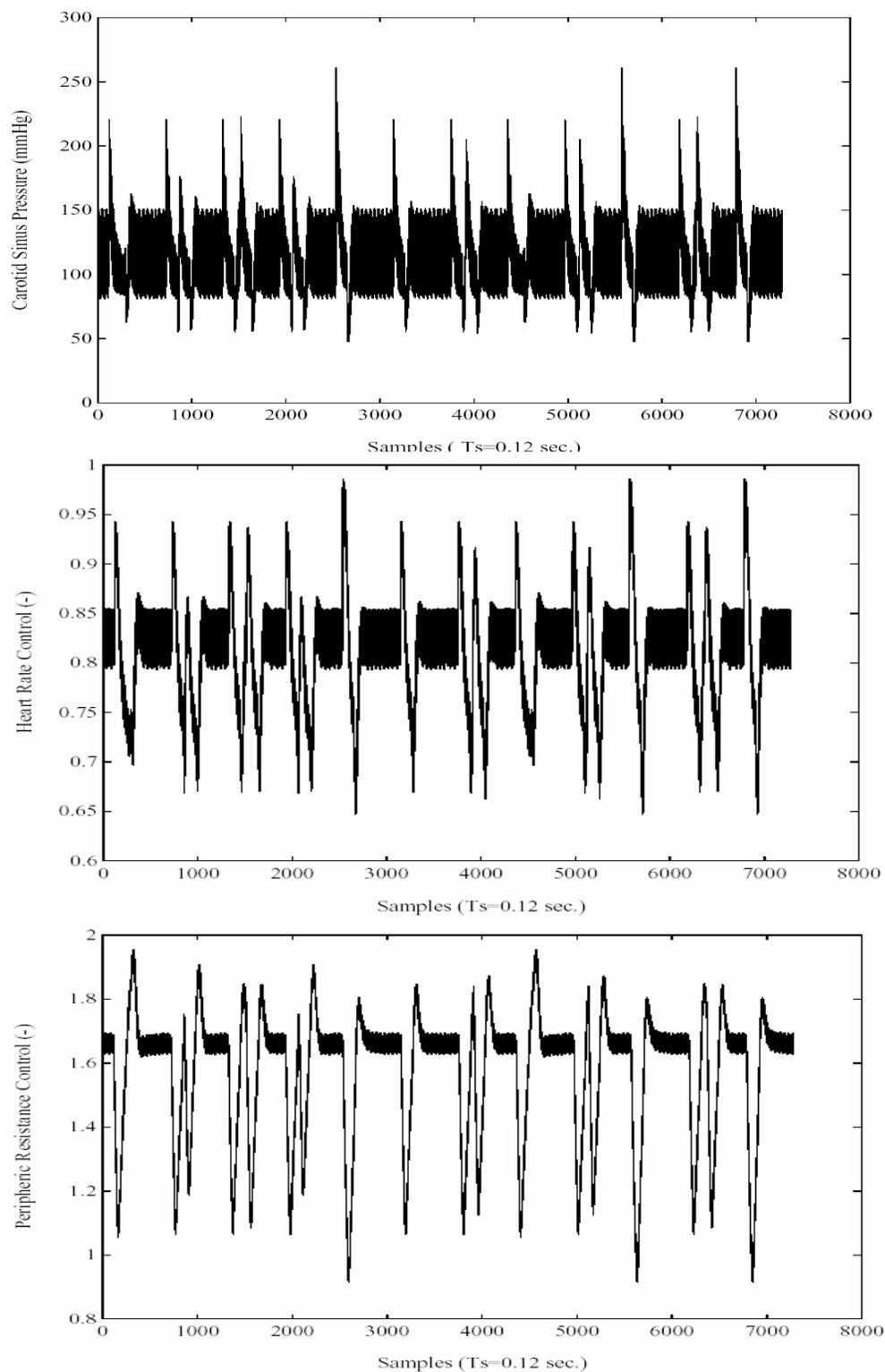
### 6.1.1 Descripción del problema

Desde el punto de vista médico, el estudio y análisis del sistema cardiovascular humano (SCH) por medio de metodologías de modelado y simulación es de gran relevancia porque permite a los doctores adquirir un mejor conocimiento de la fisiología cardiovascular, pudiendo ofrecer un diagnóstico más preciso y seleccionar la terapia más adecuada. El SCH está conformado por el sistema hemodinámico (SH) y el Sistema Nervioso Central (SNC). El SH opera esencialmente como una bomba hidro-mecánica, y por lo tanto, se conoce bien su funcionamiento y estructura. Por este motivo, se pueden encontrar en la literatura un gran número de modelos cuantitativos que representan este sistema de forma precisa [SMSS88, Val93]. El SNC genera las señales que son transmitidas desde el cerebro al corazón y a los vasos sanguíneos, controlando al SH. Estas señales de regulación actúan sobre el corazón y el lecho vascular para mantener un flujo de sangre apropiado siguiendo los cambios de requerimientos del SH en cada período de tiempo. Por consiguiente, el funcionamiento del SNC presenta una gran complejidad y no se tienen aún buenos modelos deductivos capaces de describirlo desde principios físicos y fisiológicos.

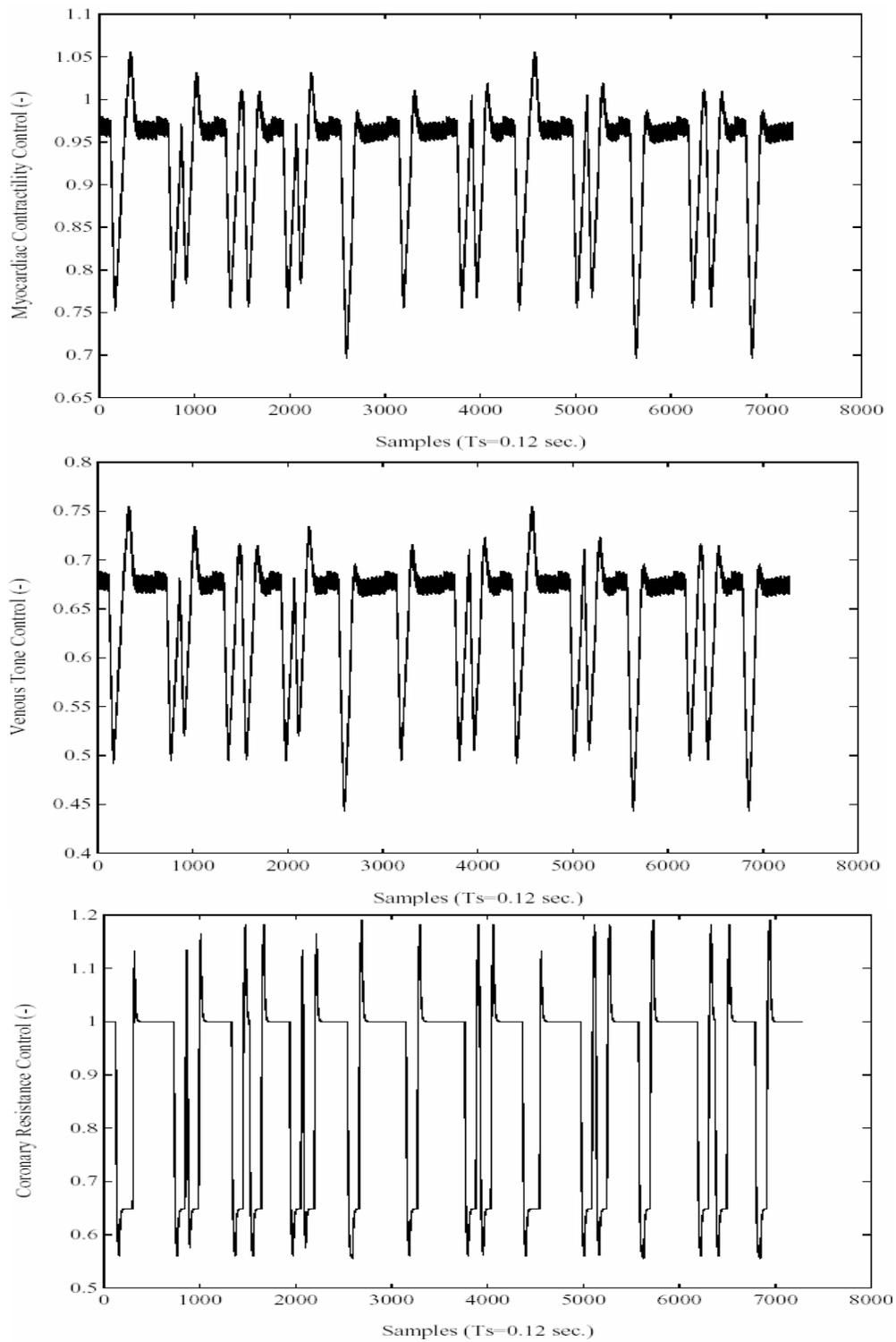


**Figura 13.** Diagrama simplificado del modelo del sistema cardiovascular, compuesto del sistema hemodinámico (SH) y del Sistema Nervioso Central (SNC).

La figura 13 muestra un diagrama simplificado del modelo del SCH compuesto del SH y del SNC. El SCH es un complejo sistema dinámico en el cual varias acciones deben ser consideradas: llenado y vaciado del ventrículo, flujo sanguíneo a través de todo el organismo, las acciones de control hechas por el SNC, y la interacción del sistema cardiovascular con otros sistemas del cuerpo.



**Figura 14.** Señales de entrenamiento de la entrada PSC y los controladores RCA y RP.



**Figura 15.** Señales de entrenamiento de los controladores CM, TV y RC.

La principal tarea del SH es asegurar el flujo continuo de sangre en el cuerpo humano para llevar el oxígeno y las sustancias metabólicas necesarias a los tejidos, así como también eliminar los productos de la oxidación. Por su parte, como ya se ha señalado el SNC tiene como función principal controlar al SH y esto lo logra mediante la generación de señales de regulación para los vasos sanguíneos y el corazón. Estas señales son transmitidas a través de los nervios simpáticos y parasimpáticos, produciendo los estímulos en los órganos correspondientes y otras partes del cuerpo.

Se ha indicado antes que el SNC está compuesto por cinco controladores diferentes que generan las señales de control del ritmo cardíaco (RCA), resistencia periférica (RP), contractilidad del miocardio (CM), tono venoso (TV) y resistencia coronaria (RC). Todos los controladores del SNC son modelos SISO<sup>15</sup> manejados por la misma variable de entrada, la presión del seno carotídeo (PSC). Adicionalmente, se especifica que la identificación del modelo se realiza para los cinco controladores de un paciente. Las señales de entrada y salida de los controladores del SNC fueron registradas con una tasa de muestreo de 0.12 segundos [NCV98], obteniendo un registro de 7279 datos para identificar cada uno de los modelos. El modelo ha sido ajustado para representar una patología específica de un paciente con una obstrucción arterial coronaria, haciendo que las cuatro variables fisiológicas (presión de la aurícula derecha, presión aórtica, flujo de sangre coronaria y ritmo cardíaco) del modelo coincidan con la medición de datos tomados del paciente real. Las *figuras 14* y *15* muestran las señales de entrenamiento de la variable de entrada y las variables de salida correspondientes a cada controlador para un paciente. Los modelos FIR obtenidos se validan mediante seis conjuntos de prueba que no han sido utilizados en el proceso de identificación. Cada conjunto de prueba tiene un tamaño de alrededor de 600 valores cada uno y representa una morfología específica permitiendo, de esta manera, la validación de cada modelo para comportamientos diferentes del sistema.

### 6.1.2 Trabajos previos

La *Tabla 3* contiene las predicciones alcanzadas para el mismo problema usando modelos NARMAX [NCV98], redes neuronales *time delay* (TDNN) [CAN97] y redes neuronales recurrentes (RNN) [CAN97]. Las columnas de la tabla especifican el promedio del error de predicción de los 6 conjuntos de prueba para cada controlador. Todas las metodologías usan el mismo conjunto de datos de entrenamiento y de prueba descrito previamente.

---

<sup>15</sup> Single input single output (siglas en inglés para nombrar los modelos de una única entrada y una única salida).

**Tabla 3.** Errores de predicción MSE de los modelos del SNC usando metodologías NARMAX, TDNN y RNN (valor medio de los 6 conjuntos de datos de prueba para cada controlador).

	RCA	RP	CM	TV	RC
NARMAX	9.3%	18.5%	22.0%	22.0%	25.5%
TDNN	15.3%	33.7%	34.0%	34.0%	55.6%
RNN	18.3%	31.1%	35.1%	34.7%	57.1%

### 6.1.3 Modelos obtenidos por los métodos *AG1*, *AG2* y *AG3*

Se han realizado 30 ejecuciones (con distinto valor de la semilla) para cada función objetivo y criterio de parada en cada uno de los métodos genéticos desarrollados. Se estudian dos funciones objetivo diferentes: a) 1-Q, y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). En esta aplicación el último 25% de la señal de entrenamiento se usa para la evaluación de la función objetivo  $MSE_{entr}$  y solamente el 75% inicial de la señal se usa para la obtención de los modelos FIR (máscaras y bases de reglas patrón).

En la ecuación 46 se muestra la máscara candidata de profundidad 3 (tres filas) utilizada para cada controlador, la cual permite encontrar las relaciones causales y temporales entre las variables de entrada y salida. Los elementos -1 de la máscara candidata representan las relaciones causales y temporales potenciales con la variable de salida (elemento +1). Por ejemplo para el controlador RC sería:

$t \setminus x$	PSC	RC
$t - 2\delta t$	-1	-1
$t - \delta t$	-1	-1
$t$	-1	+1

(46)

Por lo tanto, cuando se muestren las máscaras óptimas obtenidas en notación de posición será necesario tener en cuenta la estructura de la máscara candidata mostrada en la ecuación 46. De manera que las posiciones impares (1, 3 y 5) corresponderán a valores de la variable de entrada, PSC, en los diferentes intervalos de tiempo y las posiciones pares (2, 4 y 6) corresponderán a valores de la variable de salida, RC. Recordemos que en este tipo de notación las posiciones de la máscara se enumeran de arriba abajo y de izquierda a derecha.

Los valores escogidos para los parámetros de los AGs se muestran en la *Tabla 4*. Los valores de la probabilidad de cruce y de mutación han sido definidos de

acuerdo a lo señalado por [Gre86]. Se utilizó un computador Pentium III (0.6 GHz) para realizar todas las simulaciones. En [ANF04] se encuentran depositados todos los resultados obtenidos para los métodos **AG1** y **AG2**, en cambio para el **AG3** se exponen en [ANF05a]. Los resultados obtenidos (considerando los valores 1-Q y  $MSE_{entr}$ ) por los métodos propuestos para el problema de modelado del sistema nervioso central (SNC) humano se muestran en las *Tablas 5 a 11*.

**Tabla 4.** Parámetros de los AGs para el problema de modelado del sistema nervioso central (SNC) humano.

Parámetro	Tipo AG	Valor
<i>Tamaño de la población</i> (# individuos)	AG1	9
	AG2, AG3	50
<i>Probabilidad de cruce</i>	Todos	0.6
<i>Probabilidad de mutación</i>	Todos	0.1
<i>Criterio de parada</i> (# evaluaciones de cromosomas)	AG1	{40, 64}
	AG2	{3500, 7000}
	AG3	{7000, 14000}

Los resultados son presentados en tres sub-apartados denominados en base al método genético empleado para la resolución del problema bajo estudio.

#### 6.1.3.1 Aprendizaje del número óptimo de clases para cada variable del sistema (AG1)

En esta sección se muestran los resultados obtenidos por el **AG1** utilizado como una herramienta de pre-procesamiento de la metodología FIR en la identificación de buenos modelos para cada uno de los cinco controladores del SNC. La Tabla 5 muestra los mejores resultados de granularidad para el controlador de resistencia coronaria (RC) cuando se usan ambas funciones objetivo (primera columna). Un \* significa que la partición es la mejor posible, y por lo tanto, es la solución óptima. La tabla ha sido organizada de la siguiente manera:

- # *Eval.* : indica el número de evaluaciones de cromosomas hechas por el AG.
- *Gran.* : señala la granularidad sugerida por el AG cuando finaliza su ejecución. Nótese que aquí la granularidad sugerida es el parámetro de entrada al proceso de fusificación de la metodología FIR.
- *Opt. Mask y Q*: contienen la máscara óptima (en notación de posición) obtenida por FIR para esta granularidad específica y la calidad asociada a esta máscara, respectivamente.
- *1-Q*: corresponde al valor de la función objetivo de la calidad.
- $MSE_{entr}$ : corresponde al valor de la función objetivo del error de predicción

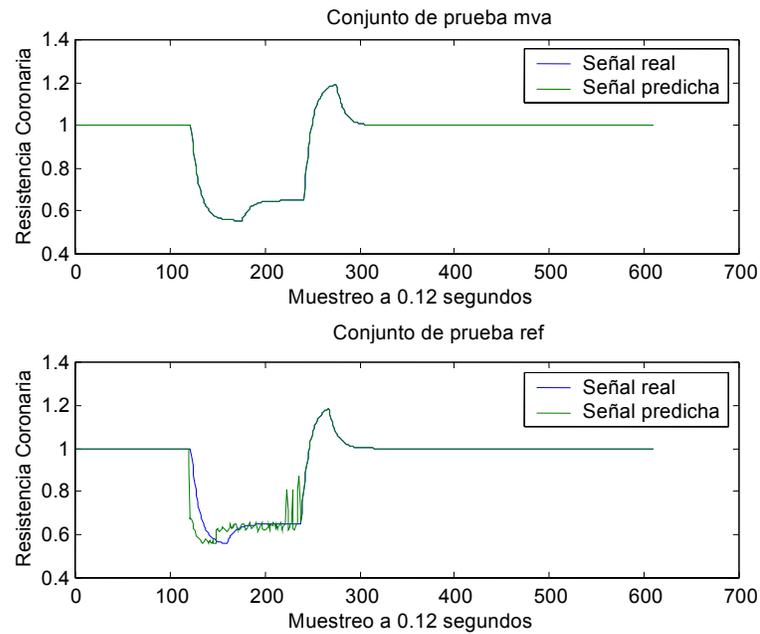
MSE del último 25% de los datos del conjunto de entrenamiento.

- $MSE_{prue}$ : indica el valor medio del error de predicción (MSE) obtenido para los 6 conjuntos de datos de prueba (Véase ecuación 35).
- $Tiempo$ : contiene el tiempo promedio de CPU (en segundos), el cual se calcula tomando en cuenta el menor y el mayor tiempo consumido por el AG para encontrar la solución sugerida.

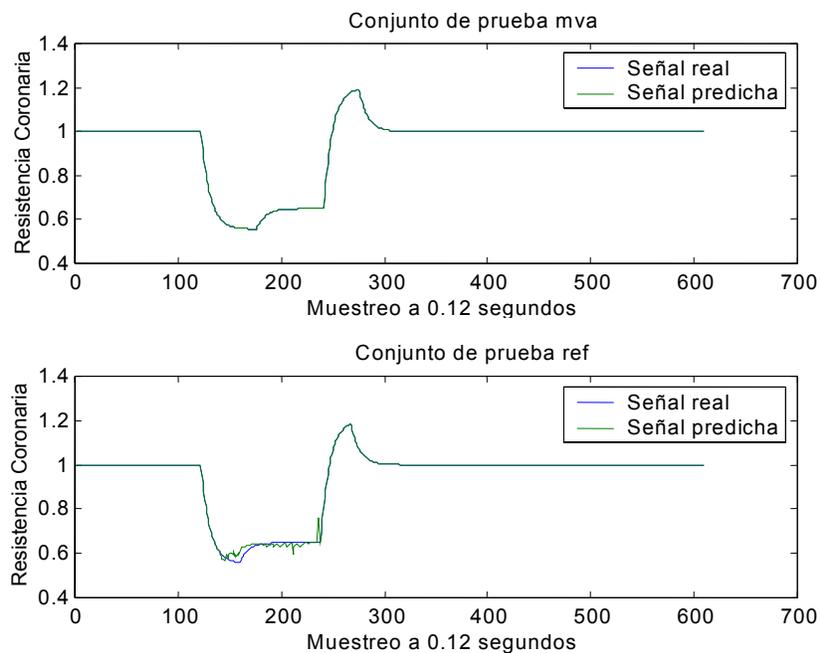
**Tabla 5.** Resultados de la granularidad para el controlador RC usando las funciones objetivo: a) 1-Q y, b) Predicción del 25% de la señal de entrenamiento ( $MSE_{entr}$ ). Método EFP para las funciones de pertenencia. (\* Solución óptima).

1-Q	# eval	Gran.	Opt. Mask	Q	1-Q	$MSE_{prue}$	Tiempo
	64	(9,3)*	(1,4,6)	0.9787	0.0213	3.92%	18
	64	(8,3)	(1,4,6)	0.9776	0.0224	4.33%	14
	64	(7,3)	(1,4,6)	0.9776	0.0224	4.85%	13
	64	(4,3)	(1,4,6)	0.9748	0.0252	1.36%	10
	40	(9,3)*	(1,4,6)	0.9787	0.0213	3.92%	12
	40	(8,3)	(1,4,6)	0.9776	0.0224	4.33%	9
	40	(7,3)	(1,4,6)	0.9776	0.0224	4.85%	9
	40	(6,3)	(1,4,6)	0.9762	0.0238	1.79%	7
	40	(5,3)	(1,4,6)	0.9749	0.0251	2.38%	9
Solución óptima: Partición= (9,3); Q=0.9787							
$MSE_{entr}$	# eval	Gran.	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$	Tiempo
	64	(2,5)*	(1,4,5,6)	0.9642	0.08%	0.16%	114
	64	(3,4)	(4,5,6)	0.9638	0.13%	0.29%	111
	64	(2,4)	(3,4,5,6)	0.9630	0.17%	0.39%	93
	64	(7,4)	(4,5,6)	0.9677	0.19%	0.42%	84
	40	(2,5)*	(1,4,5,6)	0.9642	0.08%	0.16%	49
	40	(3,4)	(4,5,6)	0.9638	0.13%	0.29%	55
	40	(2,4)	(3,4,5,6)	0.9630	0.17%	0.39%	57
	40	(7,4)	(4,5,6)	0.9677	0.19%	0.42%	54
Solución óptima: Partición= (2,5); $MSE_{entr}$ =0.08%							

Las figuras 16 y 17 muestran el seguimiento de las mejores y peores señales de predicción de los conjuntos de prueba de la solución óptima sugerida por el AG1 para el controlador RC en cada función objetivo, respectivamente.



**Figura 16.** Mejor y peor predicción usando los modelos FIR para el controlador RC con granularidad (9,3). Función objetivo  $1-Q$  (errores  $mva=1.614 \times 10^{-26}\%$   $ref=6.66\%$ ). Método AGI.



**Figura 17.** Mejor y peor predicción usando los modelos FIR para el controlador RC con granularidad (2,5). Función objetivo  $MSE_{entr}$  (errores  $mva=6.8793 \times 10^{-29}\%$   $ref=0.34\%$ ). Método AGI.

Las *Tablas 6 y 7* contienen los mejores resultados de granularidad para el resto de los controladores del SNC. La primera columna de las tablas indica el tipo de función objetivo utilizada, es decir, 1-Q ó el error de predicción del 25% final del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). La segunda columna se refiere al criterio de parada usado, es decir, el número de evaluaciones de cromosomas realizadas por el AG. La tercera columna muestra la granularidad sugerida para cada una de las variables del sistema bajo estudio. La cuarta columna indica el valor de la función objetivo (1-Q ó  $MSE_{entr}$ ). La última columna representa el error MSE promedio de los 6 conjuntos de datos de prueba para cada controlador.

**Tabla 6.** Resultados de la granularidad para el controlador TV y RP usando las funciones objetivo: a) 1-Q y, b) Predicción del 25% de la señal de entrenamiento ( $MSE_{entr}$ ). Método EFP para las funciones de pertenencia. (\* Solución óptima).

		TV			
1-Q	# eval	Gran.	1-Q	$MSE_{prue}$	
	64	(8,7)*	0.1858	9.49%	
	64	(7,7)	0.1952	40.20%	
	40	(8,7)*	0.1858	9.49%	
	40	(7,7)	0.1952	40.20%	
$MSE_{entr}$	# eval	Gran.	$MSE_{entr}$	$MSE_{prue}$	
	64	(2,5)*	0.61%	1.69%	
	64	(2,8)	0.64%	1.58%	
	64	(2,7)	0.77%	2.09%	
	64	(4,9)	0.91%	2.06%	
	40	(2,5)*	0.61%	1.69%	
	40	(2,8)	0.64%	1.58%	
	40	(2,7)	0.77%	2.09%	
	40	(7,7)	1.34%	2.08%	
		RP			
1-Q	# eval	Gran.	1-Q	$MSE_{prue}$	
	64	(8,7)*	0.1448	6.10%	
	64	(7,7)	0.1505	4.66%	
	40	(8,7)*	0.1448	6.10%	
	40	(7,7)	0.1505	4.66%	
$MSE_{entr}$	# eval	Gran.	$MSE_{entr}$	$MSE_{prue}$	
	64	(4,9)*	0.93%	3.05%	
	64	(7,7)	1.08%	3.40%	
	40	(4,9)*	0.93%	3.05%	
	40	(7,7)	1.08%	3.40%	

**Tabla 7.** Resultados de la granularidad para el controlador CM y RCA usando las funciones objetivo: a) 1-Q y, b) Predicción del 25% de la señal de entrenamiento ( $MSE_{entr}$ ). Método EFP para las funciones de pertenencia. (\* Solución óptima).

		CM			
1-Q	# eval	Gran.	1-Q	$MSE_{prue}$	
	64	(8,7)*	0.1866	8.37%	
	64	(7,7)	0.1950	40.77%	
	40	(8,7)*	0.1866	8.37%	
	40	(7,7)	0.1950	40.77%	
$MSE_{entr}$	# eval	Gran.	$MSE_{entr}$	$MSE_{prue}$	
	64	(4,9)*	0.60%	2.53%	
	64	(2,5)	0.63%	2.76%	
	64	(3,9)	1.10%	3.91%	
	40	(4,9)*	0.60%	2.53%	
	40	(2,5)	0.63%	2.76%	
	40	(3,9)	1.10%	3.91%	
	40	(2,6)	1.13%	3.04%	

		RCA			
1-Q	# eval	Gran.	1-Q	$MSE_{prue}$	
	64	(7,2)*	0.1674	13.76%	
	64	(8,2)	0.1861	12.94%	
	64	(6,2)	0.1968	15.99%	
	64	(7,4)	0.2739	2.68%	
	64	(9,4)	0.2756	13.28%	
	64	(8,7)	0.2774	12.39%	
	40	(7,2)*	0.1674	13.76%	
	40	(8,2)	0.1861	12.94%	
	40	(6,2)	0.1968	15.99%	
	40	(9,2)	0.1973	12.24%	
	40	(7,4)	0.2739	2.68%	
	40	(8,4)	0.2763	2.91%	
	40	(8,7)	0.2774	12.39%	
	40	(7,8)	0.2804	5.69%	
40	(9,7)	0.2811	3.59%		
$MSE_{entr}$	# eval	Gran.	$MSE_{entr}$	$MSE_{prue}$	
	64	(3,7)*	0.89%	9.37%	
	64	(5,9)	1.02%	2.85%	
	64	(4,6)	1.03%	3.48%	
	64	(4,7)	1.34%	3.82%	
	40	(3,7)*	0.89%	9.37%	
	40	(5,9)	1.02%	2.85%	
	40	(4,6)	1.03%	3.48%	
	40	(6,7)	1.15%	13.73%	
40	(4,4)	1.65%	2.79%		

### 6.1.3.2 Aprendizaje de las funciones de pertenencia (AG2)

Este apartado presenta los resultados obtenidos por el algoritmo genético propuesto para la optimización de las funciones de pertenencia (**AG2**). Este AG considera el número óptimo de clases obtenido por el **AG1** mostrado previamente, por lo tanto, se tiene la combinación **AG1+AG2**. En este caso, el espacio de búsqueda se incrementa considerablemente.

La *Tabla 8* presenta los resultados obtenidos por el **AG2** para los cinco controladores del SNC cuando se usa la función objetivo 1-Q. La tabla se organiza como se describe a continuación. La primera columna se divide en dos secciones. La Sección A corresponde al mejor resultado obtenido por el **AG2** mientras que la Sección B corresponde a su peor resultado. La segunda columna indica el número de evaluaciones de cromosomas realizadas por el **AG2**. La tercera columna muestra la proporción de datos para la variable de entrada (PSC) y las variables de salida de cada controlador (RC,TV,RP,CM,RCA). El número de elementos de la proporción de datos va de acuerdo a la granularidad de cada variable (mostrada en la parte superior de cada controlador). La proporción de datos es la salida del **AG2**. La cuarta columna presenta la máscara óptima (en notación de posición) encontrada por FIR cuando se emplea esa proporción de datos para el cómputo de los landmarks. La quinta columna corresponde a la calidad asociada a la máscara óptima. La sexta columna es el valor de la función objetivo 1-Q. La última columna indica el valor medio del error de predicción (MSE) obtenido para los 6 conjuntos de datos de prueba.

El tiempo de CPU necesario por el **AG2** para efectuar las 30 ejecuciones está directamente relacionado con el número predefinido de clases por variable. Para los controladores CM, TV y RP que tenían una partición con un alto número de clases (8,7) en la función objetivo 1-Q, el tiempo requerido por el **AG2** para realizar una ejecución fue de 1:45 horas.

Por su parte, la *Tabla 9* muestra los resultados obtenidos por este AG para todos los controladores del SNC usando la función objetivo  $MSE_{entr}$ . La tabla está organizada como la *Tabla 8*. La única diferencia es que la sexta columna contiene los valores de la función objetivo  $MSE_{entr}$  en vez de 1-Q. En este caso, el tiempo promedio de CPU necesario por el **AG2** para realizar una ejecución para cada controlador fue de 5 horas.

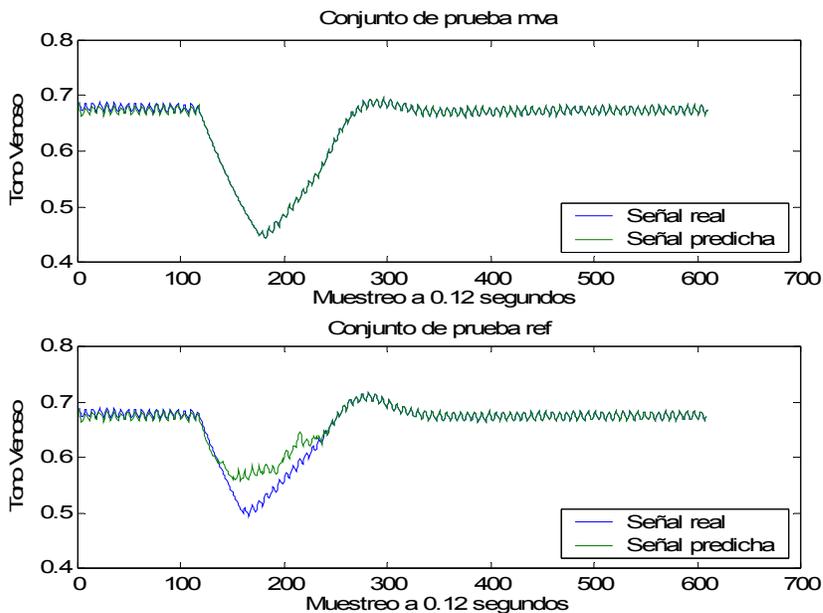
Las *figuras 18 y 19* muestran las mejores y peores señales de predicción de los conjuntos de prueba de la solución óptima sugerida por el **AG2** para el controlador TV en cada función objetivo, respectivamente.

**Tabla 8.** Resultados de las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1).

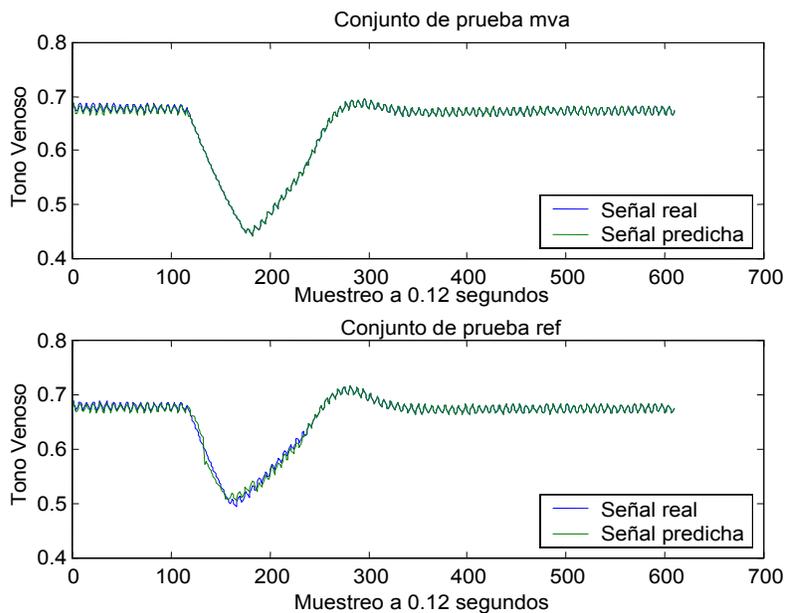
<b>Controlador RC; Núm. Clases: (9,3)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A 7000	PSC:(0.06,0.09,0.18,0.12,0.09,0.22,0.11,0.08,0.05) RC:(0.38,0.43,0.19)	(1,4,6)	0.9845	0.0155	2.03%
3500	PSC:(0.06,0.07,0.10,0.19,0.11,0.23,0.07,0.12,0.05) RC:(0.40,0.41,0.19)	(1,4,6)	0.9844	0.0156	2.03%
B 7000	PSC:(0.07,0.14,0.17,0.17,0.18,0.10,0.07,0.05,0.05) RC:(0.38,0.43,0.19)	(1,4,6)	0.9840	0.0160	1.36%
3500	PSC:(0.11,0.20,0.09,0.25,0.12,0.07,0.06,0.05,0.05) RC:(0.38,0.39,0.23)	(1,4,6)	0.9836	0.0164	1.56%
<b>Controlador TV; Núm. Clases: (8,7)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A 7000	PSC:(0.21,0.17,0.19,0.07,0.13,0.07,0.11,0.05) TV:(0.06,0.08,0.05,0.05,0.06,0.62,0.08)	(3,4,6)	0.9332	0.0668	6.17%
3500	PSC:(0.17,0.26,0.11,0.14,0.11,0.07,0.08,0.06) TV:(0.10,0.07,0.07,0.06,0.60,0.05,0.05)	(3,4,6)	0.9256	0.0744	73.77%
B 7000	PSC:(0.14,0.11,0.14,0.11,0.11,0.19,0.10,0.10) TV:(0.17,0.12,0.12,0.14,0.15,0.20,0.10)	(3,4,6)	0.8401	0.1599	4.87%
3500	PSC:(0.13,0.13,0.12,0.11,0.09,0.15,0.11,0.16) TV:(0.13,0.18,0.13,0.16,0.15,0.15,0.10)	(3,4,6)	0.8308	0.1692	9.82%
<b>Controlador RP; Núm. Clases: (8,7)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A 7000	PSC:(0.13,0.13,0.11,0.09,0.15,0.12,0.14,0.13) RP:(0.18,0.11,0.15,0.14,0.15,0.11,0.16)	(4,5,6)	0.8831	0.1169	19.56%
3500	PSC:(0.13,0.14,0.10,0.10,0.12,0.13,0.15,0.13) RP:(0.18,0.11,0.15,0.14,0.14,0.12,0.16)	(4,5,6)	0.8807	0.1194	3.61%
B 7000	PSC:(0.19,0.15,0.13,0.12,0.09,0.12,0.13,0.07) RP:(0.13,0.16,0.20,0.17,0.18,0.10,0.06)	(3,4,6)	0.8687	0.1313	8.21%
3500	PSC:(0.12,0.13,0.11,0.15,0.10,0.21,0.09,0.09) RP:(0.20,0.09,0.29,0.15,0.11,0.09,0.07)	(4,5,6)	0.8671	0.1329	3.66%
<b>Controlador CM; Núm. Clases: (8,7)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A 7000	PSC:(0.17,0.19,0.20,0.08,0.11,0.07,0.07,0.11) CM:(0.06,0.05,0.06,0.06,0.05,0.67,0.05)	(4,5,6)	0.9371	0.0629	3.63%
3500	PSC:(0.18,0.21,0.10,0.08,0.10,0.10,0.15,0.08) CM:(0.07,0.05,0.05,0.05,0.08,0.65,0.05)	(3,4,6)	0.9334	0.0666	82.52%
B 7000	PSC:(0.15,0.15,0.13,0.08,0.09,0.14,0.13,0.13) CM:(0.17,0.15,0.12,0.15,0.15,0.16,0.10)	(3,4,6)	0.8347	0.1653	4.03%
3500	PSC:(0.14,0.10,0.14,0.10,0.08,0.16,0.19,0.09) CM:(0.17,0.14,0.17,0.15,0.16,0.11,0.10)	(3,4,6)	0.8345	0.1655	2.53%
<b>Controlador RCA; Núm. Clases: (7,2)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A 7000	PSC:(0.05,0.20,0.27,0.08,0.09,0.05,0.26) RCA:(0.06,0.94)	(4,5,6)	0.9421	0.0579	11.49%
3500	PSC:(0.05,0.05,0.05,0.36,0.19,0.05,0.25) RCA:(0.06,0.94)	(4,5,6)	0.9420	0.0580	11.98%
B 7000	PSC:(0.15,0.14,0.11,0.14,0.10,0.14,0.22) RCA:(0.07,0.93)	(4,5,6)	0.9212	0.0788	12.48%
3500	PSC:(0.20,0.19,0.13,0.10,0.07,0.09,0.22) RCA:(0.06,0.94)	(4,5,6)	0.9192	0.0808	11.49%

**Tabla 9.** Resultados de las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1).

<b>Controlador RC; Núm. Clases: (2,5)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$
A 7000	PSC:(0.54,0.46) RC:(0.13,0.16,0.55,0.11,0.05)	(1,4,5,6)	0.9488	0.05%	0.11%
3500	PSC:(0.49,0.51) RC:(0.12,0.17,0.54,0.09,0.08)	(1,4,5,6)	0.9533	0.05%	0.11%
B 7000	PSC:(0.53,0.47) RC:(0.20,0.18,0.20,0.29,0.13)	(1,4,5,6)	0.9650	0.07%	0.13%
3500	PSC:(0.53,0.47) RC:(0.20,0.23,0.22,0.18,0.17)	(1,4,5,6)	0.9646	0.07%	0.13%
<b>Controlador TV; Núm. Clases: (2,5)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$
A 7000	PSC:(0.38,0.62) TV:(0.11,0.14,0.31,0.22,0.22)	(1,4,5,6)	0.7475	0.19%	0.46%
3500	PSC:(0.39,0.61) TV:(0.11,0.12,0.25,0.30,0.22)	(1,4,5,6)	0.7400	0.20%	0.47%
B 7000	PSC:(0.41,0.59) TV:(0.14,0.10,0.20,0.22,0.34)	(3,4,5,6)	0.7567	0.29%	0.76%
3500	PSC:(0.43,0.57) TV:(0.13,0.24,0.46,0.09,0.08)	(1,4,5,6)	0.7825	0.33%	0.81%
<b>Controlador RP; Núm. Clases: (4,9)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$
A 7000	CSP:(0.23,0.20,0.23,0.34) PR:(0.09,0.10,0.08,0.13,0.14,0.08,0.13,0.10,0.15)	(4,5,6)	0.8064	0.20%	0.80%
3500	CSP:(0.24,0.22,0.22,0.32) PR:(0.09,0.09,0.10,0.11,0.14,0.13,0.14,0.10,0.10)	(3,4,6)	0.7995	0.22%	0.72%
B 7000	CSP:(0.21,0.15,0.50,0.14) PR:(0.14,0.13,0.12,0.13,0.09,0.10,0.07,0.07,0.15)	(4,5,6)	0.7983	0.41%	3.88%
3500	CSP:(0.27,0.16,0.28,0.29) PR:(0.14,0.11,0.14,0.09,0.08,0.10,0.10,0.15,0.09)	(4,5,6)	0.7758	0.41%	1.06%
<b>Controlador CM; Núm. Clases: (4,9)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$
A 7000	PSC:(0.20,0.29,0.31,0.20) CM:(0.09,0.08,0.12,0.13,0.12,0.11,0.11,0.11,0.13)	(4,5,6)	0.7452	0.20%	0.45%
3500	P:(0.27,0.18,0.32,0.23) CM:(0.09,0.08,0.11,0.11,0.15,0.11,0.15,0.10,0.10)	(4,5,6)	0.7899	0.21%	0.48%
B 7000	PSC:(0.22,0.16,0.35,0.27) CM:(0.16,0.12,0.11,0.13,0.10,0.10,0.10,0.10,0.08)	(4,5,6)	0.7616	0.33%	2.98%
3500	PSC:(0.13,0.18,0.40,0.29) CM:(0.16,0.10,0.13,0.09,0.10,0.09,0.15,0.10,0.08)	(4,5,6)	0.7507	0.37%	1.59%
<b>Controlador RCA; Núm. Clases: (3,7)</b>					
# eval	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$
A 7000	PSC:(0.32,0.30,0.38) RCA:(0.13,0.10,0.12,0.12,0.29,0.12,0.12)	(4,5,6)	0.6837	0.44%	1.49%
3500	PSC:(0.30,0.32,0.38) RCA:(0.13,0.14,0.17,0.16,0.13,0.14,0.13)	(4,5,6)	0.6517	0.44%	1.52%
B 7000	PSC:(0.32,0.30,0.38) RCA:(0.19,0.12,0.16,0.14,0.12,0.12,0.15)	(4,5,6)	0.6711	0.60%	2.30%
3500	PSC:(0.41,0.23,0.36) RCA:(0.16,0.09,0.21,0.20,0.14,0.10,0.10)	(4,5,6)	0.6998	0.69%	2.73%



**Figura 18.** Mejor y peor predicción usando los modelos FIR para el controlador VT con granularidad (8,7) y distribución de landmarks CSP:(0.21,0.17,0.19,0.07,0.13,0.07,0.11,0.05)- VT:(0.06,0.08,0.05,0.05,0.06,0.62,0.08). Función objetivo  $I-Q$  (errores mva=0.11% ref=12.46%). Método AG2.



**Figura 19.** Mejor y peor predicción usando los modelos FIR para el controlador VT con granularidad (2,5) y distribución de landmarks CSP:(0.38,0.62)- VT:(0.11,0.14,0.31,0.22,0.22). Función objetivo  $MSE_{entr}$  (errores mva=0.05% ref=0.81%). Método AG2.

### 6.1.3.3 Aprendizaje global del número óptimo de clases y de las funciones de pertenencia de las clases (AG3)

A continuación se muestran los resultados obtenidos por el AG propuesto para el aprendizaje conjunto de los parámetros de fusificación de FIR. El número de evaluaciones de cromosomas aumenta mucho más aún, ya que el espacio de búsqueda es muy grande.

Las *Tablas 10* y *11* muestran los resultados obtenidos para los cinco controladores (RC, TV, RP, CM y RCA) usando las funciones objetivo 1-Q y el error de predicción del 25% del conjunto de datos de entrenamiento ( $MSE_{entr}$ ), respectivamente. Ambas tablas están organizadas como se describe a continuación. La primera columna se divide en cinco secciones. Las secciones *A* y *B* corresponden a los mejores y peores resultados obtenidos por el **AG3**, respectivamente. La sección *C* muestra el resultado cuando se utilizan los parámetros de fusificación por defecto de FIR (**FIR DEFECTO**), es decir, todas las variables se discretizan en tres clases y se emplea el método EFP para obtener funciones de pertenencia uniformes. Las secciones *D* y *E* muestran los mejores resultados cuando se emplean los enfoques **AG1** y **AG2**. El primero se encarga de obtener un nivel óptimo de granularidad (número de clases) para cada variable y el segundo tiene como objetivo la determinación de las funciones de pertenencia de las clases. El **AG1** usa el método EFP para determinar las funciones de pertenencia para cada variable. El **AG2** sugiere funciones de pertenencia no uniformes manteniendo siempre invariable la granularidad dada por el **AG1** (**AG1+AG2**). *C*, *D* y *E* se usan más adelante como referencia en el análisis de los resultados. La segunda columna indica el número de evaluaciones de cromosomas realizadas por el **AG3**. La tercera y cuarta columnas presentan la solución sugerida por el **AG3** (granularidad y proporción de datos) para la variable de entrada (PSC) y las variables de salida de cada controlador (RC, TV, RP, CM, RCA). El número de elementos de la proporción de datos va de acuerdo a la granularidad de cada variable. La quinta columna contiene la máscara óptima (en notación de posición) encontrada por FIR cuando se utiliza esa granularidad y se emplea esa proporción de datos para el cómputo de los landmarks. La sexta columna corresponde a la calidad asociada a la máscara óptima. La séptima columna es el valor de la función objetivo utilizada, 1-Q ó  $MSE_{entr}$ . La última columna indica el valor medio del error de predicción (MSE) obtenido para los 6 conjuntos de datos de prueba.

El tiempo de CPU necesario para realizar las 30 ejecuciones del controlador RC (usando 7000 evaluaciones) para las funciones objetivo 1-Q y  $MSE_{entr}$ , fue de 44:00 horas y 357:55 horas, respectivamente. Las *figuras 20* y *21* exhiben el seguimiento de las mejores y peores señales de predicción de los conjuntos de prueba de la solución óptima sugerida por el **AG3** para el controlador RCA en cada función objetivo.

**Tabla 10.** Resultados de la granularidad y las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo 1-Q.

<b>Controlador RC</b>						
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A	14000	(9,3) PSC:(0.06,0.06,0.15,0.13,0.11,0.15,0.21,0.08,0.05) RC:(0.45,0.36,0.19)	(1,4,6)	0.9845	0.0155	2.02%
	7000	(9,3) PSC:(0.07,0.07,0.06,0.25,0.06,0.25,0.11,0.08,0.05) RC:(0.44,0.37,0.19)	(1,4,6)	0.9846	0.0154	2.26%
B	14000	(5,2) PSC:(0.06,0.40,0.21,0.28,0.05); RC:(0.31,0.69)	(3,4,6)	0.9838	0.0162	79.89%
	7000	(6,3) PSC:(0.06,0.09,0.20,0.42,0.18,0.05); RC:(0.40,0.42,0.18)	(1,4,6)	0.9839	0.0161	1.07%
C	(3,3)	Equal Frequency Partition method	(3,4,6)	0.9725	0.0275	1.19%
D	(9,3)	Equal Frequency Partition method	(1,4,6)	0.9787	0.0213	3.92%
E	(9,3)	PSC:(0.06,0.09,0.18,0.12,0.09,0.22,0.11,0.08,0.05) RC:(0.38,0.43,0.19)	(1,4,6)	0.9845	0.0155	2.03%
<b>Controlador TV</b>						
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A	14000	(9,2) PSC:(0.06,0.08,0.06,0.09,0.10,0.22,0.16,0.18,0.05); TV:(0.30,0.70)	(3,4,6)	0.9731	0.0269	16.21%
	7000	(8,2) PSC:(0.06,0.15,0.08,0.18,0.17,0.13,0.18,0.05); TV:(0.30,0.70)	(3,4,6)	0.9729	0.0271	6.88%
B	14000	(9,7) PSC:(0.09,0.06,0.07,0.11,0.12,0.12,0.10,0.24,0.09) TV:(0.30,0.34,0.15,0.05,0.05,0.06,0.05)	(3,4,6)	0.8446	0.1554	7.08%
	7000	(9,6) PSC:(0.16,0.13,0.07,0.16,0.08,0.19,0.11,0.05,0.05) TV:(0.06,0.05,0.06,0.05,0.06,0.72)	(4,5,6)	0.9504	0.0496	0.78%
C	(3,3)	Equal Frequency Partition method	(3,4,6)	0.6971	0.3029	73.05%
D	(8,7)	Equal Frequency Partition method	(3,4,6)	0.8142	0.1858	9.49%
E	(8,7)	PSC:(0.21,0.17,0.19,0.07,0.13,0.07,0.11,0.05) TV:(0.06,0.08,0.05,0.05,0.06,0.62,0.08)	(3,4,6)	0.9332	0.0668	6.17%
<b>Controlador RP</b>						
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A	14000	(9,2) PSC:(0.20,0.08,0.12,0.08,0.13,0.09,0.17,0.06,0.07); RP:(0.05,0.95)	(4,5,6)	0.9750	0.0250	63.20%
	7000	(9,2) PSC:(0.20,0.08,0.12,0.08,0.09,0.10,0.11,0.15,0.07); RP:(0.05,0.95)	(4,5,6)	0.9750	0.0250	49.13%
B	14000	(6,7) PSC:(0.15,0.12,0.11,0.15,0.32,0.15) RP:(0.13,0.16,0.13,0.14,0.14,0.14,0.16)	(4,5,6)	0.8774	0.1226	24.74%
	7000	(8,7) PSC:(0.13,0.10,0.07,0.08,0.07,0.19,0.18,0.18) RP:(0.13,0.16,0.14,0.14,0.14,0.13,0.16)	(4,5,6)	0.8800	0.1200	13.55%
C	(3,3)	Equal Frequency Partition method	(3,4,6)	0.6662	0.3338	36.08%
D	(8,7)	Equal Frequency Partition method	(4,5,6)	0.8552	0.1448	6.10%
E	(8,7)	PSC:(0.13,0.13,0.11,0.09,0.15,0.12,0.14,0.13) RP:(0.18,0.11,0.15,0.14,0.15,0.11,0.16)	(4,5,6)	0.8831	0.1169	19.56%
<b>Controlador CM</b>						
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A	14000	(9,2) PSC:(0.06,0.15,0.15,0.18,0.18,0.07,0.10,0.06,0.05); CM:(0.19,0.81)	(3,4,6)	0.9739	0.0261	66.81%
	7000	(9,2) PSC:(0.07,0.14,0.13,0.09,0.18,0.11,0.07,0.11,0.10); CM:(0.19,0.81)	(3,4,6)	0.9731	0.0269	76.76%
B	14000	(8,7) PSC:(0.13,0.13,0.13,0.11,0.16,0.10,0.13,0.11) CM:(0.09,0.05,0.05,0.05,0.06,0.65,0.05)	(4,5,6)	0.9347	0.0653	1.03%
	7000	(7,6) PSC:(0.15,0.12,0.11,0.19,0.12,0.20,0.11) CM:(0.10,0.14,0.06,0.60,0.05,0.05)	(4,5,6)	0.9291	0.0709	58.53%
C	(3,3)	Equal Frequency Partition method	(3,4,6)	0.6970	0.3030	73.05%
D	(8,7)	Equal Frequency Partition method	(3,4,6)	0.8166	0.1834	8.37%
E	(8,7)	PSC:(0.17,0.19,0.20,0.08,0.11,0.07,0.07,0.11) CM:(0.06,0.05,0.06,0.06,0.05,0.67,0.05)	(4,5,6)	0.9371	0.0629	3.63%
<b>Controlador RCA</b>						
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	1-Q	MSE <sub>prue</sub>
A	14000	(3,2) PSC:(0.72,0.23,0.05); RCA:(0.92,0.08)	(1,4,5,6)	0.9714	0.0286	6.88%
	7000	(3,2) PSC:(0.78,0.17,0.05); RCA:(0.92,0.08)	(3,4,5,6)	0.9702	0.0298	6.31%
B	14000	(8,2) PSC:(0.12,0.12,0.13,0.11,0.14,0.12,0.14,0.12); RCA:(0.23,0.77)	(4,5,6)	0.9023	0.0977	5.58%
	7000	(8,5) PSC:(0.22,0.08,0.10,0.11,0.10,0.09,0.11,0.19) RCA:(0.10,0.05,0.05,0.05,0.75)	(4,5,6)	0.8927	0.1073	5.78%
C	(3,3)	Equal Frequency Partition method	(4,5,6)	0.6254	0.3746	17.86%
D	(7,2)	Equal Frequency Partition method	(4,5,6)	0.8326	0.1674	13.76%
E	(7,2)	PSC:(0.05,0.20,0.27,0.08,0.09,0.05,0.26) RCA:(0.06,0.94)	(4,5,6)	0.9421	0.0579	11.49%

**Tabla 11.** Resultados de la granularidad y las funciones de pertenencia para los controladores RC, TV, RP, CM y RCA usando la función objetivo  $MSE_{entr}$ .

<b>Controlador RC</b>							
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$	
A	14000	(2.4) PSC:(0.53.0.47) ; RC:(0.10.0.19.0.55.0.16)	(1.4.5.6)	0.9541	0.05%	0.10%	
	7000	(2.4) PSC:(0.93.0.07) ; RC:(0.12.0.17.0.54.0.17)	(1.3.4.6)	0.9460	0.05%	0.10%	
B	14000	(2.5) PSC:(0.53.0.47) ; RC:(0.20.0.22.0.24.0.26.0.08)	(1.4.5.6)	0.9566	0.07%	9.93%	
	7000	(2.5) PSC:(0.53.0.47) ; RC:(0.20.0.24.0.20.0.20.0.16)	(1.4.5.6)	0.9644	0.07%	0.13%	
C	(3.3)	Equal Frequency Partition method	(3.4.6)	0.9725	---	1.19%	
D	(2.5)	Equal Frequency Partition method	(1.4.5.6)	0.9642	0.08%	0.16%	
E	(2.5)	PSC:(0.54.0.46) ; RC:(0.13.0.16.0.55.0.11.0.05)	(1.4.5.6)	0.9488	0.05%	0.11%	

<b>Controlador TV</b>							
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$	
A	14000	(2.9) PSC:(0.44.0.56)	(1.4.5.6)	0.7496	0.12%	0.28%	
	14000	(2.8) TV:(0.09.0.08.0.08.0.19.0.09.0.10.0.14.0.14.0.09) PSC:(0.44.0.56)	(1.4.5.6)	0.7563	0.13%	0.29%	
	14000	(2.6) TV:(0.09.0.08.0.08.0.20.0.11.0.23.0.11.0.10)	(3.4.5.6)	0.7682	0.24%	0.57%	
B	14000	(2.6) PSC:(0.41.0.59)	(3.4.5.6)	0.7682	0.24%	0.57%	
	7000	(4.6) TV:(0.12.0.11.0.13.0.36.0.16.0.12) PSC:(0.13.0.22.0.11.0.54)	(4.5.6)	0.7964	0.33%	0.75%	
	7000	(3.9) TV:(0.16.0.16.0.18.0.16.0.17.0.17)	(4.5.6)	0.7752	0.29%	0.73%	
C	(3.3)	Equal Frequency Partition method	(3.4.6)	0.6971	---	73.05%	
D	(2.5)	Equal Frequency Partition method	(1.3.4.6)	0.7315	0.61%	1.69%	
E	(2.5)	PSC:(0.38.0.62) ; TV:(0.11.0.14.0.31.0.22.0.22)	(1.4.5.6)	0.7475	0.19%	0.46%	

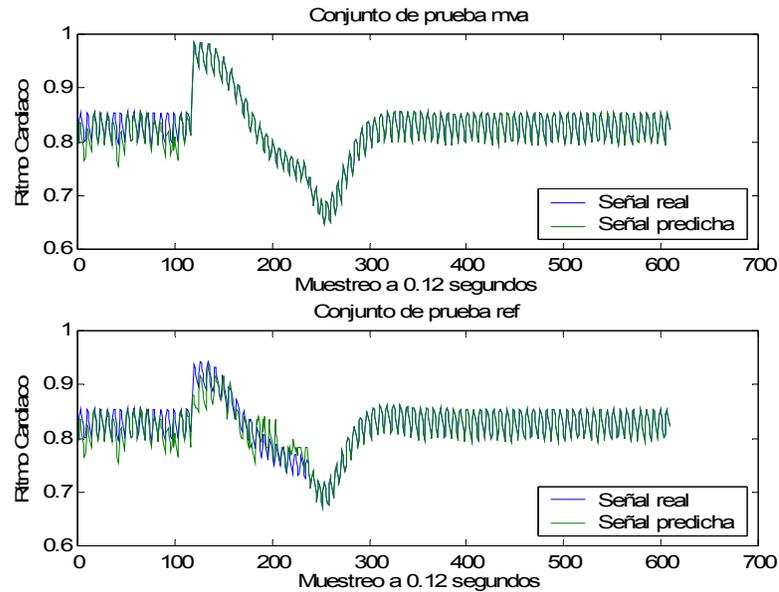
<b>Controlador RP</b>							
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$	
A	14000	(2.8) PSC:(0.27.0.73)	(3.4.5.6)	0.7541	0.06%	0.15%	
	7000	(2.9) RP:(0.06.0.07.0.08.0.08.0.17.0.21.0.14.0.19) PSC:(0.26.0.74)	(3.4.5.6)	0.7910	0.07%	0.19%	
B	14000	(2.7) RP:(0.08.0.10.0.05.0.05.0.2.0.09.0.16.0.10.0.17) PSC:(0.17.0.83)	(4.5.6)	0.7418	0.37%	0.96%	
	7000	(3.9) RP:(0.14.0.15.0.15.0.14.0.13.0.14.0.15) PSC:(0.18.0.18.0.64)	(4.5.6)	0.7752	0.29%	0.73%	
	7000	(3.9) RP:(0.12.0.08.0.09.0.16.0.08.0.12.0.09.0.11.0.15)	(4.5.6)	0.8064	0.20%	0.80%	
C	(3.3)	Equal Frequency Partition method	(3.4.6)	0.6662	---	36.08%	
D	(4.9)	Equal Frequency Partition method	(3.4.6)	0.7497	0.93%	3.05%	
E	(4.9)	PSC:(0.23.0.20.0.23.0.34)	(4.5.6)	0.8064	0.20%	0.80%	
	7000	(3.9) RP:(0.09.0.10.0.08.0.13.0.14.0.08.0.13.0.10.0.15)	(4.5.6)	0.8064	0.20%	0.80%	

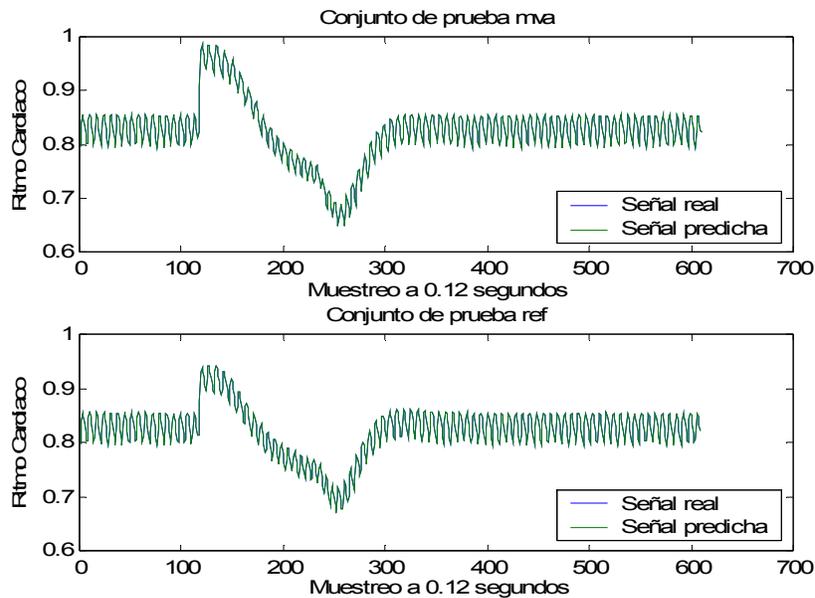
<b>Controlador CM</b>							
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$	
A	14000	(2.9) PSC:(0.42.0.58)	(3.4.5.6)	0.7959	0.14%	0.30%	
	7000	(2.9) CM:(0.07.0.05.0.05.0.06.0.06.0.11.0.20.0.34.0.06) PSC:(0.44.0.56)	(1.4.5.6)	0.7465	0.13%	0.30%	
B	14000	(3.7) CM:(0.09.0.08.0.08.0.21.0.11.0.14.0.08.0.12.0.09) PSC:(0.34.0.07.0.59)	(4.5.6)	0.7661	0.27%	0.65%	
	7000	(7.7) CM:(0.14.0.14.0.14.0.12.0.15.0.16.0.15) PSC:(0.11.0.31.0.09.0.09.0.10.0.08.0.22)	(4.5.6)	0.7654	0.35%	0.70%	
	7000	(7.7) CM:(0.13.0.05.0.12.0.23.0.28.0.12.0.07)	(4.5.6)	0.7452	0.20%	0.45%	
C	(3.3)	Equal Frequency Partition method	(3.4.6)	0.6970	---	73.05%	
D	(4.9)	Equal Frequency Partition method	(3.4.6)	0.7346	0.60%	2.53%	
E	(4.9)	PSC:(0.20.0.29.0.31.0.20)	(4.5.6)	0.7452	0.20%	0.45%	
	7000	(7.7) CM:(0.09.0.08.0.12.0.13.0.12.0.11.0.11.0.11.0.13)	(4.5.6)	0.7452	0.20%	0.45%	

<b>Controlador RCA</b>							
# Eval	Gran.	Proporción de Datos	Opt. Mask	Q	$MSE_{entr}$	$MSE_{prue}$	
A	14000	(8.5) PSC:(0.10.0.12.0.10.0.11.0.11.0.15.0.14.0.17)	(1.5.6)	0.6166	8.85e-30%	9.47e-30%	
	7000	(7.5) RCA:(0.11.0.42.0.32.0.09.0.06) PSC:(0.13.0.13.0.11.0.13.0.12.0.13.0.25)	(1.5.6)	0.5987	9.71e-30%	8.72e-30%	
B	14000	(5.9) RCA:(0.25.0.39.0.19.0.12.0.05) PSC:(0.14.0.40.0.18.0.14.0.14)	(4.5.6)	0.6798	0.49%	1.24%	
	7000	(4.8) RCA:(0.11.0.11.0.11.0.11.0.11.0.11.0.13.0.10.0.11) PSC:(0.28.0.29.0.27.0.16)	(4.5.6)	0.7090	0.60%	1.77%	
	7000	(4.8) RCA:(0.12.0.12.0.12.0.12.0.13.0.17.0.11.0.11)	(4.5.6)	0.6254	---	17.86%	
C	(3.3)	Equal Frequency Partition method	(4.5.6)	0.6254	---	17.86%	
D	(3.7)	Equal Frequency Partition method	(4.5.6)	0.6761	0.89%	9.37%	
E	(3.7)	PSC:(0.32.0.30.0.38)	(4.5.6)	0.6837	0.44%	1.49%	
	7000	(4.8) RCA:(0.13.0.10.0.12.0.12.0.29.0.12.0.12)	(4.5.6)	0.6837	0.44%	1.49%	



**Figura 20.** Mejor y peor predicción usando los modelos FIR para el controlador RCA con granularidad (3,2) y distribución de landmarks: CSP:(0.72,0.23,0.05)-RCA:(0.92, 0.08). Función objetivo  $1-Q$  (errores mva=2.69% ref=11.19%). Método AG3.



**Figura 21.** Mejor y peor predicción usando los modelos FIR para el controlador RCA con granularidad (8,5) y distribución de landmarks: CSP:(0.10,0.12,0.10,0.11,0.11,0.15,0.14, 0.17)- RCA:(0.11,0.42,0.32,0.09,0.06). Función objetivo  $MSE_{entr}$  (errores mva=9.41e-30% y ref=1.24e-29%). Método AG3.

### 6.1.4 Análisis de resultados y conclusiones

En este apartado se analizan los resultados obtenidos por los métodos **AG1**, **AG2** y **AG3** propuestos en esta memoria para la resolución del problema de modelado del sistema nervioso central humano. Asimismo, se presentan algunas conclusiones.

#### 6.1.4.1 Método AG1

En la *Tabla 5* se observa que para el controlador RC el **AG1** sugiere al alcanzar 40 evaluaciones, 5 y 4 particiones diferentes para las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente. Cuando el criterio de parada es realizar 64 evaluaciones, el **AG1** encuentra, para ambas funciones objetivo, 4 particiones diferentes. La solución óptima para la función objetivo 1-Q corresponde a la partición (9,3) con una calidad de máscara de 0.9787. Las particiones encontradas para ambos criterios de parada aconsejan siempre 3 clases para la variable de salida y 5,6,7,8 o 9 clases para la variable de entrada. Nótese que las calidades ( $Q$ ) de todas las particiones sugeridas están muy cercanas a la óptima.

La función objetivo  $MSE_{entr}$  corresponde al error de predicción de una parte del conjunto de datos de entrenamiento. Como ya se mencionó anteriormente la máscara se obtiene usando exclusivamente el primer 75% de los datos de la señal de entrenamiento. Por consiguiente, los datos usados para la evaluación de la función objetivo no han sido usados en el proceso de modelado. Esta es la razón por la cual los menores valores  $MSE_{entr}$  no correspondan necesariamente a las particiones con la máscara óptima asociada de más alta calidad. Sin embargo, la calidad de las máscaras óptimas encontradas para las particiones sugeridas es aún alta. La solución óptima es la partición (2,5) con un  $MSE_{entr}$  de 0.08%, que es realmente muy bajo. La partición (3,4) es la mejor solución subóptima, es decir, con el segundo error  $MSE_{entr}$  más bajo. Por lo tanto, el **AG1** es capaz de obtener las mejores particiones posibles. Nótese que el tiempo de CPU se incrementa considerablemente con respecto a la función objetivo 1-Q. Esto se debe al hecho que ahora el proceso de predicción de la metodología FIR necesita ejecutarse para cada partición evaluada.

En las *Tablas 6* y *7* puede observarse que la solución óptima (denotada por un \*) se alcanza para todos los controladores, para ambas funciones objetivo y criterios de parada.

Es interesante analizar el  $MSE_{prue}$  de las *Tablas 5*, *6* y *7*. Según lo esperado, la función objetivo  $MSE_{entr}$  es capaz de obtener particiones con una eficiencia más alta en la predicción de los conjuntos de datos de prueba que la función objetivo 1-Q. Sin embargo, los resultados obtenidos son muy buenos si los comparamos tanto con los de otras metodologías inductivas (Véase *Tabla 3*) como con los de **FIR DEFECTO**

(Véase secciones *C* de la *Tabla 11*). Si analizamos los  $MSE_{prue}$  de los mejores resultados obtenidos por cada método para el controlador RC, vemos que el menor  $MSE_{prue}$  se obtiene usando el **AG1** con la función objetivo  $MSE_{entr}$  (0.16%), luego viene el **FIR DEFECTO** con un error de 1.19%, seguido por el **AG1** con función objetivo 1-Q con un error de 3.92%. Finalmente el menor error obtenido con otras metodologías inductivas para este controlador es de 25.5% y se consigue usando un sistema NARMAX. La *Tabla 12* presenta el resultado de realizar este mismo análisis para todos los controladores del SNC.

**Tabla 12.**  $MSE_{prue}$  de los mejores resultados obtenidos (desde la perspectiva de las funciones objetivo) usando: AG1 con ambas funciones objetivo, FIR DEFECTO y otras metodologías inductivas.

	<b>AG1 (<math>MSE_{entr}</math>)</b>	<b>AG1 (1-Q)</b>	<b>FIR DEFECTO</b>	<b>OTRAS MET. IND.</b>
RC	0.16%	3.92%	1.19%	25.5%
TV	1.69%	9.49%	73.05%	22%
RP	3.05%	6.10%	36.08%	18.5%
CM	2.53%	8.37%	73.05%	22%
RCA	9.37%	13.76%	17.86%	9.3%

La *Tabla 12* hace evidente el buen comportamiento de los modelos FIR obtenidos usando los parámetros de discretización recomendados por el **AG1**. Para la mayoría de los controladores, los mejores resultados se obtienen cuando se usa el **AG1** con la función objetivo  $MSE_{entr}$  y en muchas ocasiones los siguientes mejores se consiguen con la función objetivo 1-Q. La excepción se encuentra en el controlador RCA, donde el mejor resultado de predicción se logra con un modelo NARMAX. Sin embargo, el error de 9.3% obtenido por éste es prácticamente el mismo que el obtenido con FIR usando el **AG1** con la función objetivo  $MSE_{entr}$ . En general los errores obtenidos usando otras metodologías inductivas son menores a los conseguidos con **FIR DEFECTO**.

Se observa en todos los controladores que la función objetivo 1-Q necesita menos tiempo para ser evaluada pero tiene menor eficiencia en la predicción del conjunto de datos de prueba. Contrariamente, la función objetivo  $MSE_{entr}$  es más costosa desde la perspectiva del tiempo de CPU pero la eficiencia es más alta.

Claramente la aplicación biomédica considerada es un problema de optimización simple, debido al hecho de que solamente intervienen dos variables y se permite un máximo de nueve clases por variable (de hecho hay solamente ocho, debido a que la clase 1 no se usa). Por lo tanto, existen únicamente 64 posibles soluciones y una búsqueda exhaustiva se puede hacer fácilmente. Sin embargo, estudios tales como [Sch99, GP01] sugieren la escalabilidad de los AGs. Por

consiguiente, puede asumirse que este enfoque trabajará bien en problemas de optimización complejos donde una búsqueda exhaustiva sería impracticable. Esto se demuestra más adelante por medio de las otras dos aplicaciones, eléctrica y ecológica, consideradas en este capítulo.

#### 6.1.4.2 Método AG2

Si analizamos las *Tablas 8 y 9* vemos que para todos los controladores, los peores resultados obtenidos por el **AG2**, valores máximos de las columnas 1-Q y  $MSE_{entr}$  en la fila *B* de cada tabla, son inferiores al mejor resultado obtenido cuando se utiliza **AG1+EFP** para obtener las funciones de pertenencia de cada clase (valor de referencia *D* en las *Tablas 10 y 11*, respectivamente). Por consiguiente, aunque el AG no asegura la solución óptima, todas las soluciones encontradas son mejores cuando éste se usa. Es decir, para esta aplicación se obtienen mejores resultados usando **AG1+AG2** que usando únicamente **AG1+EFP**. Cabe señalar aquí que cuando hablamos de mejores resultados nos referimos a los valores que estamos minimizando (funciones objetivo): 1-Q y  $MSE_{entr}$  y que un menor valor de estas funciones no implica necesariamente un menor error de predicción en el conjunto de datos de prueba ( $MSE_{prue}$ ). Recordemos que los modelos FIR se identifican usando los datos de entrenamiento y por lo tanto sólo garantizan su optimalidad para estos datos.

Si se analiza la última columna,  $MSE_{prue}$ , de cada tabla, se puede observar nuevamente que la función objetivo  $MSE_{entr}$  es capaz de obtener funciones de pertenencia que implican mayor eficiencia en la predicción del conjunto de datos de prueba que la función objetivo 1-Q, pero con el inconveniente de tener un mayor costo computacional. Aunque en la función objetivo 1-Q los  $MSE_{prue}$  son usualmente menores cuando se utiliza el **AG2**, esto no sucede siempre ya que el conjunto de datos de prueba no ha sido usado en el proceso de modelado cualitativo de FIR. Sin embargo, los mejores resultados obtenidos por el **AG2** (fila *A* de cada tabla) son muy buenos si se comparan con los mejores obtenidos por el **AG1**, **FIR DEFECTO** y otras metodologías inductivas de la *Tabla 3*.

#### 6.1.4.2 Método AG3

Si analizamos las *Tablas 10 y 11* vemos que los peores resultados obtenidos, valores máximos de las columnas 1-Q y  $MSE_{entr}$  de la fila *B* de cada tabla, usando el **AG3** son siempre inferiores que el resultado de referencia *D* (cuando se usa **AG1+EFP**). Por otra parte, se puede señalar que las mejores soluciones obtenidas, valores mínimos de las columnas 1-Q y  $MSE_{entr}$  de la fila *A* de cada tabla, son inferiores al resultado de referencia *E* (cuando se usa **AG1+AG2**).

Resulta interesante también analizar los  $MSE_{prue}$  (última columna) para ver la precisión de cada modelo obtenido. Nótese que de nuevo los  $MSE_{prue}$  son usualmente

menores cuando se utiliza el **AG3**. Sin embargo, esto no sucede siempre para la función objetivo 1-Q ya que los  $MSE_{prue}$  para los controladores RP y CM son menores cuando se utilizan las opciones **AG1+EFP** y **AG1+AG2**. Como se ha señalado anteriormente, esto se debe a que el conjunto de datos de prueba no es utilizado en el proceso de modelado cualitativo de FIR y por lo tanto no se puede garantizar que se obtendrá la mejor predicción de los datos de prueba mediante el modelo FIR óptimo identificado usando los datos de entrenamiento.

Por otro lado, es necesario destacar que para la función objetivo 1-Q el modelo FIR para el controlador RP con mayor poder de predicción se ha conseguido usando el **AG1** ( $MSE_{prue}=6.10\%$ ) en donde se utiliza el método EFP para obtener las funciones de pertenencia. Ni los obtenidos con el **AG3** ( $MSE_{prue}=49.13\%$ ) ni con la combinación **AG1+AG2** ( $MSE_{prue}=19.56\%$ ) han mejorado su habilidad de predicción.

Una vez más la función objetivo  $MSE_{entr}$  superó a 1-Q en obtener particiones difusas con mayor eficiencia en la predicción en los conjuntos de datos de prueba. Los  $MSE_{prue}$  obtenidos por la metodología FIR en este caso son significativamente más bajos que los obtenidos por las metodologías inductivas de la *Tabla 3* y los resultados de referencia (C), (D) y (E), tal y como se resume en la *Tabla 13*.

**Tabla 13.**  $MSE_{prue}$  de los mejores resultados obtenidos (desde la perspectiva de las funciones objetivo) usando: AG3, AG1+AG2, AG1+EFP, FIR DEFECTO y otras metodologías inductivas. Se muestran únicamente los resultados de los AG con función objetivo  $MSE_{entr}$ .

	<b>AG3</b>	<b>AG1+AG2</b>	<b>AG1+EFP</b>	<b>FIR DEF.</b>	<b>OTRAS MET. IND.</b>
RC	0.10%	0.11%	0.16%	1.19%	25.5%
TV	0.28%	0.46%	1.69%	73.05%	22%
RP	0.15%	0.80%	3.05%	36.08%	18.5%
CM	0.30%	0.45%	2.53%	73.05%	22%
RCA	9.47e-30%	1.49%	9.37%	17.86%	9.3%

Estos modelos FIR poseen particiones difusas con diferente número de clases y funciones de pertenencia a los resultados de referencia indicados, y esto se debe, a que el **AG3** exploró otras zonas prometedoras del espacio de búsqueda y encontró soluciones que poseen una buena capacidad de predicción. Sin embargo, como ya se ha señalado en repetidas ocasiones la función objetivo  $MSE_{entr}$  tiene un mayor costo computacional. Por lo tanto, es el usuario quien debe decidir qué función objetivo usar teniendo en cuenta el tamaño del problema de optimización y sus propias necesidades y requerimientos.

Como comentario final se puede indicar que en general el **AG3** para la aplicación biomédica estudiada tuvo un rendimiento mejor que la combinación **AG1+AG2**.

## **6.2 Caso de estudio: Costos óptimos de la línea de media tensión en núcleos urbanos**

### **6.2.1 Descripción del problema**

La energía eléctrica constituye un recurso muy valioso en la economía de un país. Representa un servicio imprescindible para la vida diaria de los ciudadanos y que incorpora un valor estratégico innegable al resto de los sectores de la economía. Se transporta desde los generadores hasta los puntos de consumo mediante líneas eléctricas de distinta capacidad. Se distingue así entre líneas de alta tensión (100KV-400KV), líneas de media tensión (10KV-100KV) y líneas de baja tensión (menor de 10KV).

A partir del 1 de enero de 1998, la organización del sector eléctrico español experimenta una profunda renovación a través de una nueva ley, la cual supone un cambio importante en el funcionamiento y regulación del sistema que hasta esa fecha se había adoptado [CHS98, Mar02]. Estas disposiciones legales tienen una decidida voluntad de propiciar una mayor liberalización del sector eléctrico, asegurando la competencia empresarial y garantizando un menor costo de la energía eléctrica. Se persigue que las empresas compitan por precios en vez de por cuota de mercado.

La explotación del sistema eléctrico nacional es un problema complejo. El sistema español está compuesto por varios cientos de generadores (centrales térmicas, hidroeléctricas y nucleares, principalmente) y por una cantidad de consumidores con requisitos de potencia distintos, que abarcan desde viviendas familiares hasta grandes demandantes de energía como las empresas productoras de aluminio o las siderúrgicas. Existen cuatro grandes agentes (Unión FENOSA, HidroCantábrico, ENDESA, IBERDROLA) que poseen la mayoría de las plantas de generación de energía y las redes de distribución en España. Cualquiera de estas empresas podría tener una gran influencia sobre el precio de la energía eléctrica. Es por ello, que el gobierno español decidió nacionalizar hace algunos años las líneas de alta tensión y separar los mercados de distribución y de generación, obligando así a las compañías antes mencionadas a proceder en cada situación como dos actores diferentes. De esta manera, las plantas de generación venden la energía que producen en un mercado parcialmente regulado y las compañías de distribución compran la energía en ese mismo mercado (se permiten en ciertos casos los contratos bilaterales entre

proveedores de energía y consumidores).

Las líneas de baja y media tensión pertenecen a compañías determinadas, que se hacen cargo de su instalación y corren con los gastos de su mantenimiento. Como las líneas de alta tensión (en las que acaban derivando todas) son de titularidad compartida, existen problemas para repartir los ingresos por los abonos de la tarifa eléctrica entre las distintas compañías generadoras. Dichos ingresos se redistribuyen según cierto criterio complejo (cantidad de generación de energía de la compañía, número de clientes, etc) con el objetivo de equilibrar el mercado.

El sistema de reparto debe ser revisado cada cierto tiempo, ya que algunas compañías no comparten su formulación. Se alegan diferentes razones, entre tales se hace mención de que por motivos históricos y políticos, los criterios de división de los costos de instalación y mantenimiento de las líneas de media y baja tensión no se basan en el tendido idóneo del cableado (red óptima), sino en estimaciones de la longitud actual que benefician a las redes más antiguas e ineficientes (con la consiguiente falta de incentivos para su modernización). El crecimiento de las líneas de media y baja tensión que existen en un núcleo urbano ha sido desordenado, por lo cual resulta muy difícil y costoso medirlas. Es por ello, que se cuestiona simplemente la veracidad de la fórmula legal que estima los costos a partir de los descriptores del núcleo de la población.

Los costos de mantenimiento dependen (entre otros factores) de la longitud total de línea eléctrica que cada compañía posee, y de su tipo. Es evidente la importancia de los modelos de longitud de líneas en el mercado eléctrico español. Todas las compañías desarrollan sus propios modelos y el gobierno también usa algunos de estos modelos para decidir la parte del pago que las compañías reciben. Se necesita que se resuelva el siguiente problema: a) estimar el costo de mantenimiento de la red de media tensión instalada en ciertos pueblos, y b) relacionar ese costo de mantenimiento con otras características de estos pueblos [CHS98, San97].

El problema planteado no se basa en datos reales sino en estimaciones de los costos mínimos de mantenimiento de una línea de media tensión, las cuales están basadas en un modelo de la red eléctrica óptima para un pueblo, es decir, como si se pudiese tender toda la línea de nuevo.

Nos han sido proporcionados datos que incluyen cuatro características diferentes de los pueblos (Véase *Tabla 14*) y su costo mínimo de mantenimiento en una muestra de 1059 núcleos urbanos. En trabajos previos [CHS98, CHS99], estos datos fueron divididos en 847 ejemplos (conjunto de entrenamiento) y 212 ejemplos (conjunto de prueba), es decir, el 80% y el 20% del conjunto de datos disponible, respectivamente. Esta misma distribución de datos se utiliza en esta investigación con

el objetivo de poder comparar de manera precisa los resultados obtenidos por los métodos genéticos propuestos en esta memoria con los logrados en investigaciones previas. Las *figuras 22 y 23* muestran las señales de entrenamiento de las variables de entrada (LTC, AOC, AOE y ESC) y de la variable de salida CMT. El conjunto de prueba permite validar los modelos obtenidos, ya que son datos que no han sido utilizados en el proceso de identificación.

**Tabla 14.** Notación considerada para las variables del problema de estimación de los costos de mantenimiento de la línea de media tensión.

Símbolo	Significado	Rango
$x_1$ (LTC)	Suma de la longitud de todas las calles de la ciudad (Kms)	[0.5:11]
$x_2$ (AOC)	Área total ocupada por la ciudad (Kms <sup>2</sup> )	[0.15:8.55]
$x_3$ (AOE)	Área ocupada por edificios (Kms <sup>2</sup> )	[1.64:142.5]
$x_4$ (ESC)	Energía suministrada a la ciudad (MWh)	[1:165]
Y (CMT)	Costo de mantenimiento de la línea de media tensión (millones PTA)	[0:8546]

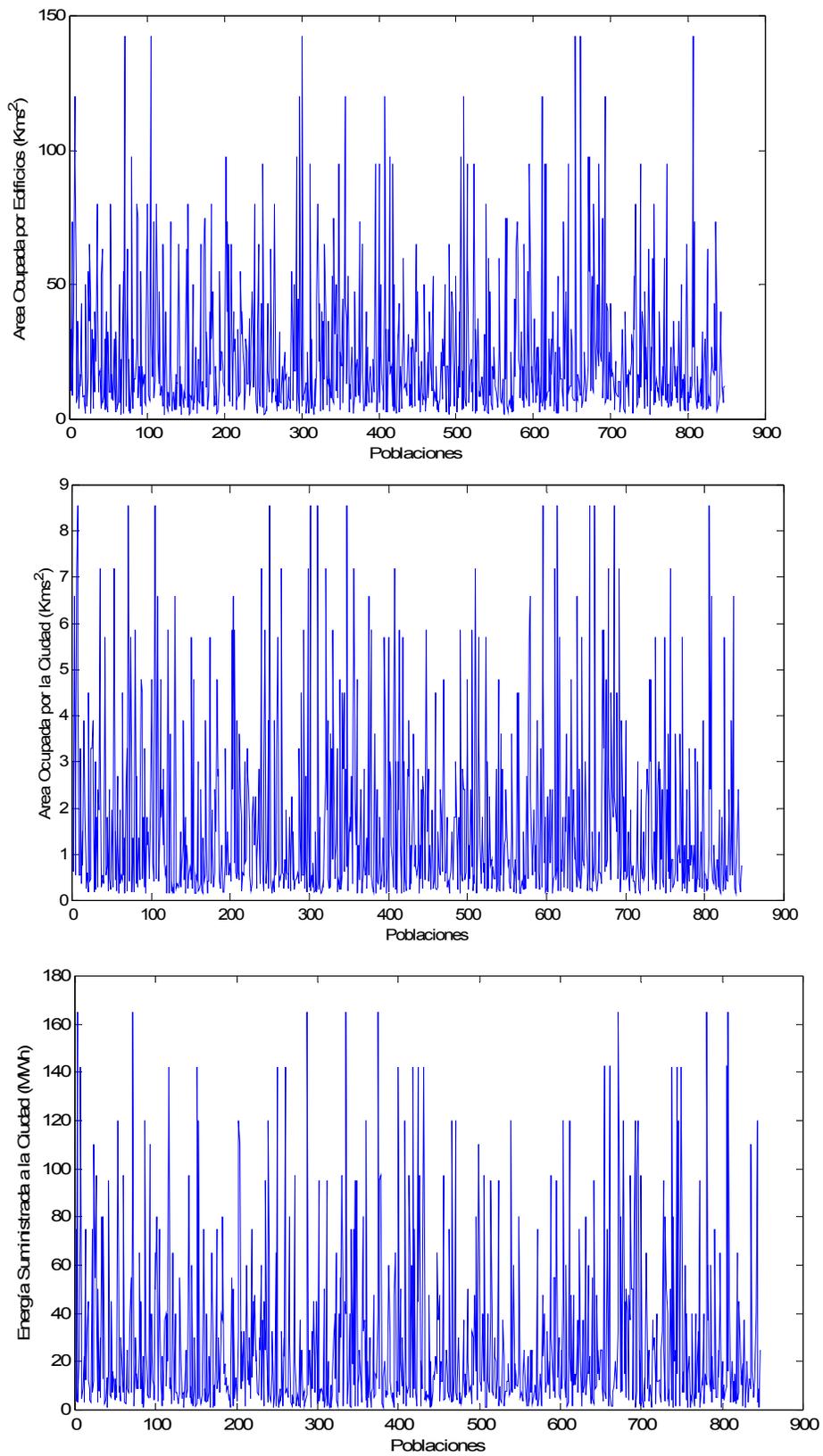
Es necesario señalar que no existen relaciones temporales entre dos ejemplos consecutivos del sistema, debido a que cada ejemplo representa un núcleo urbano concreto. En este estudio se persigue estimar y relacionar la variable costo de mantenimiento (CMT) con las otras cuatro variables (LTC, AOC, AOE y ESC) aplicando la metodología FIR. Es la primera vez que FIR se aplica a un sistema no dinámico. En este caso, las relaciones causales pueden ser encontradas definiendo una máscara de profundidad 1 (una fila), como se muestra en la ecuación 47. Los elementos -1 de la máscara candidata representan las relaciones causales potenciales con la variable de salida (elemento +1).

$$\begin{array}{c} t \backslash x \\ t \end{array} \left| \begin{array}{ccccc} \text{LTC} & \text{AOC} & \text{AOE} & \text{ESC} & \text{CMT} \\ -1 & -1 & -1 & -1 & +1 \end{array} \right| \quad (47)$$

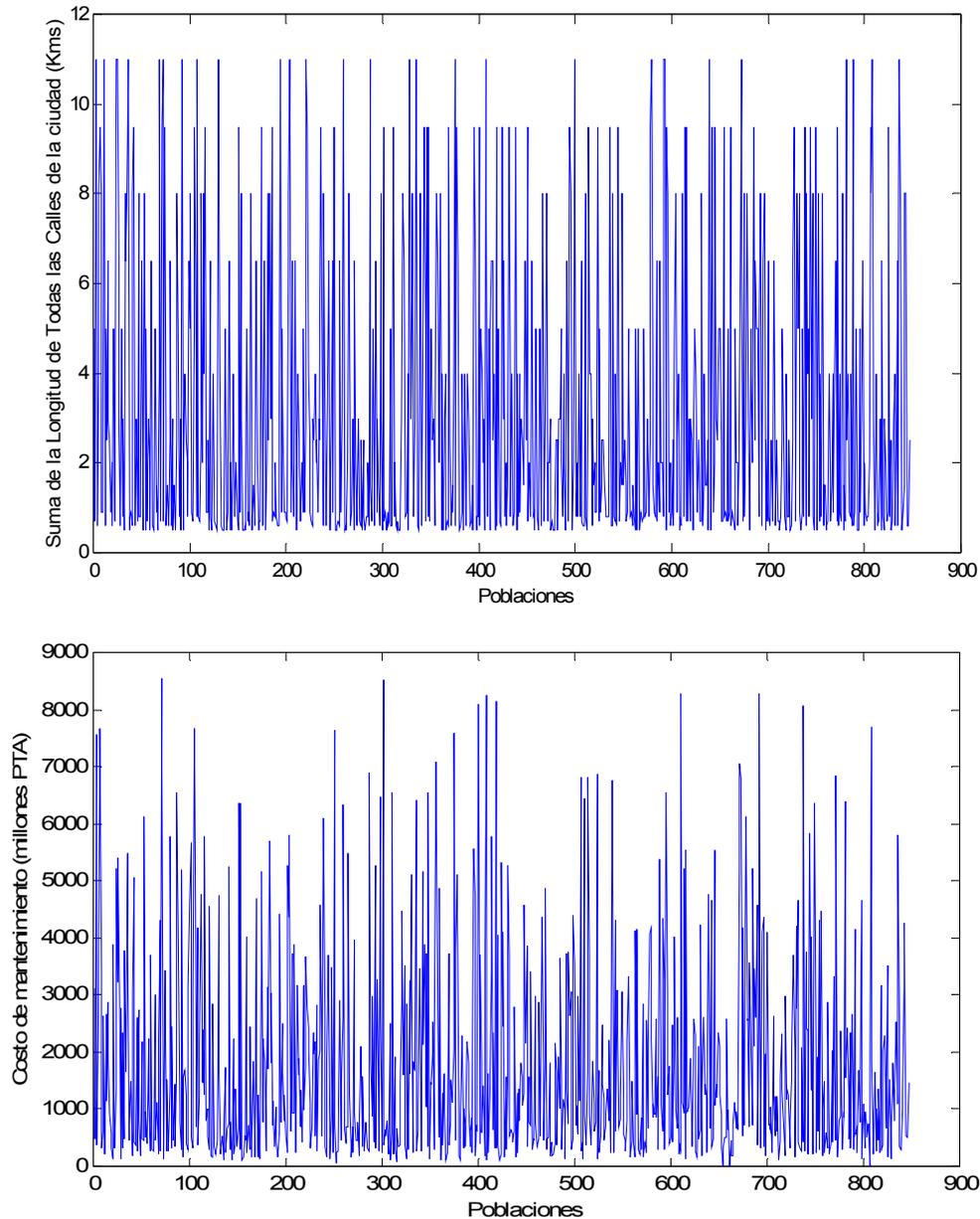
Como ya se explicó en el Capítulo I, una vez propuesta la máscara candidata el proceso de modelado cualitativo de la metodología FIR calcula todas las posibles máscaras y aquella con el máximo valor de la medida de calidad  $Q$  es considerada la óptima.

### 6.2.2 Trabajos previos

La *Tabla 15* contiene las predicciones obtenidas para el problema de estimación de los costos de mantenimiento de la línea de media tensión, primero aplicando métodos clásicos y luego usando técnicas híbridas evolutivas de análisis de datos [CHS99]. En lo que concierne a métodos clásicos, se consideraron modelos lineales, polinomiales y redes neuronales perceptrónicas multicapa (MLP). Por su parte, en



*Figura 22. Señales de entrenamiento para las variables de entrada AOE, AOC y ESC.*



**Figura 23.** Señales de entrenamiento para la variable de entrada LTC y la de salida CMT.

relación a las otras técnicas de análisis de datos se usaron: a) algoritmos GA-P (un híbrido entre programación genética y algoritmos genéticos) para realizar regresiones simbólicas, y b) sistemas genéticos basados en reglas difusas (SGBRD) para el diseño y optimización de modelos difusos (Wang-Mendel, Mandami y TSK). Una descripción completa de un algoritmo GA-P se puede encontrar en [HD95]. Todas las metodologías usan el mismo conjunto de datos de entrenamiento y de prueba descrito

previamente.

**Tabla 15.** Errores de predicción ECM de los modelos obtenidos del problema de estimación de costos de mantenimiento de la línea de media tensión usando tanto métodos clásicos como técnicas híbridas evolutivas de análisis de datos.

Método	ECM <sub>entr</sub>	ECM <sub>prue</sub>	Complejidad
Lineal	164662	36819	17 nodos, 5 part.
Polinomial (Segundo orden)	103032	45332	77 nodos, 15 part.
MLP (perceptrón 4-5-1)	86469	33105	35 part.
GA-P	18168	21884	50 nodos, 5 part.
GA-P Intervalo	16263	18325	15 nodos, 4 part.
Modelo difuso WM	20318	27615	66 reglas
Modelo difuso Mandami	19679	22591	63 reglas
Modelo difuso TSK ( $\alpha = 0$ )	25579	26450	268 reglas
Modelo difuso TSK ( $\alpha = 0,2$ )	11074	11836	268 reglas

Los errores presentados corresponden: a) Error cuadrático medio del modelo con respecto al conjunto de datos de entrenamiento (ECM<sub>entr</sub>), y b) Error cuadrático medio del modelo con respecto al conjunto de datos de prueba (ECM<sub>prue</sub>). En esta aplicación usaremos como medida de error el error cuadrático medio, ECM, definido de la siguiente manera:

$$ECM = \frac{1}{2 \cdot N} \sum_{i=1}^N (y_i(t) - \hat{y}(t))^2 \quad (48)$$

donde  $\hat{y}(t)$  es la salida predicha,  $y(t)$  la salida del sistema y  $N$  denota el número de ejemplos. Se usa esta medida en lugar de la empleada en la sección anterior (MSE) porque es la medida que se reporta en los trabajos previos y consideramos interesante poder comparar los resultados de los métodos propuestos en esta tesis con los publicados anteriormente por otros autores.

### 6.2.3 Modelos obtenidos por los métodos *AG1*, *AG2* y *AG3*

Al igual que en la aplicación anterior, en el problema planteado sobre la estimación de los costos óptimos de mantenimiento de la línea de media tensión se han realizado 30 ejecuciones (con distinto valor de la semilla) para cada función objetivo y criterio de parada en cada método genético propuesto en esta memoria. Además, se estudian dos funciones objetivo diferentes: a) 1-Q, y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento (MSE<sub>entr</sub>). En esta aplicación eléctrica el último 20% de la señal de entrenamiento se usa para la evaluación de la función objetivo MSE<sub>entr</sub> y solamente el 80% inicial de la señal se usa para la obtención de los modelos FIR (máscaras y base de reglas patrón).

Los valores seleccionados para los parámetros de los AGs se exponen en la *Tabla 16*. Las probabilidades de cruce y de mutación se han establecido de acuerdo a [Gre86]. Se utilizó un computador Pentium III (0.6 GHz) para realizar todas las simulaciones. En [AN05] se muestran todos los resultados obtenidos para los métodos **AG1** y **AG2**, mientras que los resultados del **AG3** se presentan en [ANF05b]. Los mejores resultados obtenidos por ambas funciones objetivo, 1-Q y  $MSE_{entr}$ , por los métodos propuestos se presentan en las *Tablas 17* a *23*.

**Tabla 16.** Parámetros de los AGs para el problema de estimación de costos óptimos de mantenimiento de la línea de media tensión en núcleos urbanos.

Parámetro	Tipo AG	Valor
<i>Tamaño de la población (# individuos)</i>	Todos	50
<i>Probabilidad de cruce</i>	Todos	0.6
<i>Probabilidad de mutación</i>	Todos	0.1
<i>Criterio de parada (# evaluaciones de cromosomas)</i>	AG1	{500, 1000, 2000}
	AG2	{5000, 10000, 20000}
	AG3	{5000, 10000, 20000, 40000, 80000, 160000, 320000}

Los resultados son mostrados en tres sub-apartados denominados en base al método empleado para la resolución del problema en estudio.

#### 6.2.3.1 Aprendizaje del número óptimo de clases para cada variable del sistema (AG1)

En este apartado se evalúa la utilidad del algoritmo genético propuesto para la optimización del número de clases (**AG1**) aplicado al problema de estimación de costos óptimos de mantenimiento en una línea de media tensión. Este AG se ejecuta para obtener el número óptimo de clases tanto para las variables de entrada como para la variable de salida del problema en estudio. El método EFP (*Equal Frequency Partition*) se usa para obtener las funciones de pertenencia de las clases. La *Tabla 17* presenta los mejores resultados de granularidad para ambas funciones objetivo. La tabla ha sido organizada de la siguiente manera:

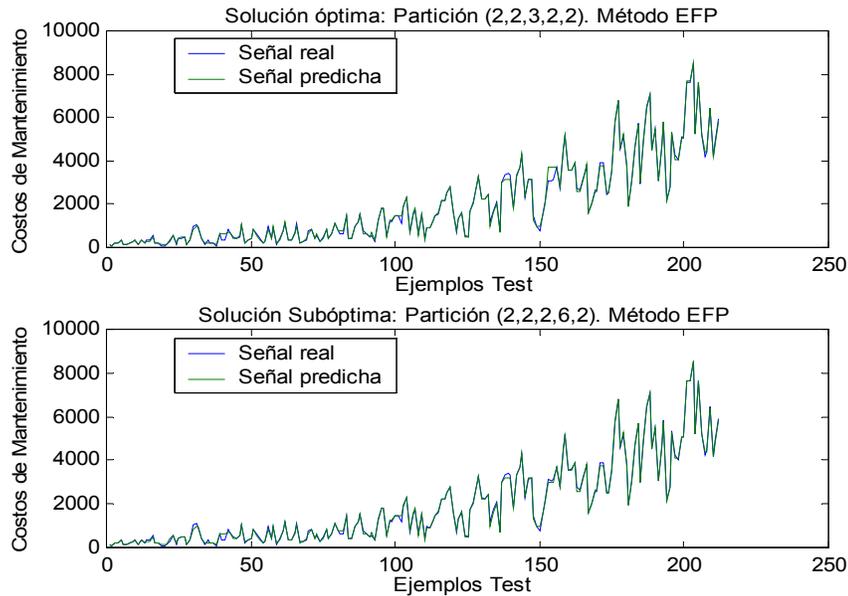
- *# Eval.* : es el número de evaluaciones de cromosomas realizadas por el AG.
- *Gran.*: señala la granularidad sugerida por el AG cuando finaliza su ejecución. Nótese que la granularidad sugerida es el parámetro de entrada al proceso de fusificación de la metodología FIR
- *Opt. Mask* y *Q*: contienen la máscara óptima (en notación de posición) identificada por FIR para esta granularidad específica y la calidad asociada a esta máscara, respectivamente.
- *1-Q*: corresponde al valor de la función objetivo de la calidad.

- $MSE_{entr}$ : corresponde al valor de la función objetivo del error de predicción MSE del último 20% de los datos del conjunto de entrenamiento.
- $ECM_{prue}$ : indica el error de predicción ECM con respecto al conjunto de datos de prueba
- *Tiempo*: contiene el tiempo promedio de CPU (en segundos), el cual se calcula tomando en cuenta el menor y el mayor tiempo consumido por el AG para encontrar la solución sugerida.

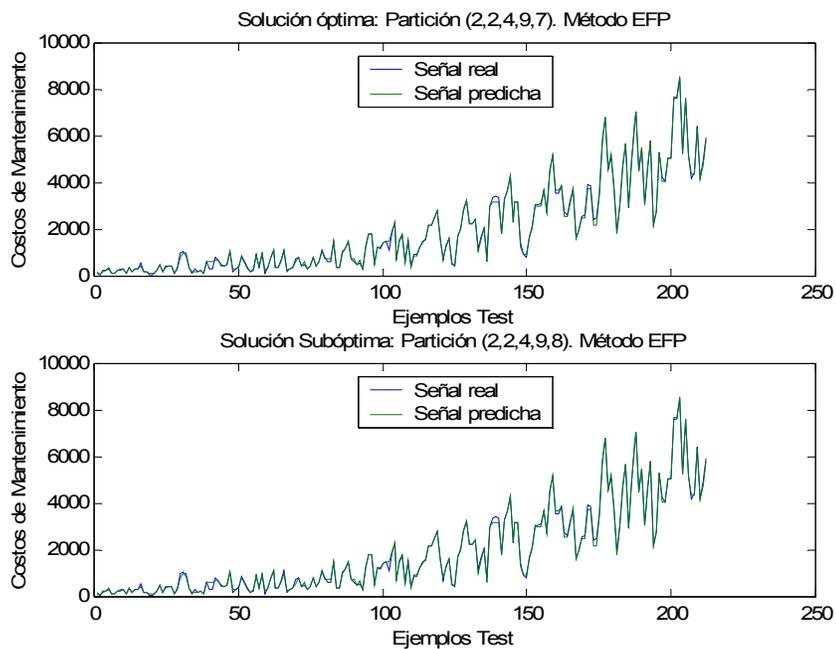
**Tabla 17.** Resultados de granularidad para el problema de estimación de costos óptimos de mantenimiento usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 20% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. (X significa cualquier valor entre [2..9]).

1-Q	# eval.	Gran.	Opt. Mask	Q	1-Q	$ECM_{prue}$	Tiempo
	2000	(X,X,3,2,2)	(3,4,5)	0.8365	0.1635	5111	85
	1000	(X,X,3,2,2)	(3,4,5)	0.8365	0.1635	5111	56
	1000	(2,2,2,6,2)	(3,4,5)	0.8308	0.1692	2805	90
	500	(X,X,3,2,2)	(3,4,5)	0.8365	0.1635	5111	29
	500	(X,X,2,6,2)	(3,4,5)	0.8308	0.1692	2805	30
Solución óptima sugerida: Partición = (X,X,3,2,2); Q=0.8365							
$MSE_{entr}$	# eval.	Gran.	Opt. Mask	Q	$MSE_{entr}$	$ECM_{prue}$	Tiempo
	2000	(X,X,4,9,7)	(3,4,5)	0.5704	0.1625	3305	150
	2000	(X,X,4,9,4)	(3,4,5)	0.6417	0.1667	3379	155
	2000	(5,3,7,9,7)	(3,4,5)	0.5742	0.1696	3572	159
	1000	(X,X,4,9,7)	(3,4,5)	0.5704	0.1625	3305	76
	1000	(X,X,4,9,8)	(3,4,5)	0.5452	0.1648	3375	80
	1000	(7,6,7,9,7)	(3,4,5)	0.5742	0.1696	3572	84
	1000	(X,X,2,2,2)	(3,4,5)	0.8010	0.1923	7851	85
	500	(X,X,4,9,7)	(3,4,5)	0.5704	0.1625	3305	39
	500	(X,X,4,9,8)	(3,4,5)	0.5452	0.1648	3375	39
	500	(X,X,4,9,9)	(3,4,5)	0.5433	0.1651	3342	41
	500	(6,9,4,9,5)	(3,4,5)	0.6224	0.1675	3349	41
	500	(7,8,7,9,9)	(3,4,5)	0.5484	0.1721	3614	43
	500	(7,5,5,2,7)	(3,4,5)	0.5273	0.2005	7236	40
Solución óptima sugerida: Partición = (X,X,4,9,7); $MSE_{entr}$ =0.1625							

Las figuras 24 y 25 muestran las señales de predicción para el conjunto de datos de prueba de las soluciones óptima y mejor subóptima de ambas funciones objetivo.



**Figura 24.** Mejores predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento. Función objetivo  $1-Q$ .  $ECM_{prue}=5111$  y  $ECM_{prue}=2805$ , respectivamente. Método AGI.



**Figura 25.** Mejores predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento. Función objetivo  $MSE_{entr}$ .  $ECM_{prue}=3305$  y  $ECM_{prue}=3375$ , respectivamente. Método AGI.

### 6.2.3.2 Aprendizaje de las funciones de pertenencia de las clases (AG2)

Este apartado muestra los resultados obtenidos por el AG cuyo fin es la optimización de las funciones de pertenencia (**AG2**). En esta aplicación, este AG considera el número óptimo de clases obtenido por el **AG1** mostrado previamente, por lo tanto, se tiene nuevamente la combinación **AG1+AG2**. Asimismo, se señala que el espacio de búsqueda se vuelve aún más grande de acuerdo a las características del problema bajo consideración.

**Tabla 18.** Resultados de las funciones de pertenencia para el problema de estimación de costos óptimos de mantenimiento usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1).

Núm. Clases: (2,2,3,2,2)							
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	ECM <sub>prue</sub>	
A	20000	LTC:(0.85,0.15)	(1,3,4,5)	0.9475	0.0525	2727	
		AOC:(0.49,0.51)					
		AOE: (0.79,0.13,0.08)					
	10000	ESC:(0.81,0.19)	(1,3,4,5)	0.9505	0.0495	2728	
		CMT:(0.85,0.15)					
		LTC:(0.92,0.08)					
	5000	AOC:(0.58,0.42)	(1,3,4,5)	0.9386	0.0614	2721	
		AOE: (0.82,0.10,0.08)					
		ESC:(0.85,0.15)					
	B	20000	CMT:(0.85,0.15)	(1,3,4,5)	0.8886	0.1114	5094
			LTC:(0.75,0.25)				
			AOC:(0.63,0.37)				
10000		AOE: (0.66,0.12,0.22)	(1,2,4,5)	0.8844	0.1156	94458	
		ESC:(0.78,0.22)					
		CMT:(0.73,0.27)					
5000		LTC:(0.52,0.48)	(1,3,4,5)	0.8723	0.1277	5063	
		AOC:(0.51,0.49)					
		AOE: (0.35,0.24,0.41)					
		ESC:(0.46,0.54)	(1,3,4,5)	0.8723	0.1277	5063	
		CMT:(0.49,0.51)					
		LTC:(0.63,0.37)					
	AOC:(0.60,0.40)	(1,3,4,5)	0.8723	0.1277	5063		
	AOE: (0.43,0.26,0.31)						
	ESC:(0.59,0.41)						
	CMT:(0.63,0.37)	(1,3,4,5)	0.8723	0.1277	5063		
	LTC:(0.51,0.49)						
	AOC:(0.46,0.54)						
	AOE: (0.31,0.26,0.43)	(1,3,4,5)	0.8723	0.1277	5063		
	ESC:(0.53,0.47)						
	CMT:(0.58,0.42)						

La *Tabla 18* presenta los resultados obtenidos por el **AG2** para el problema de estimación de costos de mantenimiento cuando se usa la función objetivo 1-Q. La tabla se organiza como se describe a continuación. La primera columna se divide en dos secciones. La Sección *A* corresponde a los mejores resultados obtenidos por el **AG2** mientras que la Sección *B* corresponde a sus peores resultados. La segunda columna indica el número de evaluaciones de cromosomas realizadas por el **AG2**. La

tercera columna muestra la proporción de datos para las variables de entrada (LTC, AOC, AOE, ESC) y la variable de salida (CMT). El número de elementos de la proporción de datos va de acuerdo a la granularidad de cada variable. La proporción de datos es la salida del **AG2** y el segundo parámetro de entrada del proceso de fusificación de la metodología FIR. La cuarta columna presenta la máscara óptima, en notación de posición, encontrada por FIR cuando se emplea esa proporción de datos para el cómputo de los landmarks. La quinta columna corresponde a la calidad asociada a la máscara óptima. La sexta columna contiene el valor de la función objetivo 1-Q. La última columna muestra el error de predicción ECM (Véase ecuación 48) del conjunto de datos de prueba.

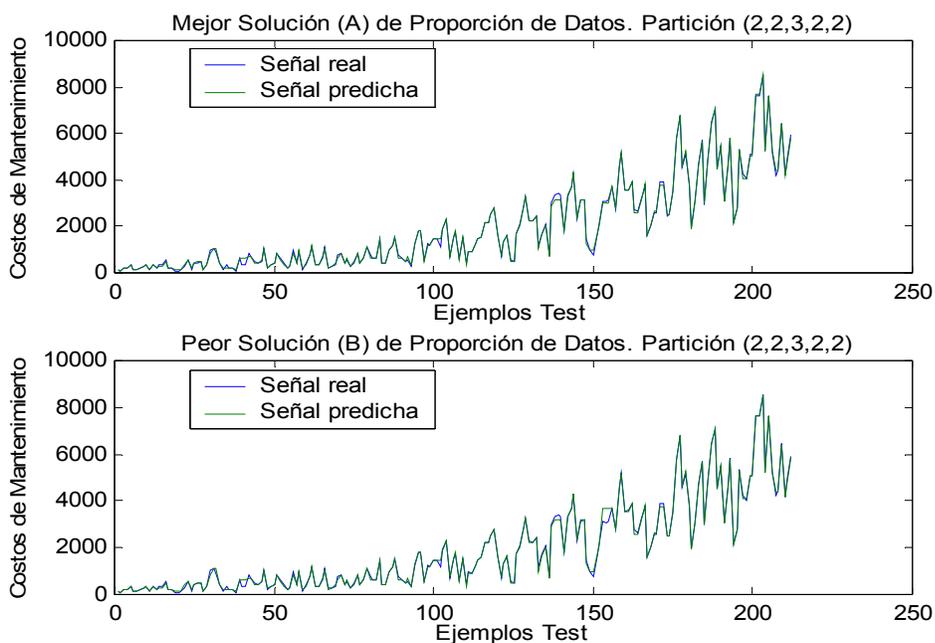
**Tabla 19.** Resultados de las funciones de pertenencia para el problema de estimación de costos óptimos de mantenimiento usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1).

Núm. Clases: (2,2,4,9,7)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$ECM_{prue}$
A	20000	LTC:(0.52,0.48)	(3,4,5)	0.5245	0.1228	2973
		AOC:(0.59,0.41)				
		AOE (0.19,0.33,0.29,0.19)				
	10000	ESC:(0.11,0.12,0.12,0.12,0.12,0.14,0.08,0.08,0.11)	(3,4,5)	0.5343	0.1246	3028
		CMT:(0.17,0.07,0.11,0.19,0.08,0.25,0.13)				
		LTC:(0.56,0.44)				
	5000	AOC:(0.49,0.51)	(3,4,5)	0.5212	0.1272	3016
		AOE:(0.19,0.35,0.28,0.18)				
		ESC:(0.11,0.11,0.14,0.12,0.11,0.14,0.09,0.07,0.11)				
	20000	CMT:(0.09,0.08,0.09,0.09,0.26,0.31,0.08)	(3,4,5)	0.5741	0.1306	3005
		LTC:(0.59,0.41)				
		AOC:(0.42,0.58)				
B	10000	AOE (0.20,0.34,0.26,0.20)	(3,4,5)	0.5465	0.1332	2985
		ESC:(0.10,0.16,0.15,0.09,0.09,0.14,0.09,0.07,0.11)				
		CMT:(0.13,0.10,0.12,0.40,0.10,0.08,0.07)				
	5000	LTC:(0.40, 0.60)	(3,4,5)	0.5124	0.1331	2975
		AOC:(0.55,0.45)				
		AOE (0.27,0.27,0.23,0.23)				
	10000	ESC:(0.13,0.11,0.11,0.08, 0.10,0.13,0.15,0.14,0.05)	(3,4,5)	0.5465	0.1332	2985
		CMT:(0.12,0.11, 0.20,0.12,0.21,0.05,0.19)				
		LTC:(0.52,0.48)				
	5000	AOC:(0.58,0.42)	(3,4,5)	0.5124	0.1331	2975
		AOE (0.34,0.26,0.22,0.18)				
		ESC:(0.10,0.11,0.12,0.11,0.18,0.14, 0.10,0.05,0.09)				
	10000	CMT:(0.19,0.24,0.17,0.12,0.15,0.08,0.05)	(3,4,5)	0.5124	0.1331	2975
		LTC:(0.40,0.60)				
		AOC:(0.64,0.36)				
	5000	AOE (0.32,0.33,0.16,0.19)	(3,4,5)	0.5124	0.1331	2975
		ESC:(0.15,0.13,0.12,0.11,0.12,0.13,0.08,0.07,0.09)				
		CMT:(0.15,0.13,0.16,0.14, 0.10,0.15,0.17)				

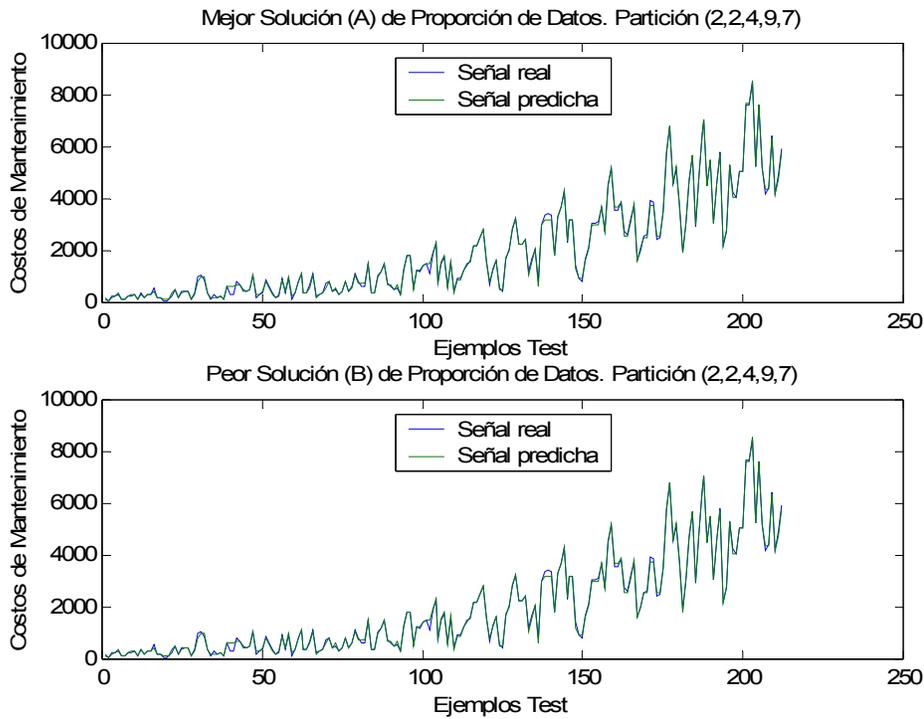
Por otro lado, la *Tabla 19* muestra los resultados obtenidos por éste AG usando la función objetivo  $MSE_{entr}$ . La tabla se organiza como la *Tabla 18*. La única diferencia es que la sexta columna contiene los valores de la función objetivo  $MSE_{entr}$  en vez de 1-Q.

Como se ha señalado antes, el tiempo de CPU necesario por el **AG2** cuando se utiliza la función objetivo  $MSE_{entr}$  es mayor que el consumido cuando se emplea la función objetivo 1-Q. Por ejemplo, el **AG2** para alcanzar las 30 ejecuciones usando un criterio de parada de 20000 evaluaciones necesitó de 6:36 horas y 11:39 horas para la funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente.

Las *figuras 26 y 27* muestran el comportamiento de las predicciones de los modelos FIR para las funciones objetivo 1-Q y  $MSE_{entr}$ , correspondientes a la mejor solución (A) y a la peor solución (B) encontradas por el **AG2**, usando como criterio de parada 20000 evaluaciones de cromosomas.



**Figura 26.** Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidad (2,2,3,2,2). Mejor y peor solución de proporción de datos (20000 evaluaciones). Función objetivo 1-Q.  $ECM_{prue}=2727$  y  $ECM_{prue}=5094$ , respectivamente. Método AG2.



**Figura 27.** Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidad (2,2,4,9,7). Mejor y peor solución de proporción de datos (20000 evaluaciones). Función objetivo  $MSE_{entr}$ .  $ECM_{prue}=2973$  y  $ECM_{prue}=3005$ , respectivamente. Método AG2.

### 6.2.3.3 Aprendizaje global del número óptimo de clases y de las funciones de pertenencia de las clases (AG3)

En esta sección se presentan los resultados obtenidos por el **AG3** propuesto para el aprendizaje conjunto de los parámetros de fusificación de FIR. El número de evaluaciones de cromosomas se incrementa de manera considerable, ya que el espacio de búsqueda es notable.

La *Tabla 20* muestra los resultados obtenidos para la estimación de los costos de mantenimiento de la línea de media tensión usando la función objetivo 1-Q. La *Tabla 22* muestra los resultados obtenidos cuando se utiliza la función objetivo del error de predicción del 20% final del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). Ambas tablas están organizadas como se describe a continuación. La primera columna se divide en cinco secciones. Las secciones *A* y *B* corresponden a los mejores y a los peores resultados obtenidos por el **AG3**, respectivamente. La sección

*C* muestra el resultado cuando se utilizan los parámetros de fusificación por defecto de FIR (**FIR DEFECTO**), es decir, todas las variables se discretizan en tres clases y se emplea el método EFP para obtener funciones de pertenencia uniformes. Las secciones *D* y *E* muestran los mejores resultados cuando se emplean los enfoques **AG1** y **AG2**. El primero se encarga de obtener un nivel óptimo de granularidad (número de clases) para cada variable y el segundo tiene como objetivo la determinación de las funciones de pertenencia de las clases. El **AG1** usa el método EFP para determinar las funciones de pertenencia para cada variable. El **AG2** sugiere funciones de pertenencia no uniformes manteniendo siempre invariable la granularidad dada por el **AG1** (combinación **AG1+AG2**). *C*, *D* y *E* se usan más adelante como referencia en el análisis de los resultados. La segunda columna indica el número de evaluaciones de cromosomas realizadas por el **AG3**. La tercera y cuarta columnas presentan la solución sugerida por el **AG3** (granularidad y proporción de datos) para las variables de entrada (LTC, AOC, AOE, ESC) y la variable de salida (CMT) cuando se cumple el criterio de parada estipulado. El número de elementos de la proporción de datos va de acuerdo a la granularidad (etiquetas) de cada variable. Nótese que la solución sugerida representa los parámetros de entrada al proceso de fusificación de la metodología FIR. La quinta columna contiene la máscara óptima (en notación de posición) encontrada por FIR cuando se utiliza esa granularidad y se emplea esa proporción de datos para el cómputo de los landmarks. La sexta columna corresponde a la calidad asociada a la máscara óptima. La séptima columna es el valor de la función objetivo utilizada, 1-Q ó  $MSE_{entr}$ . La última columna muestra el error de predicción ECM para el conjunto de datos de prueba.

**Tabla 21.** Tiempo consumido por el AG3 para realizar las treinta ejecuciones para cada criterio de parada. Funciones objetivo 1-Q y  $MSE_{entr}$ .

Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas)	
	1-Q	$MSE_{entr}$
5000	2:01	3:35
10000	4:06	7:11
20000	8:43	15:38
40000	16:17	32:32
80000	32:06	61:00
160000	61:16	132:17
320000	124:01	308:24

El número de evaluaciones de cromosomas hechas por el AG condiciona obviamente el tiempo de CPU necesario para realizar las treinta ejecuciones. En la *Tabla 21* se muestran los tiempos para las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente.

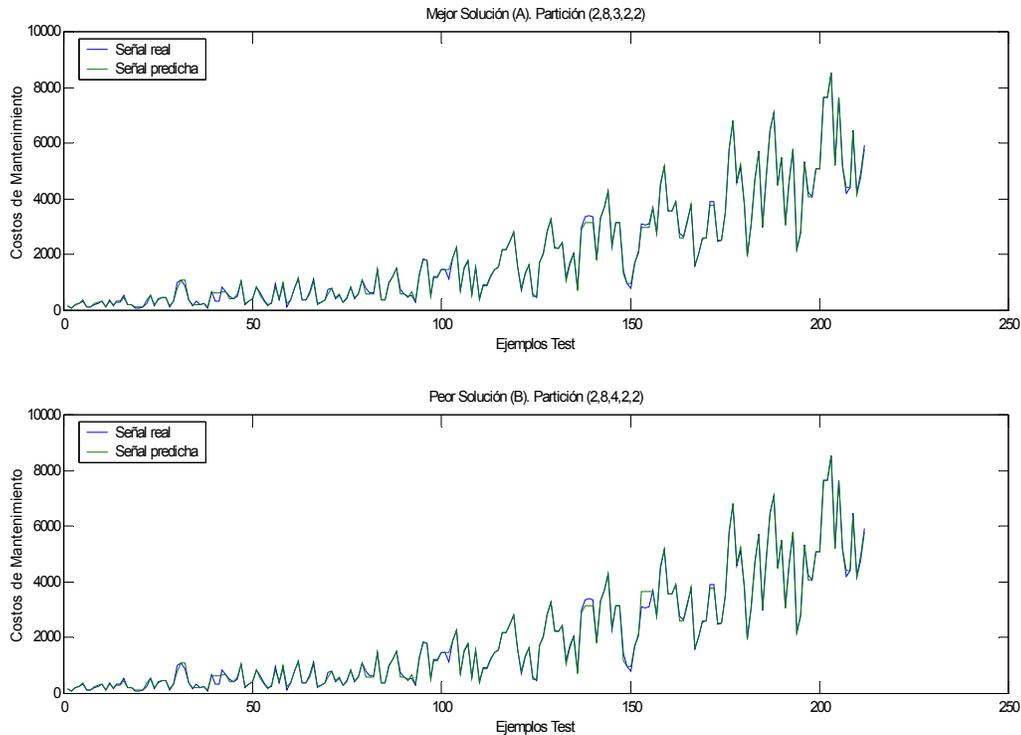
**Tabla 20.** Resultados de la granularidad y las funciones de pertenencia para el problema de estimación de costos de mantenimiento de la línea de media tensión usando la función objetivo 1-Q.

	# eval	Gran.	Proporción de datos	Opt. Mask	Q	1-Q	ECM <sub>prue</sub>
A	320000	(2,8,3,2,2)	LTC:(0.91,0.09) AOC:(0.10,0.10,0.20,0.10,0.12,0.14,0.12) AOE: (0.88,0.07,0.05) ; ESC:(0.88,0.12) CMT:(0.92,0.08)	(1,3,4,5)	0.9716	0.0284	2728
	160000	(2,7,2,2,2)	LTC:(0.91,0.09) AOC:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) AOE: (0.94,0.06) ; ESC:(0.93,0.07) CMT:(0.92,0.08)	(1,3,4,5)	0.9638	0.0362	2728
	80000	(7,6,6,3,2)	LTC:(0.12,0.12,0.12,0.12,0.13,0.25,0.14) AOC:(0.13,0.13,0.14,0.14,0.33,0.13) AOE: (0.73,0.06,0.06,0.05,0.05,0.05) ESC:(0.83,0.09,0.08) ; CMT:(0.90,0.10)	(3,4,5)	0.9777	0.0223	2759
	40000	(2,5,3,2,2)	LTC:(0.91,0.09) AOC:(0.48,0.12,0.13,0.13,0.14) AOE: (0.88,0.06,0.06) ; ESC:(0.93,0.07) CMT:(0.92,0.08)	(1,3,4,5)	0.9721	0.0279	2729
	20000	(8,3,6,2,2)	LTC:(0.11,0.11,0.11,0.11,0.12,0.18,0.13,0.13) AOC:(0.34,0.33,0.33) AOE: (0.74,0.06,0.05,0.05,0.05,0.05) ESC:(0.92,0.08) ; CMT:(0.90,0.10)	(3,4,5)	0.9559	0.0441	2759
	10000	(7,4,3,3,2)	LTC:(0.15,0.13,0.13,0.15,0.13,0.17,0.14) AOC:(0.25,0.25,0.25,0.25) AOE:(0.48,0.20,0.32) ; ESC:(0.48,0.19,0.33) CMT:(0.54,0.46)	(3,4,5)	0.9503	0.0497	5159
	5000	(2,5,5,2,2)	LTC:(0.29,0.71) AOC:(0.20,0.20,0.20,0.20,0.20) AOE:(0.62,0.17,0.11,0.05,0.05) ; ESC:(0.84,0.16) CMT:(0.90,0.10)	(3,4,5)	0.9423	0.0577	2759
B	320000	(2,8,4,2,2)	LTC: (0.53,0.47) AOC: (0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) AOE: (0.38,0.09,0.13,0.40) ; ESC: (0.47,0.53) CMT: (0.54,0.46)	(1,3,4,5)	0.9031	0.0969	5132
	160000	(9,8,6,2,2)	LTC: (0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) AOC: (0.11,0.13,0.11,0.11,0.12,0.12,0.18,0.12) AOE: (0.11,0.14,0.19,0.13,0.12,0.31) ESC: (0.46,0.54) ; CMT: (0.53,0.47)	(3,4,5)	0.8934	0.1066	5136
	80000	(2,8,2,2,2)	LTC:(0.53,0.47) AOC:(0.14,0.10,0.16,0.10,0.10,0.10,0.20) AOE: (0.49,0.51) ; ESC:(0.47,0.53) CMT:(0.54,0.46)	(1,3,4,5)	0.8979	0.1021	5125
	40000	(4,5,5,2,2)	LTC:(0.08,0.69,0.14,0.09) AOC:(0.33,0.13,0.13,0.3,0.11) AOE: (0.15,0.19,0.21,0.17,0.28) ESC:(0.47,0.53) ; CMT:(0.48,0.52)	(3,4,5)	0.8821	0.1179	5119
	20000	(6,7,8,8,2)	LTC:(0.13,0.13,0.14,0.14,0.15,0.31) AOC:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) AOE: (0.08,0.38,0.29,0.05,0.05,0.05,0.05,0.05) ESC:(0.15,0.12,0.12,0.12,0.15,0.13,0.10,0.11) CMT:(0.90,0.10)	(3,5)	0.8848	0.1152	225848
	10000	(5,6,8,7,3)	LTC:(0.25,0.23,0.06,0.23,0.23) AOC:(0.16,0.16,0.17,0.17,0.17,0.17) AOE:(0.07,0.14,0.16,0.33,0.08,0.08,0.07,0.07) ESC:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) CMT:(0.87,0.05,0.08)	(3,5)	0.8548	0.1452	210251
	5000	(6,9,9,2,2)	LTC:(0.16,0.16,0.17,0.17,0.17,0.17) AOC:(0.10,0.10,0.10,0.10,0.12,0.11,0.10,0.17,0.10) AOE:(0.10,0.36,0.10,0.09,0.05,0.09,0.08,0.06,0.07) ESC:(0.50,0.50) ; CMT:(0.91,0.09)	(3,5)	0.8764	0.1236	216862
C		(3,3,3,3,3)	Equal Frequency Partition method	(3,4,5)	0.6116	0.3884	5134
D		(2,2,3,2,2)	Equal Frequency Partition method	(3,4,5)	0.8365	0.1635	5111
E		(2,2,3,2,2)	LTC:(0.92,0.08) ; AOC:(0.58,0.42) AOE: (0.82,0.10,0.08) ; ESC:(0.85,0.15) CMT:(0.85,0.15)	(1,3,4,5)	0.9505	0.0495	2728

**Tabla 22.** Resultados de la granularidad y las funciones de pertenencia para el problema de estimación de costos de mantenimiento de la línea de media tensión usando la función objetivo  $MSE_{entr}$ .

	# eval	Gran.	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$ECM_{pntc}$
A	320000	(2,2,6,7,8)	LTC:(0.10,0.90) ; AOC:(0.50,0.50) AOE: (0.05,0.47,0.19,0.07,0.12,0.10) ESC: (0.21,0.10,0.28,0.15,0.10,0.09,0.07) CMT: (0.12,0.06,0.05,0.12,0.19,0.08,0.22,0.16)	(1,3,4,5)	0.4404	0.1078	2880
	160000	(2,4,5,7,8)	LTC:(0.10,0.90) ; AOC:(0.21,0.24,0.27,0.28) AOE: (0.19,0.33,0.31,0.09,0.08) ESC: (0.17,0.20,0.16,0.17,0.11,0.14,0.05) CMT: (0.13,0.11,0.11,0.12,0.15,0.09,0.06,0.23)	(1,3,4,5)	0.4779	0.1090	2936
	80000	(2,7,5,5,8)	LTC:(0.10,0.90) ; AOC:(0.11,0.11,0.11,0.17,0.12,0.12,0.26) AOE: (0.12,0.4,0.24,0.10,0.14) ESC:(0.25,0.21,0.25,0.22,0.07) CMT:(0.13,0.11,0.11,0.12,0.19,0.11,0.13,0.10)	(1,3,4,5)	0.4953	0.1116	8252
	40000	(2,8,7,9,9)	LTC:(0.08,0.92) AOC:(0.12,0.12,0.12,0.13,0.13,0.13,0.12) AOE: (0.29,0.12,0.12,0.12,0.12,0.11,0.12) ESC:(0.14,0.07,0.08,0.13,0.08,0.20,0.09,0.16,0.05) CMT:(0.08,0.11,0.05,0.06,0.07,0.08,0.14,0.18,0.23)	(1,3,4,5)	0.5080	0.1152	3022
	20000	(2,7,7,8,9)	LTC:(0.10,0.90) ; AOC:(0.11,0.27,0.11,0.12,0.12,0.13,0.14) AOE: (0.12,0.12,0.30,0.11,0.11,0.12,0.12) ESC:(0.11,0.11,0.19,0.11,0.15,0.11,0.11,0.11) CMT:(0.09,0.09,0.05,0.12,0.15,0.10,0.10,0.19,0.11)	(1,3,4,5)	0.5217	0.1165	3066
	10000	(2,3,5,8,7)	LTC:(0.57,0.43) ; AOC:(0.20,0.51,0.29) AOE: (0.18,0.18,0.18,0.27,0.19) ESC: (0.18,0.10,0.10,0.10,0.11,0.17,0.12,0.12) CMT:(0.08,0.08,0.08,0.11,0.28,0.21,0.16)	(3,4,5)	0.5748	0.1258	3025
	5000	(2,2,7,9,9)	LTC:(0.10,0.90) ; AOC:(0.50,0.50) AOE: (0.32,0.10,0.12,0.07,0.17,0.12,0.10) ESC:(0.12,0.22,0.06,0.09,0.12,0.12,0.11,0.05,0.11) CMT:(0.12,0.05,0.15,0.06,0.07,0.18,0.13,0.18,0.06)	(1,3,4,5)	0.4999	0.1180	2976
B	320000	(3,2,3,5,9)	LTC:(0.32,0.36,0.32) ; AOC:(0.59,0.41) AOE: (0.19,0.64,0.17) ESC:(0.25,0.19,0.15,0.30,0.11) CMT:(0.12,0.06,0.07,0.05,0.13,0.19,0.10,0.12,0.16)	(3,4,5)	0.4096	0.1329	2970
	160000	(5,7,7,6,7)	LTC:(0.17,0.22,0.24,0.17,0.20) AOC:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) AOE: (0.10,0.08,0.11,0.12,0.33,0.13,0.13) ESC: (0.09,0.27,0.09,0.21,0.27,0.07) CMT: (0.06,0.07,0.30,0.33,0.06,0.10,0.08)	(3,4,5)	0.5357	0.1341	2939
	80000	(4,5,4,9,8)	LTC:(0.25,0.25,0.25,0.25) ; AOC:(0.20,0.20,0.20,0.20,0.20) AOE: (0.18,0.07,0.55,0.20) ESC:(0.09,0.16,0.09,0.09,0.09,0.09,0.18,0.12) CMT:(0.11,0.11,0.18,0.05,0.12,0.12,0.13,0.18)	(3,4,5)	0.4470	0.1359	2963
	40000	(4,2,4,6,8)	LTC:(0.16,0.33,0.24,0.27) ; AOC:(0.54,0.46) AOE: (0.39,0.16,0.27,0.18) ESC:(0.06,0.16,0.16,0.36,0.21,0.05) CMT:(0.05,0.11,0.07,0.12,0.09,0.32,0.05,0.19)	(3,4,5)	0.4646	0.1353	3007
	20000	(4,3,7,9,7)	LTC:(0.25,0.25,0.25,0.25) ; AOC:(0.36,0.32,0.32) AOE: (0.14,0.14,0.14,0.14,0.14,0.15,0.15) ESC:(0.06,0.11,0.13,0.15,0.07,0.19,0.14,0.07,0.08) CMT:(0.18,0.05,0.09,0.11,0.32,0.13,0.12)	(3,4,5)	0.5390	0.1361	2978
	10000	(4,4,4,3,3)	LTC:(0.11,0.53,0.11,0.25) ; AOC:(0.25,0.25,0.25,0.25) AOE: (0.13,0.17,0.35,0.35) ; ESC:(0.45,0.49,0.06) CMT:(0.15,0.28,0.57)	(3,4,5)	0.5566	0.1390	3053
	5000	(8,7,2,2,2)	LTC:(0.11,0.12,0.12,0.12,0.14,0.14,0.13,0.12) AOC:(0.11,0.11,0.12,0.12,0.28,0.13,0.13) AOE: (0.52,0.48) ; ESC:(0.50,0.50) CMT:(0.35,0.65)	(3,4,5)	0.6616	0.1847	7896
C		(3,3,3,3,3)	Equal Frequency Partition method	(3,4,5)	0.6116	---	5134
D		(2,2,4,9,7)	Equal Frequency Partition method	(3,4,5)	0.5711	0.1653	3259
E		(2,2,4,9,7)	LTC:(0.52,0.48) ; AOC:(0.59,0.41) AOE: (0.19,0.33,0.29,0.19) ESC:(0.11,0.12,0.12,0.12,0.14,0.08,0.08,0.11) CMT:(0.17,0.07,0.11,0.19,0.08,0.25,0.13)	(3,4,5)	0.5245	0.1228	2973

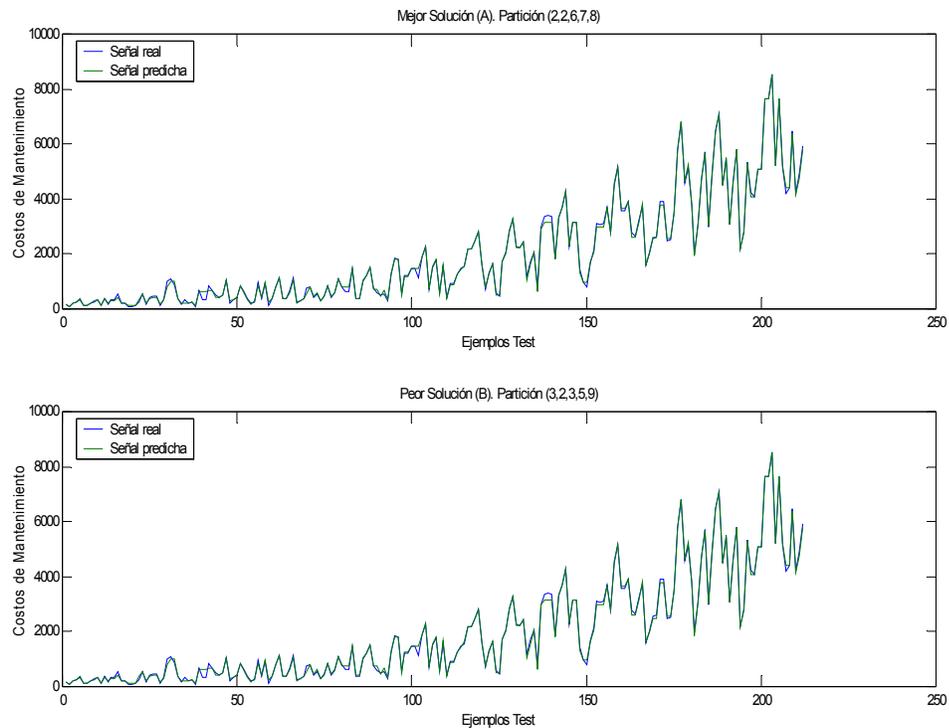
Las *figuras 28 y 29* presentan el comportamiento de las predicciones de los modelos FIR para las funciones objetivo 1-Q y  $MSE_{entr}$ , correspondientes a la mejor solución (A) y a la peor solución (B) encontradas por el **AG3**, usando como criterio de parada 320000 evaluaciones de cromosomas.



**Figura 28.** Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidades (2,8,3,2,2) y (2,8,4,2,2). Mejor y peor solución (320000 evaluaciones). Función objetivo 1-Q.  $ECM_{prue}=2728$  y  $ECM_{prue}=5132$ , respectivamente. Método AG3.

#### 6.2.4 Análisis de resultados y conclusiones

Esta sección comprende un análisis de los resultados obtenidos por los métodos **AG1**, **AG2** y **AG3** propuestos en esta memoria para la resolución del problema de modelado de los costos óptimos de mantenimiento de la línea de media tensión en núcleos urbanos. Las dos funciones objetivo establecidas son evaluadas y comparadas desde la perspectiva de su eficiencia y tiempo computacional. Por último, se dan algunas conclusiones.



**Figura 29.** Predicciones usando los modelos FIR para el problema estimación de costos óptimos de mantenimiento con granularidades  $(2,2,6,7,8)$  y  $(3,2,3,5,9)$ . Mejor y peor solución (320000 evaluaciones). Función objetivo  $MSE_{entr}$ .  $ECM_{prue}=2880$  y  $ECM_{prue}=2970$ , respectivamente. Método AG3.

#### 6.2.4.1 Método AG1

De la *Tabla 17* se desprende que cuando se alcanzan 2000 evaluaciones, el AG sugiere una única partición para la función objetivo 1-Q y tres particiones diferentes para la función objetivo  $MSE_{entr}$ . Cuando se usan 1000 evaluaciones como criterio de parada, el AG sugiere dos y cuatro particiones diferentes para las funciones de costo 1-Q y  $MSE_{entr}$ , respectivamente. Mientras que para el último criterio de parada el AG sugiere dos y seis particiones diferentes, respectivamente.

Es interesante destacar que en todas las particiones sugeridas por el AG, las primeras dos variables, suma de la longitud de todas las calles de la ciudad (LTC) y área total ocupada por la ciudad (AOC), no son consideradas relevantes por FIR (no forman parte de la máscara óptima encontrada) para la determinación del costo de mantenimiento de la línea de media tensión (CMT). FIR encuentra que no existen

relaciones causales importantes entre LTC y AOC con respecto a la variable de salida CMT.

Por otra parte, se puede observar que los mejores modelos FIR obtenidos por el **AG1** (soluciones óptimas) y **FIR DEFECTO** (Véase fila *C* de la *Tabla 20*) tienen mayor eficiencia tanto en la predicción del conjunto de entrenamiento como de prueba que los obtenidos por métodos clásicos y híbridos evolutivos (Véase *Tabla 15*). El  $ECM_{prue}$  más bajo logrado en trabajos previos es 11836, mientras que los más altos obtenidos por el **AG1** son de 1094 y 5111, respectivamente. Con **FIR DEFECTO** se obtiene un  $ECM_{prue}$  de 5134, error un poco más alto que los peores logrados con el AG1. La reducción de los errores es muy significativa y tiene una implicación económica enorme. FIR con el **AG1** es capaz de obtener modelos más fiables y precisos para la estimación de los costos de mantenimiento de la línea de media tensión.

Es importante analizar los resultados obtenidos usando el **AG1** por las dos funciones objetivo: 1-Q y  $MSE_{entr}$ . En la *Tabla 17* se puede advertir que cuando se usa la función objetivo  $MSE_{entr}$  los valores de  $ECM_{prue}$  son en general menores a excepción de dos soluciones subóptimas sugeridas por el AG que convergieron en una zona no prometedora con un  $MSE_{entr}$  de alrededor de 0.20% ( $ECM_{prue}$  de 7236 y 7851, respectivamente). El  $ECM_{prue}=5111$ , fue el más alto obtenido por la función objetivo 1-Q y correspondió a la solución óptima. El tiempo de CPU para la función objetivo  $MSE_{entr}$  se incrementa con respecto a la función objetivo 1-Q, debido a que se requiere ejecutar el proceso de predicción de FIR para la evaluación de cada partición. No obstante, este incremento no es realmente significativo porque en esta aplicación las máscaras óptimas encontradas por las particiones sugeridas son muy simples (la m-salida sólo depende de dos m-entradas), y por ello, el proceso de predicción es muy rápido. Nótese que en este caso se requiere aproximadamente el doble del tiempo necesario para computar la función objetivo 1-Q, mientras que en la aplicación anterior el tiempo de cómputo de la función  $MSE_{entr}$  era 7 veces superior al de la función 1-Q.

Los valores absolutos de los errores ECM son altos debido al rango [0...10000] de la variable costo de mantenimiento y a que la fórmula del ECM no está normalizada por la varianza. Sin embargo, en las *figuras 24 y 25* puede verse, para un criterio de parada de 2000 evaluaciones de cromosomas, que las predicciones de los modelos FIR identificados tanto para la función objetivo 1-Q como para  $MSE_{entr}$ , siguen la señal real del costo de mantenimiento de una manera muy precisa.

En la *figura 24*, la solución óptima sugerida por el AG1: (2,2,3,2,2), presentó un  $ECM_{prue}$  de 5111 y la solución subóptima: (2,2,2,6,2), un  $ECM_{prue}$  de 2805. En este caso, la solución subóptima estimó más idóneamente los costos de mantenimiento de

la línea de media tensión. Sin embargo, sucedió lo contrario en la *figura 25*. Se puede indicar que la solución óptima sugerida por el **AG1**: (2,2,4,9,7) presentó un  $ECM_{prue}$  de 3305, y la mejor solución subóptima: (2,2,4,9,8), un  $ECM_{prue}$  de 3379. En todo caso, como se ve en ambas figuras, la predicción de los modelos FIR sigue extraordinariamente bien la señal real.

#### 6.2.4.2 Método AG2

Aunque el **AG2** enlazado con el **AG1** (**AG1+AG2**) no asegura la solución óptima, todas las soluciones (filas *A* y *B*) de las *Tablas 18* y *19*, son preferibles a la solución óptima obtenida únicamente con el **AG1**. En este caso, y para la función objetivo 1-Q, FIR encuentra que existe relación causal entre la variable de entrada: suma de la Longitud de Todas las Calles de la ciudad (LTC) y la variable de salida Costo de ManTenimiento (CMT), cosa que no sucedía cuando se usaba el **AG1** ni cuando se usa el **AG2** con la función objetivo  $MSE_{entr}$ .

Por otra parte, si analizamos las últimas columnas de ambas tablas, una vez más se puede observar que el  $ECM_{prue}$  obtenidos por la implementación de AGs en el contexto de la metodología FIR (**AG1+AG2**), son significativamente inferiores a los obtenidos por las otras metodologías (Véase *Tabla 15*) y **FIR DEFECTO** (Véase sección *C* de la *Tabla 20*). El mejor resultado del  $ECM_{prue}$  conseguido por modelos difusos TSK (11836) es mucho más alto que los  $ECM_{prue}$  logrados (Véase *Tabla 18*) cuando se implementa el **AG2** en la metodología FIR (2721 y 2727).

Como era de esperar, los resultados obtenidos por FIR cuando se utiliza la combinación **AG1+AG2** son mejores a los registrados cuando sólo se computa el **AG1** ó se utiliza el **FIR DEFECTO**. De las *Tablas 18* y *19* se puede observar que los resultados obtenidos por **AG1 + AG2** en ambas funciones objetivo son equivalentes. En este caso, la eficiencia de los modelos FIR cuando se usa la función objetivo  $MSE_{entr}$  no es superior a la eficiencia conseguida con la función objetivo 1-Q. Por lo tanto en este caso la función objetivo 1-Q es preferible por su menor costo computacional.

En la *figura 26* se presentan las predicciones de los modelos FIR para la partición (2,2,3,2,2) correspondientes a la mejor solución (*A*) y a la peor solución (*B*) encontradas con la función objetivo 1-Q por el **AG2** (criterio de parada 20000 evaluaciones). Los errores  $ECM_{prue}$  para ambas soluciones fueron 2727 y 5094, respectivamente. En la *figura 27* se muestran las predicciones de los modelos FIR para la partición (2,2,4,9,7), correspondientes a la mejor solución (*A*) y a la peor solución (*B*) encontradas por el **AG2** usando la función objetivo  $MSE_{entr}$  (criterio de parada 20000 evaluaciones). Los  $ECM_{prue}$  para ambas soluciones fueron 2973 y 3005, respectivamente. Se observa que en ambas funciones objetivo, los modelos FIR

siguen de forma idónea el comportamiento de la señal real del costo de mantenimiento de la línea de media tensión en núcleos urbanos.

#### 6.2.4.3 Método AG3

Analicemos los resultados de los tres métodos desarrollados desde la perspectiva de la calidad de predicción de los modelos FIR identificados. La *Tabla 23* resume este análisis.

**Tabla 23.**  $ECM_{prue}$  de los mejores resultados obtenidos (desde la perspectiva de las funciones objetivo) usando: AG3, AG1+AG2, AG1+EFP, FIR DEFECTO y otros métodos (clásicos + híbridos evolutivos).

	<b>AG3</b>	<b>AG1+AG2</b>	<b>AG1+EFP</b>	<b>FIR DEFECTO</b>	<b>OTROS METODOS</b>
MSE <sub>entr</sub>	2880	2973	3305	5134	11836
1-Q	2759	2728	5111	5134	11836

Se deduce de la *Tabla 23* que el poder de predicción de los modelos FIR identificados que han hecho uso de los tres SDEs propuestos en esta memoria es superior al logrado por el modelo FIR que usa valores por defecto para los parámetros del proceso de discretización de FIR. Es más, los resultados obtenidos por los tres métodos desarrollados son muy superiores a los logrados por las 9 técnicas diferentes presentadas en la literatura (Véase *Tabla 15*), correspondiendo la mayor parte de ellas a diferentes tipos de modelos difusos. Cabe señalar que una mejora de un orden de magnitud en el error de predicción supone una implicación económica importante para las compañías eléctricas.

De los tres SDEs propuestos, el **AG3** y el **AG1+AG2** presentan resultados equivalentes (Véase *Tablas 18, 19, 20 y 22*). Cabe destacar que en este caso los mejores resultados desde el punto de vista de la predicción se obtienen usando la función objetivo 1-Q, aunque la diferencia es realmente mínima. El **AG1** queda desmarcado de los demás AG propuestos a pesar que sigue siendo una buena alternativa si lo comparamos con **FIR DEFECTO** y los otros métodos previamente publicados.

Es muy interesante prestar atención a las máscaras identificadas por el **AG3**. Al igual que pasaba con el **AG1+AG2** y el **AG1**, la mayoría de las máscaras identifican las variables Área Ocupada por Edificios (AOE) y Energía Suministrada a la Ciudad (ESC) (posiciones 3 y 4 de la máscara) como las más relevantes para predecir el comportamiento de la variable de salida CMT. En algunas ocasiones, principalmente en el **AG3** y **AG1+AG2**, la máscara identifica también encuentra relevante la variable de entrada LTC (suma de la Longitud de Todas las Calles de la ciudad). Esto enfatiza de nuevo la importancia de determinar buenos parámetros de discretización para

lograr a obtener el mejor modelo que represente el sistema estudiado.

En la *figura 28* se presentan las predicciones de los modelos FIR correspondientes a la mejor solución (*A*) y a la peor solución (*B*) encontradas por el **AG3** con la función objetivo 1-Q y usando como criterio de parada 320000 evaluaciones de cromosomas. Los errores ECM para ambas soluciones fueron 2728 y 5132, respectivamente. Por su parte, la *figura 29* muestra las predicciones de los modelos FIR para la función objetivo  $MSE_{entr}$ , en este caso los errores ECM fueron de 2880 y 2970, respectivamente. Aquí también se observó que los modelos FIR para ambas funciones objetivo siguieron de forma precisa el comportamiento de la señal real del costo de mantenimiento de la línea de media tensión.

Para finalizar se puede indicar que en esta aplicación la combinación **AG1+AG2** presenta resultados equivalentes a los obtenidos por el **AG3**. Por otro lado, ambas funciones objetivo presentan resultados muy similares, siendo en este caso recomendable usar 1-Q debido a su bajo coste computacional con respecto al de  $MSE_{entr}$ .

## 6.3 Caso de estudio: Concentraciones de ozono en Austria

### 6.3.1 Descripción del problema

La contaminación del aire representa uno de los problemas medioambientales de especial atención en las áreas urbanas y su origen es obra del hombre como resultado de sus actividades comerciales, domésticas e industriales, aunque también puede ser resultado de causas naturales (emanaciones volcánicas, radiactividad natural). El proceso contaminante más significativo es el proceso de combustión, el cual se utiliza para la obtención de calor, energía eléctrica o movimiento.

El aire puro es una mezcla de gases compuesta por un 78% de nitrógeno, un 21% de oxígeno, con proporciones que permanecen más o menos constantes; y un 1% de diferentes componentes como gases nobles (argón, neón, helio, kriptón y xenón), dióxido de carbono ( $CO_2$ ) y ozono ( $O_3$ ). Se entiende por contaminación atmosférica la presencia en el aire de sustancias y formas de energía que alteran la calidad del mismo, de modo que implique riesgos, daño o molestia grave para las personas y bienes de cualquier naturaleza. Es por ello que el estudio acerca de cualquier cambio en el equilibrio de los componentes del aire que produzca una modificación de sus propiedades físicas y químicas, resulta de gran interés científico. La contaminación del aire provocada por el aumento de los gases de efecto invernadero, principalmente el  $CO_2$ , contribuye a largo plazo al calentamiento del planeta. Además del  $CO_2$ ,

contaminan también la atmósfera los óxidos de nitrógeno ( $\text{NO}_x$ ) y el monóxido de carbono (CO), los hidrocarburos no quemados, el benceno ( $\text{C}_6\text{H}_6$ ) de los derivados oxigenados de los hidrocarburos, y el plomo procedente de los antidetonantes de la gasolina.

Los principales contaminantes del aire se clasifican en primarios y secundarios. Los primarios son los que permanecen en la atmósfera tal y como fueron emitidos por la fuente. Para fines de la evaluación de la calidad de aire se consideran: óxidos de azufre ( $\text{SO}_x$ ), CO,  $\text{NO}_x$ , hidrocarburos y partículas. Los secundarios son los que han estado sujetos a cambios químicos, o bien, son el producto de la reacción de dos o más contaminantes primarios en la atmósfera. Entre ellos destacan los oxidantes fotoquímicos y algunos radicales de corta existencia como el ozono.

El ozono llamado también *smog fotoquímico*, se forma en la baja atmósfera mediante la reacción química de compuestos volátiles orgánicos (VOC) y bióxido de nitrógeno ( $\text{NO}_2$ ) en presencia de la luz solar, es por ello que las temperaturas elevadas estimulan la reacción fotoquímica. El ozono es un gas compuesto de tres átomos de oxígeno y es el resultado de la cantidad de combustible que consumen los vehículos, industrias y servicios de cualquier centro urbano. La creación del ozono depende de la cantidad y tipo de fuentes emisoras, la localización geográfica del área y de los asentamientos humanos y fabriles. El ozono produce efectos tóxicos porque fácilmente quita electrones a otras moléculas, iniciando reacciones en cadena y perturbando estructuras vitales en las células ocasionando daños que afectan la salud de la población en general, tales como dolores de cabeza, problemas respiratorios, irritaciones en los ojos, nariz y garganta, así como el empeoramiento de las enfermedades cardiovasculares.

Para poder advertir a tiempo y de manera adecuada, es importante tener previsiones precisas y fiables de los niveles altos futuros de ozono. Por consiguiente, la construcción de modelos de ozono que capturen de la manera más precisa posible el comportamiento de este gas en la atmósfera resulta ser el principal interés no sólo de los científicos medioambientales sino de las agencias gubernamentales.

Hay diferentes modelos disponibles para la predicción de la calidad del aire a escala local y para la previsión de niveles de ozono. Algunos de estos usan métodos clásicos basados en algoritmos numéricos y aproximaciones estadísticas [Loi96, SMPA06, WLWL03, YB06]. Otros usan conocimiento químico/físico [SWKK96]. Recientemente se están usando para el mismo propósito otros paradigmas tales como las redes neuronales [WW99, LHP+04, YB06, Ghi05], árboles de decisión o reglas de asociación [WW01] y sistemas basados en lógica difusa [LHP+04, Ghi05, HK04], entre los que se encuentra la metodología FIR [GNMW01,GNRA+03, NM06].

En la región Oriental Austriaca, específicamente en Viena, se han detectado altos niveles de contaminación del aire por lo cual se hace necesario obtener modelos para estimar de la mejor forma posible las concentraciones locales máximas de ozono con miras a alertar a la población y tomar los correctivos que sean necesarios. La aplicación que se presenta en esta sección fue publicada inicialmente por los hermanos Wotawa en [WW99]. Con el objetivo de poder comparar la bondad de nuestros métodos con los propuestos en este estudio previo, se decidió usar exactamente los mismos datos y la misma función de error. Las mediciones se registraron principalmente durante la estación de verano. Los datos proporcionados involucran tres variables atmosféricas obtenidas mediante predicción y los niveles de concentraciones de ozono. La variable de salida es la serie de tiempo del ozono ( $O_3$ ), medida en ppb (partes por mil millones o por billones) y calculada como una media aritmética de cinco puntos de medición que proporcionan valores promedios de tres horas fijas de un día. Por lo tanto se dispone de un registro de ozono por día. Las variables de entrada son: a) la temperatura (TMP), medida en °K; b) la cobertura de nubes o nivel de nubosidad (NUB), que puede tomar valores entre el rango [0...1] donde un valor de 0 corresponde a la situación en que no hay nubes (despejado) y un valor de 1 significa que el cielo está completamente cubierto (nublado); y c) la velocidad del viento (VV), medida en m/s. Estos datos provienen del *European Center for Medium Range Weather Forecasts* [ECMWF95].

Se dispone de datos del ozono y variables atmosféricas para los períodos: 07-07-1995 a 25-09-1995 (81 valores), y 01-05-1996 a 30-09-1996 (149 valores). El objetivo de esta aplicación es la identificación de modelos que permitan predecir la concentración de ozono del día siguiente al actual. Por lo tanto, el objetivo es la predicción a corto plazo del contaminante ozono en la zona este de Austria. Para la identificación y validación de los modelos FIR usando los métodos genéticos propuestos en esta memoria, los datos disponibles fueron divididos en conjunto de identificación (entrenamiento) y de prueba. Siguiendo el trabajo de [WW99] todos los datos del año 1996 se usan para la identificación del modelo y los datos del año 1995 se usan para la prueba de éste, constituyéndose de este modo la *Partición Inicial*. Las *figuras 30 y 31* muestran las señales de entrenamiento de las variables de entrada (TMP, NUB y VV) y de la variable de salida  $O_3$  para esta distribución de datos.

Estudios previos [GNRA+03] señalan que la reducida cantidad de datos disponibles constituye una gran limitación para las metodologías inductivas como lo es FIR. Es por ello, que se aplicó la técnica *cross-validation* con la cual se pretendió extraer la máxima información posible teniendo en cuenta el escaso número de mediciones disponibles. De este modo se consideraron, adicionalmente a la *Partición Inicial*, dos particiones/distribuciones diferentes de los datos:

- *Caso 1*: El conjunto de identificación (entrenamiento) está compuesto por todos los datos de 1995 más los primeros 68 valores del conjunto de datos de

1996. El conjunto de prueba contiene los últimos 81 valores del conjunto de datos de 1996.

- *Caso 2:* El conjunto de identificación (entrenamiento) está compuesto por todos los datos de 1995 junto con los últimos 68 valores del conjunto de datos de 1996. El conjunto de prueba contiene los primeros 81 valores del conjunto de datos de 1996.

En ambos conjuntos de identificación, los datos que corresponden a diferentes años se concatenan adicionando un intervalo de valores perdidos (*missing values*) predefinidos, para evitar la creación de relaciones causales físicamente inexistentes. La metodología FIR es capaz de tratar con elementos perdidos en los procesos de identificación y simulación cualitativa.

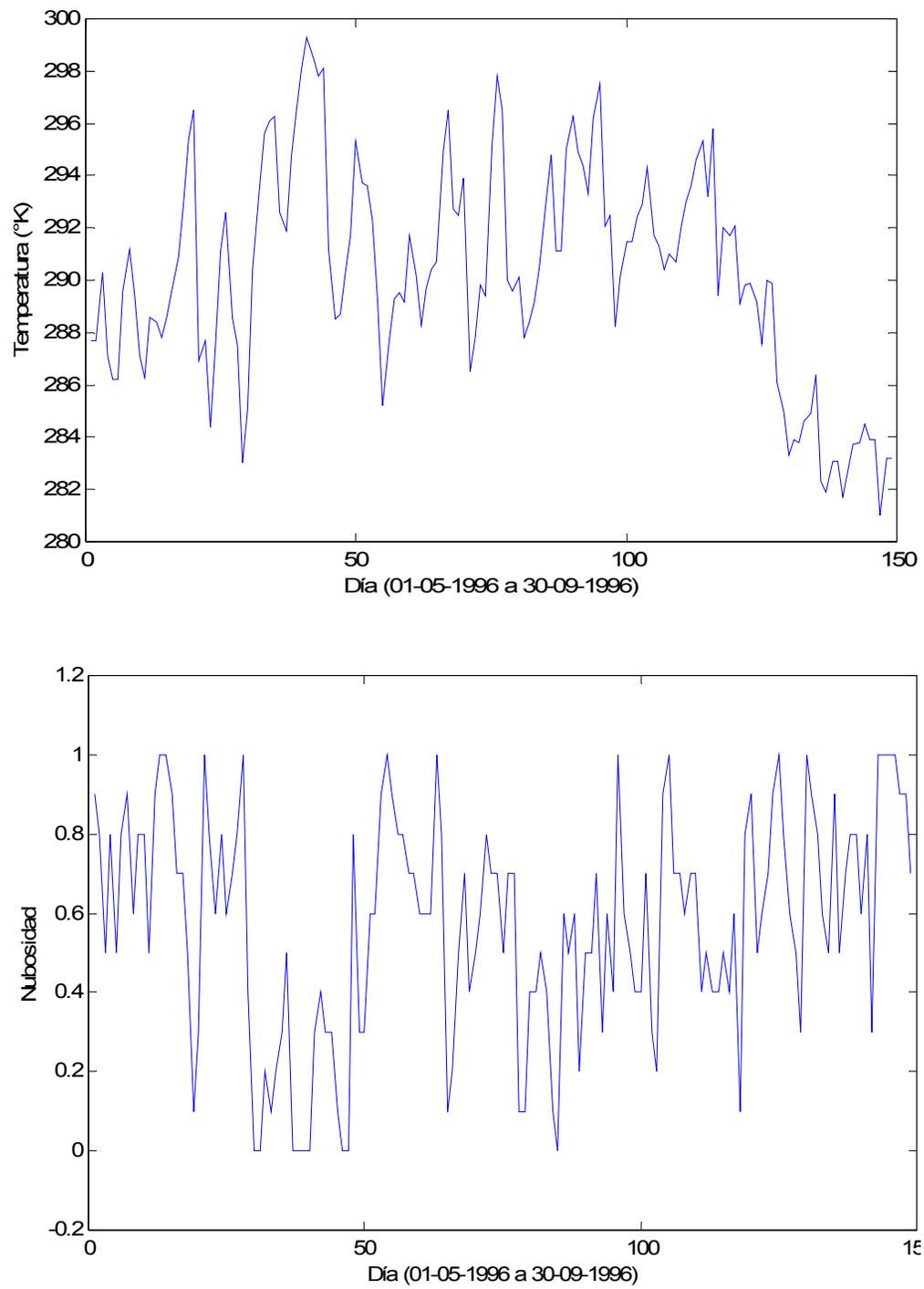
### 6.3.2 Trabajos previos

En [WW99] se realizó una investigación con redes neuronales de tipo perceptrónicas multicapas (MLP), *Elman Networks* (EN) y *Modified Elman Networks* (MEN). Las MLPs son las conocidas redes de alimentación adelantada (topología feedforward); las ENs son una clase de redes neuronales parcialmente recurrentes; y por último, las MENs son una extensión de las ENs en la cual los estados de las neuronas son concatenados con sí mismos, pero obteniendo cierta inercia que incrementa las capacidades de la red para la memorización dinámica de los datos. Todos los modelos usaron neuronas con activación sigmoideal y función de salida lineal. La función de entrada calculaba la suma de los pesos de las entradas. El algoritmo *Back-propagation* (Propagación hacia atrás o Retropropagación) se empleó para el entrenamiento de las redes.

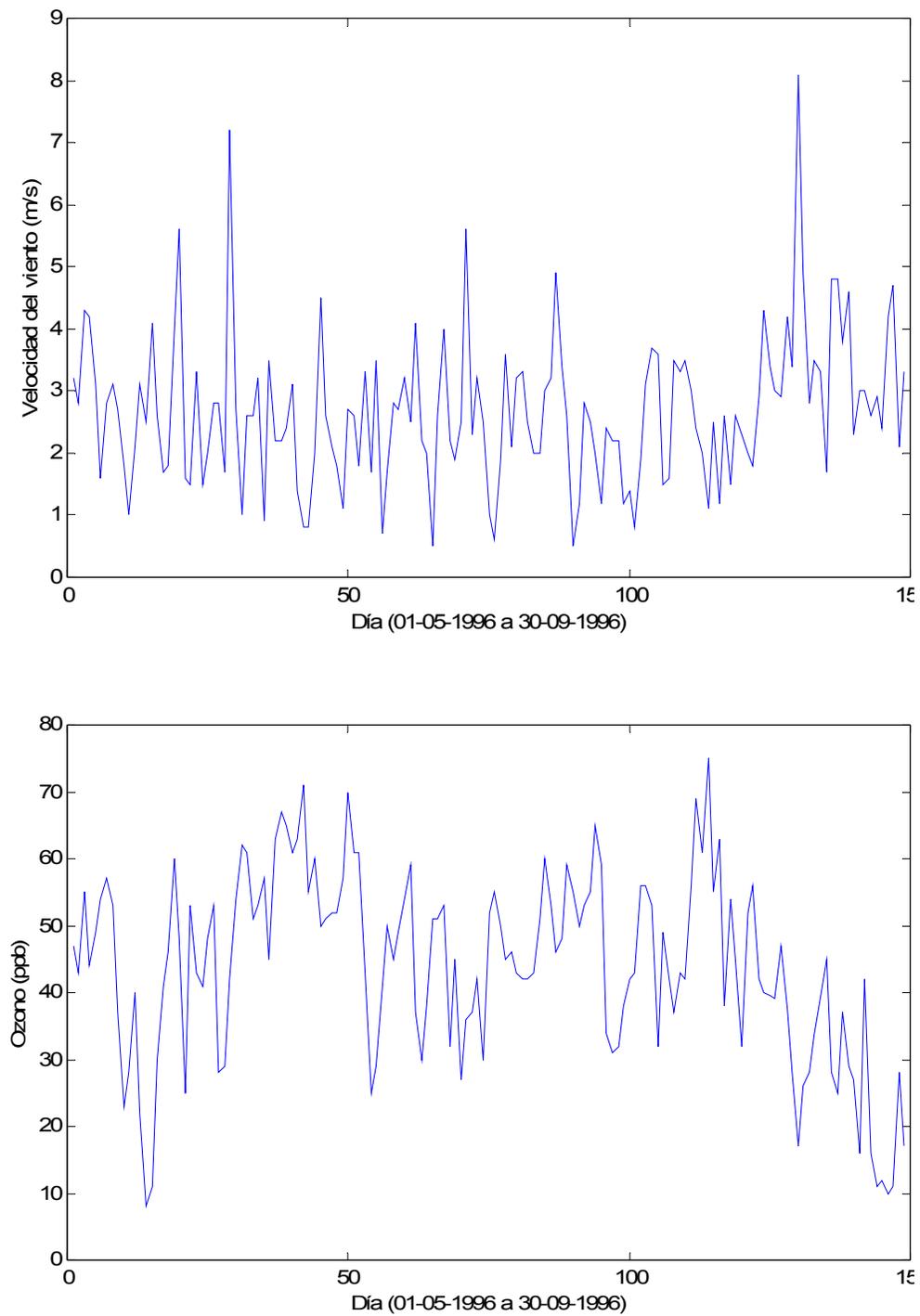
Estos modelos neuronales fueron identificados y validados con la distribución de datos denominada *Partición Inicial* que ha sido descrita previamente. Además, los resultados obtenidos con estos trabajos se compararon con los resultados logrados en otras investigaciones previas en donde se aplicaron métodos clásicos tales como algoritmos numéricos y aproximaciones estadísticas [Loi96], modelos IMPO basados en conocimiento químico/físico [SWKK96] y modelos Santiago [AJP96], mostrando que las redes neuronales tenían mayor eficiencia en la predicción. Para ello se empleó como medida de error la raíz del error cuadrático medio, *Root Mean Square Error* (RMS), descrita en la ecuación 49.

$$RMS = \sqrt{\frac{\sum_{i=1}^N (y_i(t) - \hat{y}_i(t))^2}{N}} \quad (49)$$

Donde:  $y(t)$  es la salida del sistema,  $\hat{y}(t)$  es la salida predicha y  $N$  denota el número de observaciones.



**Figura 30.** Señales de entrenamiento para las variables de entrada TMP y NUB. **Partición Inicial.**



**Figura 31.** Señales de entrenamiento para la variable de entrada  $VV$  y la variable de salida  $O_3$ .  
**Partición Inicial.**

**Tabla 24.** Modelos neuronales obtenidos por [WW99] para la predicción de concentraciones locales máximas de ozono en Austria. **Partición Inicial.**

Red Neuronal	#HU	LR	Bias	Steps	RMS <sub>prue</sub>
MLP	2	1	N	3000	15.0537
MLP	5+1	1	S	1000	11.2004
MLP	5	0.4	N	100	11.1768
MLP	5+1	0.2	S	100	10.8132
MLP	2+1	0.2	S	1000	10.6651
MEN ( $\alpha=0.2$ )	5+1	0.2	S	3000	10.5150
EN	8+1	0.2	S	5000	10.3186
MEN ( $\alpha=0.4$ )	5+1	0.2	S	5000	9.9579

La *Tabla 24* contiene los mejores modelos neuronales obtenidos por [WW99] para el problema de estimación de las concentraciones locales máximas de ozono a corto plazo de la región oriental austriaca. La tabla está organizada como se indica a continuación. La primera columna indica el tipo de red neuronal. La segunda y tercera columnas contienen el número de unidades ocultas y la tasa de aprendizaje, respectivamente. La cuarta columna señala si se ha usado o no el parámetro bias. La penúltima columna muestra el número necesario de ciclos de aprendizaje de la RN para obtener esos resultados. La última columna corresponde al valor del error de predicción del conjunto de datos de prueba (RMS<sub>prue</sub>) que viene dado por la ecuación 49.

Por otro lado, el trabajo realizado en [GNRA+03] profundizó en el mismo problema, pero aplicando la metodología FIR. En este estudio se consideraron tanto las máscaras óptimas como las subóptimas para obtener la predicción más precisa posible. Se trabajó con las tres distribuciones de datos ya descritas: *Partición Inicial*, *Caso 1* y *Caso 2*. Los parámetros de discretización de FIR se determinaron de la siguiente manera. En relación al número de clases por variable (granularidad), se realizaron diferentes pruebas de manera “manual” usando desde 2 hasta 5 clases. Debido a la reducida cantidad de datos disponibles, el número de clases (etiquetas) debía ser pequeño para poder inferir modelos con un buen nivel de predictividad. Con respecto a las funciones de pertenencia se usaron el algoritmo FCM (*Fuzzy C-Means*) y el EFP (*Equal Frequency Partition*) [WC87, BEF84]. Los mejores resultados se obtuvieron con el método EFP y discretizando la variable temperatura en 3 clases mientras que el resto de las variables fueron discretizadas en 2 clases. Se decidió fusificar la temperatura en tres clases porque FIR encontró que ésta variable tiene una alta relación causal con la concentración de ozono. Por lo tanto, una discretización más cuidada de la variable temperatura podría producir modelos más precisos.

**Tabla 25.** Máscaras FIR para la predicción de concentraciones locales máximas de ozono usando la granularidad (3,2,2,2) buscada de manera manual. Funciones de pertenencia uniformes (Método EFP). **Partición Inicial.**

Mask type	Mask relations	Q	RMS <sub>prue</sub>
Óptima	(7,12,14,15,16)	0.4306	13.6768
Subóptima	(8,12,13,16)	0.3680	10.5508
Subóptima	(1,4,13,16)	0.3158	10.5383
Subóptima	(4,13,14,16)	0.4033	10.1837
Subóptima	(7,12,13,16)	0.3456	10.1110
Subóptima	(1,12,13,16)	0.3566	10.0220
Subóptima	(9,12,13,16)	0.2830	9.8270
Subóptima	(8,9,13,16)	0.2866	9.7577

En la *Tabla 25* se presentan la máscara óptima y las mejores máscaras subóptimas encontradas para la *Partición Inicial* usando la granularidad (3,2,2,2) y el método EFP para distribuir uniformemente las funciones de pertenencia. La primera columna indica si la máscara identificada es la óptima o es una subóptima. La segunda columna describe las relaciones causales de cada máscara (en notación de posición). La tercera columna corresponde a la calidad asociada a cada máscara. La última columna muestra el error de predicción del conjunto de datos de prueba, RMS<sub>prue</sub>.

**Tabla 26.** Máscaras FIR para la predicción de concentraciones locales máximas de ozono usando la granularidad (3,2,2,2) buscada de manera manual. Funciones de pertenencia uniformes (Método EFP). **Particiones: Caso 1 y Caso 2.**

	Mask relations	RMS <sub>prue</sub>
<i>Caso 1</i>	(8,11,12,13,16)	9.9764
	(12,13,14,16)	9.9736
	(11,12,13,16)	9.9383
	(8,12,13,14,16)	9.3165
<i>Caso 2</i>	(5,12,13,14,16)	10.1063
	(9,12,13,16)	9.8964
	(1,12,13,14,16)	9.5026
	(1,12,13,16)	9.3343

Las mejores máscaras subóptimas FIR obtenidas para las distribuciones de datos *Caso 1* y *Caso 2* se muestran en la *Tabla 26*. La tabla está organizada como se describe a continuación. La primera columna indica la distribución de los datos (*Caso 1* o *Caso 2*). La segunda columna muestra las relaciones causales de cada máscara (en notación de posición). La última columna indica el valor del error de predicción de los datos de prueba, RMS<sub>prue</sub>.

Asimismo, los autores en [GNRA+03] llevaron a cabo estudios adicionales para la *Partición Inicial* usando redes MENs con funciones de activación tangente hiperbólica y un tipo de redes neuronales recurrentes llamadas *Long Short Term Memory* (LSTM) [GSC00, HS97]. La *Tabla 27* muestra los mejores modelos neuronales obtenidos con estos métodos.

**Tabla 27.** Modelos neuronales obtenidos en [GNRA+03] para la predicción de concentraciones locales máximas de ozono. **Partición Inicial.**

Red Neuronal	#HU	LR	Bias	Steps	RMS <sub>prue</sub>
MEN-HYP ( $\alpha=0.2$ )	5+1	0.2	S	1000	9.1930
LSTM	-----	0.00125	-----	3500	9.7961

### 6.3.3 Modelos obtenidos por los métodos *AG1*, *AG2* y *AG3*

Al igual que en las aplicaciones anteriores, en este problema de estimación de concentraciones máximas de ozono en Austria, se han realizado 30 ejecuciones (con distinto valor de semilla) para cada función objetivo y criterio de parada en cada método genético planteado en esta memoria. Se trabaja con tres distribuciones de datos: Partición Inicial, Caso 1 y Caso 2. Igualmente, se estudian dos funciones objetivo diferentes: a) 1-Q, y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento (MSE<sub>entr</sub>). En esta aplicación medioambiental se usa el último 8% de la señal de entrenamiento para la evaluación de la función objetivo MSE<sub>entr</sub> y el 92% inicial de la señal para la obtención de los modelos FIR (máscaras y base de reglas patrón), puesto que el tamaño del conjunto de datos disponible es inferior a 400 instancias.

**Tabla 28.** Parámetros de los AGs para el problema de estimación de las concentraciones locales máximas de ozono en Austria.

Parámetro	Tipo AG	Valor
<i>Tamaño de la población</i> (# individuos)	AG1	16
	AG2, AG3	40
<i>Probabilidad de cruce</i>	Todos	0.6
<i>Probabilidad de mutación</i>	Todos	0.1
<i>Criterio de parada</i> (# evaluaciones de cromosomas)	AG1	{200, 400, 800}
	AG2	{500, 1000, 2000, 4000}
	AG3	{500, 1000, 2000, 4000, 8000, 16000}

En la ecuación 50 se presenta la máscara candidata de profundidad 4 (cuatro filas) utilizada, la cual permite encontrar las relaciones causales y temporales entre las variables de la aplicación bajo estudio. Los elementos -1 de la máscara candidata

representan las relaciones causales y temporales potenciales con la variable de salida (elemento +1).

$t \times$	TMP	NUB	VV	03
$t - 3\delta t$	-1	-1	-1	-1
$t - 2\delta t$	-1	-1	-1	-1
$t - \delta t$	-1	-1	-1	-1
$t$	-1	-1	-1	+1

(50)

Los valores establecidos para los parámetros de los AGs se presentan en la *Tabla 28*. Las probabilidades de cruce y de mutación son escogidas de acuerdo a [Gre86]. Se emplearon computadoras Pentium IV (2.66 GHz) para realizar todas las simulaciones. En [ANF06] se muestran todos los resultados obtenidos para los métodos **MANUAL+AG2**, **AG1**, **AG1+AG2** y **AG3**. Los resultados obtenidos por los métodos propuestos se presentan en las *Tablas 29 a 52*.

El resto de la sección se dedica a mostrar los resultados logrados por las diferentes combinaciones de métodos propuestos organizados de la siguiente manera:

1. Granularidad (3,2,2,2) y funciones de pertenencia sugeridas por el **AG2**.
2. Granularidad sugerida por el **AG1** y funciones de pertenencia obtenidas con el método EFP.
3. Granularidad sugerida por el **AG1** y funciones de pertenencia sugeridas por el **AG2**.
4. Granularidad y funciones de pertenencia sugeridas por el **AG3**.

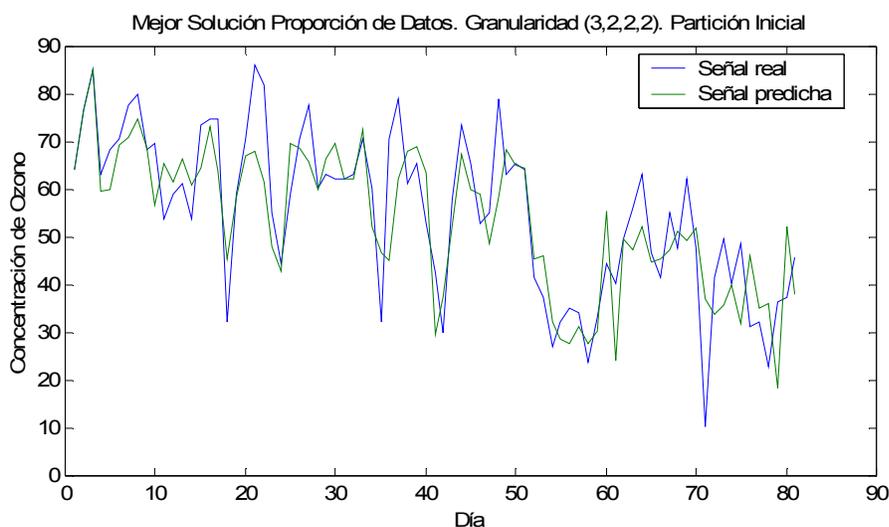
#### 6.3.3.1 Aprendizaje de las funciones de pertenencia de las clases (AG2). Granularidad (3,2,2,2) suministrada en estudios previos

En ésta parte se presentan los resultados obtenidos para el problema de estimación de concentraciones de ozono en Austria por el método genético propuesto para la determinación automática de las funciones de pertenencia de la metodología FIR (**AG2**). Se considera la granularidad (3,2,2,2) encontrada de manera manual en [GNRA+03], por lo tanto, se tiene la combinación **MANUAL+AG2**.

Las *Tablas 29, 30 y 31* presentan los resultados del **AG2** para la granularidad (3,2,2,2) cuando se usa la función objetivo 1-Q y se utilizan las distribuciones de datos: Partición Inicial, Caso 1 y Caso 2, respectivamente. Las tablas están organizadas como se especifica a continuación. La primera columna se divide en tres ó cuatro secciones. La sección *A* corresponde a los mejores resultados de la función objetivo obtenidos por la combinación **MANUAL+AG2**. La sección *B* corresponde a los mejores resultados de predicción sobre el conjunto de prueba obtenidos por **MANUAL+AG2**. La sección *C* y *D* muestran los mejores resultados de predicción

obtenidos por FIR en [GNRA+03] cuando se usa el método EFP para determinar la proporción de datos para cada variable y se consideran máscaras óptimas y/o subóptimas (Véanse *Tablas 25 y 26*). De esta forma,  $C$  y  $D$  se usan como referencia para el análisis de los resultados. La segunda columna indica el criterio de parada, es decir, el número de evaluaciones de cromosomas realizadas por el AG. La tercera columna muestra la proporción de datos para las variables de entrada (TMP,NUB,VV) y la variable de salida ( $O_3$ ). El número de elementos de la proporción de datos corresponde al número de clases para cada variable (en este caso (3,2,2,2)). La proporción de datos es la salida del **AG2** y el segundo parámetro de entrada del proceso de fusificación de la metodología FIR. La cuarta columna presenta la máscara óptima (en notación de posición) encontrada por FIR cuando la proporción de datos obtenida es usada como conjunto de landmarks. La quinta columna corresponde a la calidad asociada a la máscara óptima. La sexta columna es el valor de la función objetivo 1-Q. La última columna muestra el valor del error  $RMS_{prue}$ .

Las *figuras 32, 33 y 34* muestran como los mejores modelos obtenidos por el método **MANUAL+AG2** usando la función objetivo 1-Q siguen la señal real del sistema bajo estudio para cada distribución de datos considerada: Partición Inicial, Caso 1 y Caso 2, respectivamente.



**Figura 32.** Predicción del mejor modelo FIR para la **Partición Inicial**. Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (3,2,2,2) con  $Q=0,9707$  y distribución de landmarks: TMP:(0.23,0.46, 0.31)-NUB:(0.72,0.28)-VV:(0.35,0.65)- $O_3$ :(0.94,0.06).  $RMS_{prue}=9.6817$ . Método MANUAL+AG2.

**Tabla 29.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable obtenida manualmente en [GNRA+03]). **Partición Inicial.**

Núm. Clases: (3,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	4000	TMP:(0.58,0.36,0.06) NUB:(0.40,0.60) VVI: (0.51,0.49) O3:(0.94,0.06)	(2,6,13,16)	0.9799	0.0201	22.0337
	2000	TMP:(0.61,0.32,0.07) NUB:(0.40,0.60) VVI: (0.52,0.48) O3:(0.94,0.06)	(5,6,13,16)	0.9802	0.0198	27.5986
	1000	TMP:(0.18,0.15,0.67) NUB:(0.43,0.57) VVI: (0.63,0.37) O3:(0.94,0.06)	(2,6,14,16)	0.9783	0.0217	14.5029
	500	TMP:(0.37,0.27,0.36) NUB:(0.66,0.34) VVI: (0.50,0.50) O3:(0.94,0.06)	(3,5,7,16)	0.9779	0.0221	17.1370
B	4000	TMP:(0.35,0.36,0.29) NUB:(0.61,0.39) VVI: (0.61,0.39) O3:(0.94,0.06)	(3,13,16)	0.9791	0.0209	10.3597
	4000	TMP:(0.37,0.34,0.29) NUB:(0.41,0.59) VVI: (0.59,0.41) O3:(0.94,0.06)	(3,13,16)	0.9785	0.0215	10.5577
	2000	TMP:(0.40,0.16,0.44) NUB:(0.57,0.43) VVI: (0.52,0.48) O3:(0.92,0.08)	(1,13,16)	0.9456	0.0544	10.4012
	2000	TMP:(0.39,0.17,0.44) NUB:(0.56,0.44) VVI: (0.71,0.29) O3:(0.92,0.08)	(1,13,16)	0.9455	0.0545	10.5162
	1000	TMP:(0.23,0.46,0.31) NUB:(0.72,0.28) VVI: (0.35,0.65) O3:(0.94,0.06)	(7,13,16)	0.9707	0.0293	9.6817
	1000	TMP:(0.29,0.34,0.37) NUB:(0.45,0.55) VVI: (0.51,0.49) O3:(0.86,0.14)	(3,11,13,16)	0.8638	0.1362	9.7561
	500	TMP:(0.32,0.30,0.38) NUB:(0.76,0.24) VVI:(0.50,0.50) O3:(0.84,0.16)	(3,11,13,16)	0.8566	0.1434	10.0128
	500	TMP:(0.62,0.15,0.23) NUB:(0.90,0.10) VVI:(0.35,0.65) O3:(0.90,0.10)	(3,13,16)	0.8943	0.1057	10.3072
C		Método EFP con Máscara óptima	(7,12,14,15,16)	0.4306	0.5694	13.6768
D		Método EFP con Máscara subóptima	(8,9,13,16)	0.2866	0.7134	9.7577

**Tabla 30.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable obtenida manualmente en [GNRA+03]). **Partición: Caso 1.**

Núm. Clases: (3,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	4000	TMP:(0.40,0.22,0.38) NUB:(0.81,0.19) VV: (0.50, 0.50) O3:(0.05,0.95)	(3,6,14,16)	0.9428	0.0572	13.0978
	2000	TMP:(0.37,0.29,0.34) NUB:(0.87,0.13) VV: (0.46,0.54) O3:(0.05,0.95)	(1,3,14,16)	0.9467	0.0533	12.6310
	1000	TMP:(0.26,0.17,0.57) NUB:(0.84,0.16) VV: (0.46,0.54) O3:(0.05,0.95)	(3,6,14,16)	0.9428	0.0572	13.3636
	500	TMP:(0.39,0.28,0.33) NUB:(0.62,0.38) VV: (0.48,0.52) O3:(0.05,0.95)	(6,11,14,15,16)	0.9267	0.0733	12.3104
B	4000	TMP:(0.28,0.29,0.43) NUB:(0.39,0.61) VV: (0.60,0.40) O3:(0.05,0.95)	(11,13,15,16)	0.9191	0.0809	8.7900
	4000	TMP:(0.32,0.24,0.44) NUB:(0.14,0.86) VV: (0.51,0.49) O3:(0.05,0.95)	(11,13,15,16)	0.9105	0.0895	8.9092
	2000	TMP:(0.32,0.40,0.28) NUB:(0.59,0.41) VV: (0.44,0.56) O3:(0.07,0.93)	(10,13,14,16)	0.8634	0.1366	9.0613
	2000	TMP:(0.32,0.39,0.29) NUB:(0.61,0.39) VV: (0.34,0.66) O3:(0.09,0.91)	(10,13,14,16)	0.8227	0.1773	9.1468
	1000	TMP:(0.31,0.39,0.30) NUB:(0.46,0.54) VV: (0.53,0.47) O3:(0.05,0.95)	(11,13,15,16)	0.9247	0.0753	8.1732
	1000	TMP:(0.32,0.29,0.39) NUB:(0.58,0.42) VV: (0.37,0.63) O3:(0.07,0.93)	(10,13,14,16)	0.8660	0.1340	9.3578
	500	TMP:(0.63,0.24,0.13) NUB:(0.43,0.57) VV: (0.52,0.48) O3:(0.89,0.11)	(11,13,15,16)	0.7886	0.2114	8.7201
	500	TMP:(0.31,0.34,0.35) NUB:(0.57,0.43) VV: (0.57,0.43) O3:(0.10,0.90)	(2,11,13,16)	0.8076	0.1924	9.2707
C		Método EFP con Máscara Subóptima	(8,12,13,14,16)	----	----	9.3165

**Tabla 31.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable obtenida manualmente en [GNRA+03]). **Partición: Caso 2.**

Núm. Clases: (3,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	4000	TMP:(0.31,0.35,0.34) NUB:(0.58,0.42) VV: (0.71,0.29) O3:(0.06,0.94)	(9,11,14,16)	0.8844	0.1156	12.9258
	2000	TMP:(0.30,0.34,0.36) NUB:(0.58,0.42) VV: (0.68,0.32) O3:(0.06,0.94)	(9,11,14,16)	0.8833	0.1167	12.5201
	1000	TMP:(0.5,0.21,0.29) NUB:(0.47,0.53) VV: (0.38,0.62) O3:(0.62,0.38)	(13,14,15,16)	0.8481	0.1519	10.5341
	500	TMP:(0.50,0.21,0.29) NUB:(0.50,0.50) VV: (0.39,0.61) O3:(0.62,0.38)	(13,14,15,16)	0.8481	0.1519	10.5515
B	4000	TMP:(0.50,0.17,0.33) NUB:(0.57,0.43) VV: (0.39,0.61) O3:(0.52,0.48)	(13,14,15,16)	0.8479	0.1521	9.8372
	4000	TMP:(0.50,0.17,0.33) NUB:(0.64,0.36) VV: (0.40,0.60) O3:(0.58,0.42)	(13,14,15,16)	0.8364	0.1636	9.9246
	2000	TMP:(0.49,0.18,0.33) NUB:(0.56,0.44) VV: (0.39,0.61) O3:(0.51,0.49)	(13,14,15,16)	0.8279	0.1721	10.1254
	2000	TMP:(0.50,0.22,0.28) NUB:(0.48,0.52) VV: (0.40,0.60) O3:(0.62,0.38)	(13,14,15,16)	0.8438	0.1562	10.4209
	1000	TMP:(0.60,0.22,0.18) NUB:(0.37,0.63) VV: (0.71,0.29) O3:(0.71,0.29)	(13,14,16)	0.7967	0.2033	9.6190
	1000	TMP:(0.53,0.22,0.25) NUB:(0.46,0.54) VV: (0.72,0.28) O3:(0.58,0.42)	(10,13,14,16)	0.7855	0.2145	10.2927
	500	TMP:(0.54,0.17,0.29) NUB:(0.45,0.55) VV:(0.52,0.48) O3:(0.58,0.42)	(10,13,14,16)	0.8218	0.1782	10.4122
	500	TMP:(0.52,0.21,0.27) NUB:(0.54,0.46) VV:(0.42,0.58) O3:(0.63,0.37)	(13,14,15,16)	0.8073	0.1927	10.5162
C		Método EFP con Máscara Subóptima	(1,12,13,16)	----	----	9.3343

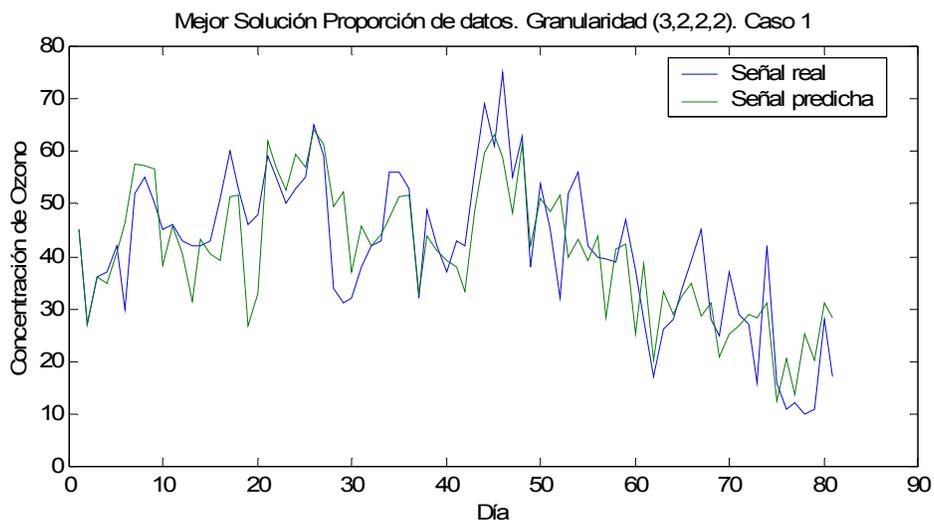
En la *Tabla 32* se presenta el tiempo de CPU usado por **MANUAL + AG2** para alcanzar las treinta ejecuciones para cada uno de los criterios de parada estipulados y distribuciones de datos (Partición Inicial, Caso 1 y Caso 2) cuando se utilizan las

funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente.

**Tabla 32.** Tiempo consumido por el MANUAL + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Distribuciones de Datos (Ozono Austria): **Partición Inicial, Caso 1 y Caso 2.** Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Función objetivo	Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas)		
		Partición Inicial	Caso 1	Caso 2
1-Q	500	0:33	0:40	0:39
	1000	1:04	1:19	1:19
	2000	2:04	2:37	2:36
	4000	4:04	5:13	5:12
$MSE_{entr}$	500	0:31	0:38	0:37
	1000	1:02	1:14	1:14
	2000	2:03	2:27	2:27
	4000	4:06	4:58	4:55

Por su parte, las *Tablas 33, 34 y 35* exhiben los resultados para la función objetivo  $MSE_{entr}$ . Estas tablas están organizadas como la *Tabla 29*. La diferencia estriba en que la sexta columna representa los valores de la función objetivo  $MSE_{entr}$  en vez de 1-Q.



**Figura 33.** Predicción del mejor modelo FIR para la **Partición Caso 1**. Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (3,2,2,2) con  $Q=0.9247$  y distribución de landmarks:  $TMP:(0.31,0.39,0.30)$ - $NUB:(0.46,0.54)$ - $VV:(0.53,0.47)$ - $O_3:(0.05,0.95)$ .  $RMS_{prue}=8.1732$ . Método MANUAL+AG2.

**Tabla 33.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo  $MSE_{entr}$  (Número de clases por variable obtenida manualmente en [GNRA+03]). **Partición Inicial.**

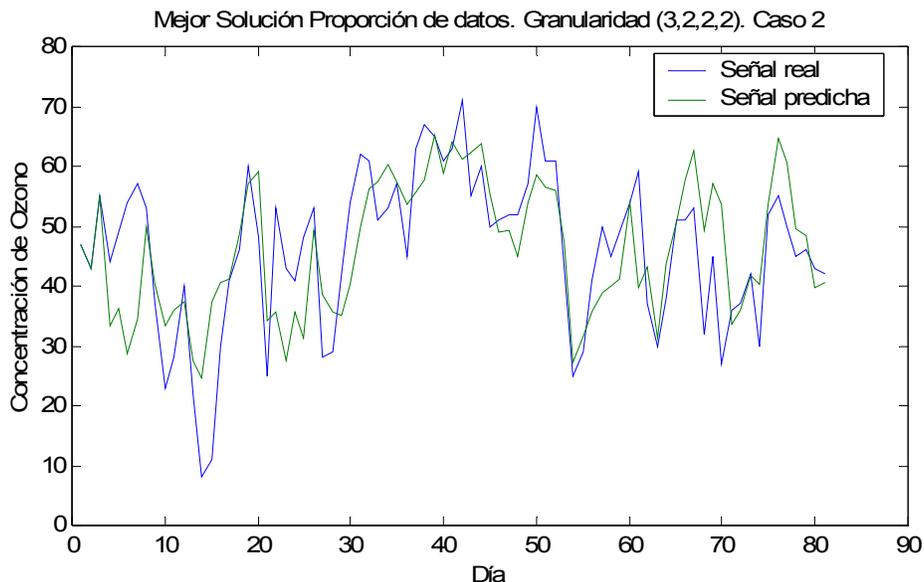
Núm. Clases: (3,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	4000	TMP:(0.16,0.42,0.42) NUB:(0.60,0.40) VV: (0.63,0.37) O3:(0.46,0.54)	(3,13,14,16)	0.3852	47.47	10.3050
	2000	TMP:(0.16,0.43,0.41) NUB:(0.60,0.40) VV: (0.66,0.34) O3:(0.48,0.52)	(3,13,14,16)	0.3936	47.48	9.8505
	1000	TMP:(0.37,0.20,0.43) NUB:(0.66,0.34) VV: (0.27,0.73) O3:(0.40,0.60)	(1,12,14,15,16)	0.4138	52.57	14.2783
	500	TMP:(0.31,0.30,0.39) NUB:(0.60,0.40) VV: (0.29,0.71) O3:(0.39,0.61)	(1,12,14,15,16)	0.4086	55.66	14.3945
B	4000	TMP:(0.36,0.45,0.19) NUB:(0.55,0.45) VV: (0.52,0.48) O3:(0.65,0.35)	(13,14,16)	0.4851	59.51	8.8040
	4000	TMP:(0.35,0.45,0.20) NUB:(0.64,0.36) VV: (0.64,0.36) O3:(0.67,0.33)	(13,14,16)	0.4870	59.69	9.3573
	2000	TMP:(0.36,0.45,0.19) NUB:(0.62,0.38) VV: (0.78,0.22) O3:(0.65,0.35)	(13,14,16)	0.4851	59.51	9.0753
	2000	TMP:(0.16,0.32,0.52) NUB:(0.63,0.37) VV: (0.69,0.31) O3:(0.57,0.43)	(3,13,14,16)	0.4079	47.49	9.4274
	1000	TMP:(0.36,0.45,0.19) NUB:(0.62,0.38) VV: (0.86,0.14) O3:(0.64,0.36)	(13,14,16)	0.4273	59.51	9.0392
	1000	TMP:(0.36,0.33,0.31) NUB:(0.58,0.42) VV: (0.61,0.39) O3:(0.68,0.32)	(13,14,16)	0.4831	59.51	9.3283
	500	TMP:(0.12,0.34,0.54) NUB:(0.63,0.37) VV: (0.25,0.75) O3:(0.65,0.35)	(13,14,16)	0.4570	59.94	9.3587
	500	TMP:(0.29,0.23,0.48) NUB:(0.71,0.29) VV: (0.25,0.75) O3:(0.46,0.54)	(12,13,14,16)	0.3530	61.12	9.4836
C		Método EFP con Máscara óptima	(7,12,14,15,16)	0.4306	---	13.6768
D		Método EFP con Máscara Subóptima	(8,9,13,16)	0.2866	---	9.7577

**Tabla 34.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo  $MSE_{entr}$  (Número de clases por variable obtenida manualmente en [GNRA+03]).  
**Partición: Caso 1.**

Núm. Clases: (3,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	4000	TMP:(0.29,0.42,0.29) NUB:(0.57,0.43) VV: (0.34,0.66) O3:(0.40,0.60)	(11,13,14,16)	0.7096	57.18	9.7850
	2000	TMP:(0.29,0.42,0.29) NUB:(0.62,0.38) VV: (0.32,0.68) O3:(0.37,0.63)	(11,13,14,16)	0.7008	57.04	9.7990
	1000	TMP:(0.29,0.42,0.29) NUB:(0.57,0.43) VV: (0.30,0.70) O3:(0.47,0.53)	(11,13,14,16)	0.6751	58.67	9.7194
	500	TMP:(0.27,0.44,0.29) NUB:(0.56,0.44) VV: (0.28,0.72) O3:(0.47,0.53)	(6,13,14,16)	0.6220	63.63	9.5303
B	4000	TMP:(0.4,0.31,0.29) NUB:(0.57,0.43) VV: (0.17,0.83) O3:(0.58,0.42)	(6,13,14,16)	0.6020	70.75	8.9693
	4000	TMP:(0.71,0.14,0.15) NUB:(0.35,0.65) VV: (0.45,0.55) O3:(0.30,0.70)	(11,13,14,16)	0.5792	67.87	9.0477
	2000	TMP:(0.30,0.45,0.25) NUB:(0.56,0.44) VV: (0.53,0.47) O3:(0.35,0.65)	(7,13,14,16)	0.6825	72.59	8.9193
	2000	TMP:(0.37,0.34,0.29) NUB:(0.55,0.45) VV: (0.47,0.53) O3:(0.63,0.37)	(6,13,14,16)	0.5605	72.78	9.2567
	1000	TMP:(0.37,0.34,0.29) NUB:(0.57,0.43) VV: (0.88,0.12) O3:(0.51,0.49)	(6,13,14,16)	0.5917	72.29	8.9777
	1000	TMP:(0.27,0.44,0.29) NUB:(0.55,0.45) VV: (0.20,0.80) O3:(0.61,0.39)	(6,13,14,16)	0.5112	63.12	9.3320
	500	TMP:(0.41,0.31,0.28) NUB:(0.57,0.43) VV: (0.40,0.60) O3:(0.67,0.33)	(6,13,14,16)	0.6511	71.60	8.8942
	500	TMP:(0.41,0.30,0.29) NUB:(0.57,0.43) VV: (0.42,0.58) O3:(0.63,0.37)	(6,13,14,16)	0.5861	71.05	8.9058
C		Método EFP con Máscara Subóptima	(8,12,13,14,16)	----	----	9.3165

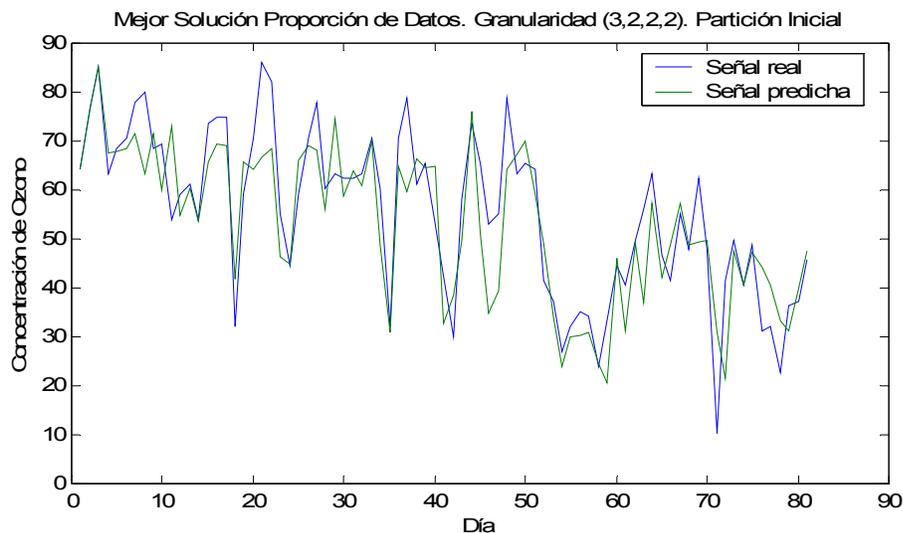
**Tabla 35.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo  $MSE_{entr}$  (Número de clases por variable obtenida manualmente en [GNRA+03]).  
**Partición: Caso 2.**

Núm. Clases: (3,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	4000	TMP:(0.28,0.41,0.31) NUB:(0.40,0.60) VV: (0.65,0.35) O3:(0.05,0.95)	(3,13,16)	0.9497	67.65	10.9785
	2000	TMP:(0.30,0.36,0.34) NUB:(0.41,0.59) VV: (0.65,0.35) O3:(0.06,0.94)	(3,13,16)	0.9301	73.46	11.2286
	1000	TMP:(0.11,0.44,0.45) NUB:(0.69,0.31) VV: (0.43,0.57) O3:(0.34,0.66)	(13,14,16)	0.5026	75.53	10.8441
	500	TMP:(0.11,0.33,0.56) NUB:(0.72,0.28) VV: (0.74,0.26) O3:(0.39,0.61)	(13,14,16)	0.4601	75.53	11.2436
B	4000	TMP:(0.32,0.29,0.39) NUB:(0.52,0.48) VV: (0.63,0.37) O3:(0.24,0.76)	(3,13,14,16)	0.6360	80.34	10.2195
	4000	TMP:(0.31,0.36,0.33) NUB:(0.48,0.52) VV: (0.63,0.37) O3:(0.24,0.76)	(3,13,14,16)	0.6552	80.18	10.3558
	2000	TMP:(0.30,0.29,0.41) NUB:(0.53,0.47) VV: (0.65,0.35) O3:(0.41,0.59)	(3,13,14,16)	0.6417	80.66	10.1238
	2000	TMP:(0.48,0.26,0.26) NUB:(0.37,0.63) VV: (0.63,0.37) O3:(0.28,0.72)	(3,13,14,16)	0.5386	78.94	10.1511
	1000	TMP:(0.30,0.34,0.36) NUB:(0.45,0.55) VV: (0.63,0.37) O3:(0.34,0.66)	(3,13,14,16)	0.6557	80.69	10.0844
	1000	TMP:(0.30,0.36,0.34) NUB:(0.45,0.55) VV: (0.64,0.36) O3:(0.47,0.53)	(3,13,14,16)	0.6658	80.66	10.2040
	500	TMP:(0.35,0.15,0.50) NUB:(0.57,0.43) VV: (0.64,0.36) O3:(0.31,0.69)	(3,13,14,16)	0.6311	88.73	10.0253
	500	TMP:(0.30,0.38,0.32) NUB:(0.54,0.46) VV: (0.64,0.36) O3:(0.41,0.59)	(3,13,14,16)	0.6457	80.66	10.3622
C		Método EFP con Máscara Subóptima	(1,12,13,16)	----	----	9.3343

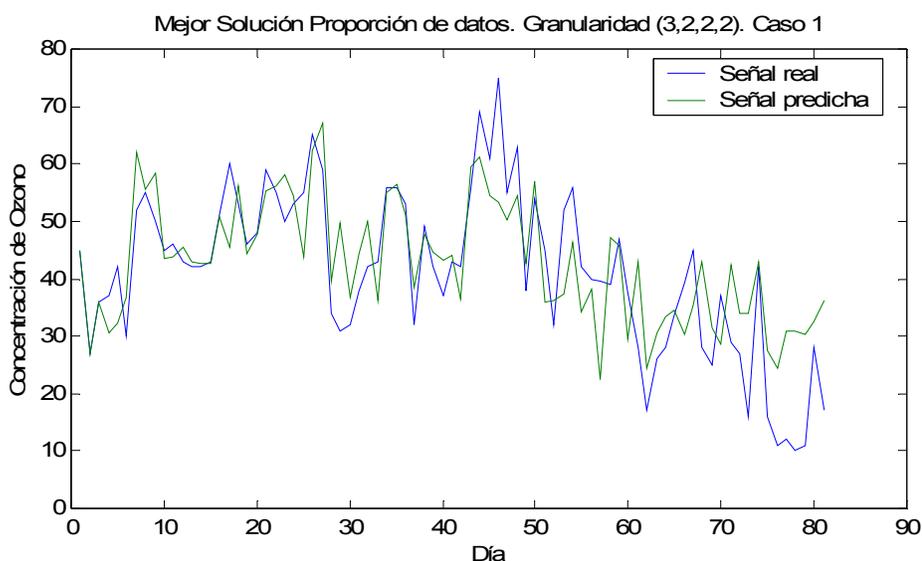


**Figura 34.** Predicción del mejor modelo FIR para la **Partición Caso 2**. Problema concentraciones de ozono en Austria. Función objetivo  $1-Q$ . Método MANUAL+AG2. Granularidad (3,2,2,2) con  $Q=0.7967$  y distribución de landmarks: TMP:(0.60,0.22,0.18)-NUB:(0.37,0.63)-VV:(0.71,0.29)- $O_3$ :(0.71,0.29).  $RMS_{prue}=9.6190$ . Método MANUAL+AG2.

Por su parte, las *figuras 35 y 36* muestran como los mejores modelos obtenidos usando la función objetivo  $MSE_{entr}$  para la Partición Inicial y Caso 1 siguen el comportamiento de la señal real.



**Figura 35.** Predicción del mejor modelo FIR para la **Partición Inicial**. Problema concentraciones de ozono en Austria. Función objetivo  $MSE_{entr}$ . Granularidad (3,2,2,2) con  $Q=0.4851$  y distribución de landmarks: TMP:(0.36,0.45, 0.19)-NUB:(0.55,0.45)-VV:(0.52,0.48)- $O_3$ :(0.65,0.35).  $RMS_{prue}=8.8040$ . Método MANUAL+AG2.



**Figura 36.** Predicción del mejor modelo FIR para la **Partición Caso 1**. Problema concentraciones de ozono en Austria. Función objetivo  $MSE_{entr}$ . Granularidad (3,2,2,2) con  $Q=0.6511$  y distribución de landmarks: TMP:(0.41,0.31, 0.28)-NUB:(0.57,0.43)-VV:(0.40,0.60)- $O_3$ :(0.67,0.33).  $RMS_{prue}=8.8942$ . Método MANUAL+AG2.

### 6.3.3.2 Aprendizaje del número óptimo de clases para cada variable del sistema (AG1)

El **AG1** se ejecuta para obtener el número óptimo de clases tanto para las variables de entrada como para la variable de salida. El método EFP se emplea para encontrar las funciones de pertenencia de las clases, por lo tanto se obtienen particiones difusas uniformes. En este bloque se muestran los resultados obtenidos por la combinación **AG1+EFP**.

Las *Tablas 36, 37 y 38* muestran los resultados obtenidos por el **AG1** para la estimación de concentraciones locales máximas de ozono en Austria usando las funciones objetivo: a) 1-Q, y b) Error de predicción MSE del último 8% del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). Como se ha señalado, el resto de la señal de entrenamiento (el primer 92% de los datos) se utiliza para obtener los modelos FIR (máscara y base de reglas patrón). Cada tabla se refiere a una de las distribuciones de datos consideradas (Partición Inicial, Caso 1 y Caso 2). Las tablas han sido organizadas de la siguiente manera:

- # *Eval.* : indica el número de evaluaciones de cromosomas hechas por el AG.
- *Gran.*: señala la granularidad sugerida por el **AG1** cuando finaliza su ejecución. Nótese que aquí la granularidad sugerida es el parámetro de entrada

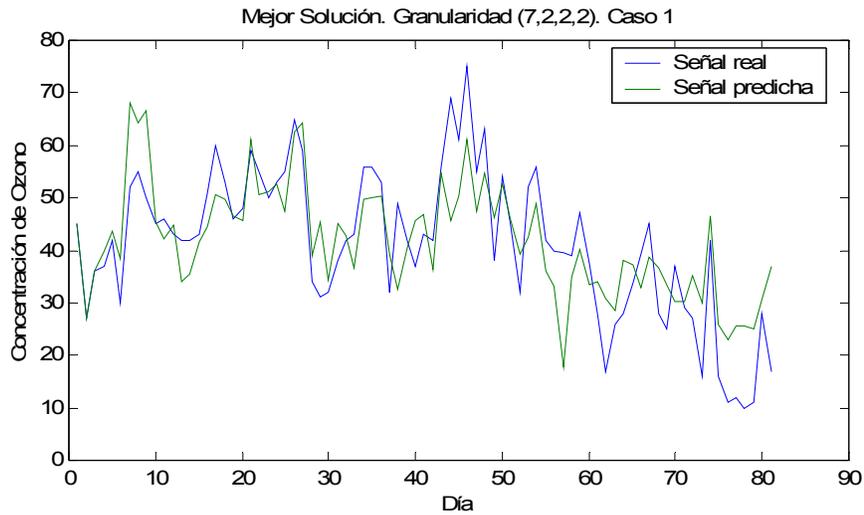
al proceso de fusificación de la metodología FIR.

- *Opt. Mask* y *Q*: contienen la máscara óptima (en notación de posición) identificada por FIR para esta granularidad específica y la calidad asociada a esta máscara, respectivamente.
- *I-Q*: corresponde al valor de la función objetivo de la calidad.
- $MSE_{entr}$ : corresponde al valor de la función objetivo del error de predicción MSE del último 8% de los datos del conjunto de entrenamiento.
- $RMS_{prue}$ : indica el error de predicción RMS con respecto al conjunto de datos de prueba.
- *Tiempo*: contiene el tiempo promedio de CPU (en segundos), el cual se calcula tomando en cuenta el menor y el mayor tiempo consumido por el AG para encontrar la solución sugerida.

**Tabla 36.** Resultados de la granularidad para la estimación de la concentración de ozono en Austria usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 8% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. **Partición Inicial.** (X significa cualquier valor entre [ 2..9]).

1-Q	# eval.	Gran.	Opt. Mask	Q	1-Q	$RMS_{prue}$	Tiempo
	800	(5,4,X,2)	(13,14,16)	0.4867	0.5133	10.5776	278
	800	(X,2,4,2)	(12,14,15,16)	0.4814	0.5186	13.7670	251
	800	(X,5,2,2)	(12,14,15,16)	0.4776	0.5224	14.5910	240
	400	(5,4,X,2)	(13,14,16)	0.4867	0.5133	10.5776	207
	400	(3,6,8,2)	(13,14,16)	0.4733	0.5267	11.2070	208
	400	(7,2,6,2)	(12,14,15,16)	0.4720	0.5280	13.3450	210
	200	(5,4,X,2)	(13,14,16)	0.4867	0.5133	10.5776	117
	200	(3,6,5,2)	(13,14,16)	0.4733	0.5267	11.2070	101
	200	(X,2,4,2)	(12,14,15,16)	0.4814	0.5186	13.7670	109
	Solución óptima sugerida: Partición = (5,4,X,2); Q=0.4867						
$MSE_{entr}$	# eval.	Gran.	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$	Tiempo
	800	(2,4,X,2)	(12,13,14,16)	0.4118	64.84	9.7832	397
	800	(3,4,5,2)	(13,14,16)	0.4032	79.37	9.8607	336
	800	(5,4,X,2)	(13,14,16)	0.4483	78.62	10.2485	337
	400	(2,4,X,2)	(12,13,14,16)	0.4118	64.84	9.7832	241
	400	(3,4,5,2)	(13,14,16)	0.4032	79.37	9.8607	269
	400	(5,4,X,2)	(13,14,16)	0.4483	78.62	10.2485	235
	200	(2,4,X,2)	(12,13,14,16)	0.4118	64.84	9.7832	124
	200	(5,4,X,2)	(13,14,16)	0.4483	78.62	10.2485	164
	200	(2,5,X,3)	(13,14,16)	0.3954	98.65	10.3223	172
	Solución óptima sugerida: Partición = (2,4,X,2); $MSE_{entr}$ =64.84						

La figura 37 muestra cómo el mejor modelo obtenido por el AG1 para la Partición Caso 1 sigue el comportamiento de la señal real de sistema bajo estudio.



**Figura 37.** Predicción del mejor modelo FIR para la **Partición Caso 1**. Problema concentraciones de ozono en Austria. Función objetivo 1-Q. Granularidad (7,2,2,2) con  $Q=0.5350$  empleo del EFP para determinar las funciones de pertenencia.  $RMS_{prue}=8.6116$ . Método AGI.

**Tabla 37.** Resultados de la granularidad para el problema de estimación de la concentración de ozono en Austria usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 8% de los datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. **Partición Caso 1**. (X significa cualquier valor entre [ 2..9]).

1-Q	# eval.	Gran.	Opt. Mask	Q	1-Q	$RMS_{prue}$	Tiempo
	800	(3,X,7,2)	(13,15,16)	0.5912	0.4088	10.3230	363
	800	(2,3,2,2)	(11,13,14,16)	0.5449	0.4552	8.6241	192
	800	(3,5,X,2)	(13,14,16)	0.5554	0.4446	9.5380	275
	400	(3,X,7,2)	(13,15,16)	0.5912	0.4088	10.3230	228
	400	(7,2,2,2)	(13,14,16)	0.5350	0.4650	8.6116	187
	400	(2,7,X,2)	(13,14,16)	0.5562	0.4438	9.4973	150
	200	(3,X,7,2)	(13,15,16)	0.5912	0.4088	10.3230	132
	200	(2,3,2,2)	(11,13,14,16)	0.5449	0.4552	8.6241	97
	200	(5,2,2,2)	(13,14,16)	0.5411	0.4589	9.4874	121
Solución óptima sugerida: Partición = (3,X,7,2); Q=0.5912							
$MSE_{entr}$	# eval.	Gran.	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$	Tiempo
	800	(2,X,X,2)	(1,12,16)	0.3880	62.33	13.7961	808
	800	(4,X,5,6)	(7,13,16)	0.4161	111.70	10.2882	524
	800	(2,7,X,5)	(9,14,16)	0.3843	98.42	10.3380	492
	400	(2,X,X,2)	(1,12,16)	0.3880	62.33	13.7961	346
	400	(7,2,5,5)	(10,13,16)	0.4413	104.91	10.2654	352
	400	(4,8,9,4)	(8,13,16)	0.3641	86.07	10.2696	332
	200	(2,X,X,2)	(1,12,16)	0.3880	62.33	13.7961	204
	200	(2,7,X,5)	(9,14,16)	0.3843	98.42	10.3380	182
	200	(7,7,2,7)	(7,13,16)	0.4000	94.14	10.4040	140
Solución óptima sugerida: Partición =(2,X,X,2); $MSE_{entr}=62.33$							

**Tabla 38.** Resultados de la granularidad para el problema de estimación de la concentración de ozono en Austria usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 8% de los datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. **Partición Caso 2.** (X significa cualquier valor entre [ 2..9]).

1-Q	# eval.	Gran.	Opt. Mask	Q	1-Q	$RMS_{prue}$	Tiempo
	800	(3,2,2,2)	(13,14,15,16)	0.7745	0.2255	11.3350	227
	800	(3,8,2,2)	(11,13,15,16)	0.7342	0.2658	10.9740	293
	800	(6,2,X,2)	(13,14,16)	0.7619	0.2381	11.0000	241
	400	(3,2,2,2)	(13,14,15,16)	0.7745	0.2255	11.3350	171
	400	(5,2,9,2)	(13,14,16)	0.7452	0.2548	10.2620	145
	400	(3,6,2,2)	(11,13,15,16)	0.7342	0.2658	10.9740	184
	200	(3,2,2,2)	(13,14,15,16)	0.7745	0.2255	11.3350	122
	200	(6,2,X,2)	(13,14,16)	0.7619	0.2381	11.0000	117
	200	(3,2,4,2)	(13,14,16)	0.7342	0.2658	11.0350	123
	Solución óptima sugerida: Partición = (3,2,2,2); Q=0.7745						
$MSE_{entr}$	# eval.	Gran.	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$	Tiempo
	800	(9,2,X,3)	(13,14,16)	0.5661	98.29	10.3630	940
	800	(7,9,9,8)	(13,16)	0.3671	133.11	11.5370	1172
	800	(5,4,4,9)	(3,13,16)	0.4343	113.14	11.9990	939
	400	(9,2,X,3)	(13,14,16)	0.5661	98.29	10.3630	354
	400	(7,8,8,8)	(13,16)	0.3671	133.11	11.5370	521
	400	(5,6,4,3)	(3,13,16)	0.5152	98.58	12.2600	378
	200	(9,2,X,3)	(13,14,16)	0.5661	98.29	10.3630	237
	200	(9,2,X,2)	(13,14,16)	0.6831	98.29	10.4320	264
	200	(8,2,9,5)	(13,14,16)	0.4789	108.94	10.7260	234
	Solución óptima sugerida: Partición = (9,2,X,3); $MSE_{entr}$ =98.29						

### 6.3.3.3 Aprendizaje de las funciones de pertenencia de las clases (AG2). Granularidad suministrada por el AG1

En esta sección se exhiben los resultados obtenidos por el **AG2** utilizado para obtener las funciones de pertenencia de las clases. En este caso, se considera el número óptimo de clases suministrado por el **AG1**. Por lo tanto, se tiene la combinación **AG1 + AG2**.

Las *Tablas 39, 40 y 41* presentan los resultados obtenidos por el **AG2** usando la función objetivo 1-Q tomando en cuenta las granularidades suministradas por el **AG1** para cada distribución de datos. Las tablas están organizadas como la *Tabla 29*. La diferencia radica en que la primera columna ahora se divide solamente en dos secciones. La sección *A* corresponde a los mejores resultados de la función objetivo obtenidos por el **AG2**. La sección *B* corresponde a los mejores resultados de predicción sobre el conjunto de prueba obtenidos por el **AG2**. Por su parte, las *Tablas 42, 43 y 44* muestran los resultados obtenidos por el **AG2** para la función objetivo

$MSE_{entr}$  y están organizadas como la *Tabla 33* pero sólo se contemplan dos secciones *A* y *B* del mismo modo que la *Tabla 39*.

**Tabla 39.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). **Partición Inicial.**

Núm. Clases: (5,4,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	4000	TMP:(0.29,0.32,0.16,0.13,0.10) NUB:(0.26,0.15,0.47,0.12) VV: (0.60,0.40) O3:(0.94,0.06)	(3,14,16)	0.9799	0.0201	14.6507
	2000	TMP:(0.15,0.42,0.17,0.19,0.07) NUB:(0.31,0.12,0.40,0.17) VV: (0.58,0.42) O3:(0.94,0.06)	(3,14,16)	0.9798	0.0202	15.9742
	1000	TMP:(0.27,0.23,0.10,0.11,0.29) NUB:(0.31,0.24,0.21,0.24) VV: (0.55,0.45) O3:(0.95,0.05)	(1,3,16)	0.9787	0.0213	17.0665
	500	TMP:(0.24,0.25,0.24,0.12,0.15) NUB:(0.32,0.15,0.27,0.26) VV: (0.35,0.65) O3:(0.94,0.06)	(2,11,15,16)	0.9768	0.0232	17.0459
B	4000	TMP:(0.14,0.15,0.10,0.32,0.29) NUB:(0.35,0.36,0.17,0.12) VV: (0.60,0.40) O3:(0.94,0.06)	(3,13,16)	0.9798	0.0202	10.1140
	4000	TMP:(0.28,0.13,0.13,0.18,0.28) NUB:(0.24,0.33,0.22,0.21) VV: (0.60,0.40) O3:(0.94,0.06)	(3,13,16)	0.9793	0.0207	10.4497
	2000	TMP:(0.18,0.12,0.22,0.19,0.29) NUB:(0.29,0.17,0.16,0.38) VV: (0.53,0.47) O3:(0.94,0.06)	(3,13,16)	0.9762	0.0238	9.9990
	2000	TMP:(0.18,0.36,0.10,0.08,0.28) NUB:(0.30,0.47,0.16,0.07) VV: (0.39,0.61) O3:(0.92,0.08)	(7,13,16)	0.9605	0.0395	11.0847
	1000	TMP:(0.17,0.11,0.21,0.23,0.28) NUB:(0.20,0.12,0.40,0.28) VV: (0.54,0.46) O3:(0.94,0.06)	(3,13,16)	0.9777	0.0223	9.8054
	1000	TMP:(0.19,0.17,0.23,0.22,0.19) NUB:(0.20,0.50,0.12,0.18) VV: (0.47,0.53) O3:(0.90,0.10)	(13,16)	0.8897	0.1103	9.9606
	500	TMP:(0.24,0.08,0.28,0.10,0.30) NUB:(0.33,0.19,0.26,0.22) VV: (0.56,0.44) O3:(0.84,0.16)	(7,13,16)	0.8534	0.1466	10.1001
	500	TMP:(0.33,0.13,0.20,0.11,0.23) NUB:(0.16,0.16,0.29,0.39) VV: (0.69,0.31) O3:(0.94,0.06)	(13,16)	0.9649	0.0351	11.2074

**Tabla 40.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). **Partición Caso 1.**

Núm. Clases: (7,2,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	4000	TMP:(0.22,0.09,0.12,0.13,0.15,0.12,0.17) NUB:(0.62,0.38) VV: (0.38,0.62) O3:(0.05,0.95)	(10,13,16)	0.9438	0.0562	11.3017
	2000	TMP:(0.21,0.15,0.17,0.11,0.11,0.15,0.10) NUB:(0.56,0.44) VV: (0.50,0.50) O3:(0.05,0.95)	(6,11,14,15,16)	0.9267	0.0733	10.8737
	1000	TMP:(0.19,0.14,0.16,0.14,0.11,0.15,0.11) NUB:(0.32,0.68) VV: (0.43,0.57) O3:(0.05,0.95)	(3,5,16)	0.9379	0.0621	14.6469
	500	TMP:(0.12,0.36,0.05,0.09,0.06,0.15,0.17) NUB:(0.55,0.45) VV:(0.31,0.69) O3:(0.05,0.95)	(3,10,15,16)	0.8900	0.1100	15.0828
B	4000	TMP:(0.13,0.12,0.26,0.13,0.11,0.13,0.12) NUB:(0.58,0.42) VV: (0.65,0.35) O3:(0.94,0.06)	(13,16)	0.8441	0.1559	9.9229
	4000	TMP:(0.13,0.18,0.14,0.09,0.19,0.15,0.12) NUB:(0.49,0.51) VV: (0.57,0.43) O3:(0.94,0.06)	(13,16)	0.8469	0.1531	10.0527
	2000	TMP:(0.18,0.10,0.09,0.18,0.27,0.11,0.07) NUB:(0.73,0.27) VV: (0.53,0.47) O3:(0.14,0.86)	(3,13,16)	0.7969	0.2031	9.5901
	2000	TMP:(0.18,0.17,0.16,0.12,0.15,0.12,0.10) NUB:(0.36,0.64) VV: (0.53,0.47) O3:(0.11,0.89)	(3,13,16)	0.7929	0.2071	9.8653
	1000	TMP:(0.12,0.13,0.16,0.14,0.22,0.11,0.12) NUB:(0.56,0.44) VV: (0.72,0.28) O3:(0.94,0.06)	(13,16)	0.8433	0.1567	9.7921
	1000	TMP:(0.17,0.13,0.16,0.14,0.11,0.16,0.13) NUB:(0.64,0.36) VV: (0.55,0.45) O3:(0.14,0.86)	(2,13,16)	0.7585	0.2415	9.8414
	500	TMP:(0.09,0.09,0.25,0.08,0.30,0.11,0.08) NUB:(0.56,0.44) VV:(0.59,0.41) O3:(0.95,0.05)	(13,16)	0.8305	0.1695	9.3463
	500	TMP:(0.12,0.05,0.17,0.40,0.05,0.07,0.14) NUB:(0.68,0.32) VV:(0.48,0.52) O3:(0.83,0.17)	(13,16)	0.7047	0.2953	9.5654

**Tabla 41.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). **Partición Caso 2.**

Núm. Clases: (5,2,9,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	4000	TMP:(0.34,0.18,0.14,0.17,0.17) NUB:(0.48,0.52) VV: (0.09,0.10,0.12,0.11,0.13,0.11,0.12,0.13,0.09) O3:(0.60,0.40)	(13,14,16)	0.8561	0.1439	11.1710
	2000	TMP:(0.34,0.18,0.14,0.17,0.17) NUB: (0.54,0.46) VV: (0.14,0.09,0.11,0.14,0.15,0.14,0.09,0.09,0.05) O3:(0.59,0.41)	(13,14,16)	0.8561	0.1439	11.1710
	1000	TMP:(0.34,0.18,0.14,0.19,0.15) NUB:(0.54,0.46) VV: (0.11,0.10,0.15,0.15,0.09,0.11,0.08,0.11, 0.10) O3:(0.59,0.41)	(13,14,16)	0.8545	0.1455	11.1608
	500	TMP:(0.27,0.25,0.12,0.19,0.17) NUB:(0.54,0.46) VV :(0.12,0.12,0.10,0.14,0.11,0.11,0.12,0.07,0.11) O3:(0.61,0.39)	(13,14,16)	0.8486	0.1514	10.5218
B	4000	TMP:(0.24,0.28,0.14,0.17,0.17) NUB:(0.51,0.49) VV: (0.11,0.12,0.15,0.09,0.16,0.09,0.10,0.08,0.10) O3:(0.61,0.39)	(13,14,16)	0.8513	0.1487	10.1966
	4000	TMP:(0.31,0.21,0.14,0.17,0.17) NUB:(0.50,0.50) VV: (0.16,0.13,0.12,0.18,0.11,0.08,0.06,0.06, 0.10) O3:(0.59,0.41)	(13,14,16)	0.8540	0.1460	10.2097
	2000	TMP:(0.24,0.30,0.16,0.17,0.13) NUB:(0.54,0.46) VV: (0.10,0.19,0.09,0.08,0.10,0.10,0.09,0.12,0.13) O3:(0.62,0.38)	(13,14,16)	0.8356	0.1644	9.9389
	2000	TMP:(0.32,0.20,0.15,0.15,0.18) NUB:(0.48,0.52) VV: (0.11,0.10,0.13,0.11,0.12,0.11,0.09,0.11,0.12) O3:(0.58,0.42)	(13,14,16)	0.8399	0.1601	10.1683
	1000	TMP:(0.24,0.28,0.16,0.18,0.14) NUB:(0.48,0.52) VV: (0.16,0.12,0.09,0.11,0.12,0.09,0.12,0.11,0.08) O3:(0.58,0.42)	(13,14,16)	0.8192	0.1808	10.1818
	1000	TMP:(0.20,0.32,0.15,0.20,0.13) NUB:(0.47,0.53) VV: (0.12,0.11,0.07,0.09,0.10,0.17,0.07,0.18,0.09) O3:(0.58,0.42)	(13,14,16)	0.8297	0.1703	10.2288
	500	TMP:(0.19,0.25,0.14,0.22,0.20) NUB: (0.57,0.43) VV:(0.09,0.11,0.12,0.11,0.17,0.10,0.12,0.09,0.09) O3:(0.53,0.47)	(13,14,16)	0.7857	0.2143	9.8198
	500	TMP:(0.21,0.31,0.14,0.16,0.18) NUB:(0.53,0.47) VV:(0.12,0.12,0.09,0.07,0.11,0.14,0.12,0.11,0.12) O3:(0.59,0.41)	(13,14,16)	0.8392	0.1608	10.3689

**Tabla 42.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1). **Partición Inicial.**

Núm. Clases: (2,4,2,2)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	4000	TMP:(0.55,0.45) NUB:(0.17,0.21,0.40,0.22) VV: (0.20,0.80) O3:(0.50,0.50)	(1,12,14,16)	0.3930	41.88	12.9503
	2000	TMP:(0.55,0.45) NUB:(0.21,0.20,0.37,0.22) VV: (0.56,0.44) O3:(0.50,0.50)	(1,12,14,16)	0.4248	42.08	13.7231
	1000	TMP:(0.54,0.46) NUB:(0.19,0.21,0.38,0.22) VV: (0.49,0.51) O3:(0.50,0.50)	(1,12,14,16)	0.4239	42.53	14.0486
	500	TMP:(0.54,0.46) NUB:(0.22,0.20,0.38,0.20) VV: (0.47,0.53) O3:(0.50,0.50)	(1,12,14,16)	0.4174	45.45	15.6726
B	4000	TMP:(0.26,0.74) NUB:(0.34,0.20,0.19,0.27) VV: (0.66,0.34) O3:(0.46,0.54)	(12,13,14,16)	0.3855	54.97	9.9243
	4000	TMP:(0.49,0.51) NUB:(0.17,0.27,0.27,0.29) VV: (0.29,0.71) O3:(0.52,0.48)	(12,13,14,16)	0.4440	57.19	9.9839
	2000	TMP:(0.41,0.59) NUB:(0.34,0.23,0.18,0.25) VV: (0.48,0.52) O3:(0.51,0.49)	(12,13,14,16)	0.3936	57.61	9.1685
	2000	TMP:(0.34,0.66) NUB:(0.44,0.13,0.16,0.27) VV: (0.39,0.61) O3:(0.50,0.50)	(12,13,14,16)	0.3497	56.38	9.5133
	1000	TMP:(0.42,0.58) NUB:(0.23,0.36,0.17,0.24) VV: (0.39,0.61) O3:(0.51,0.49)	(12,13,14,16)	0.4055	58.03	9.3154
	1000	TMP:(0.41,0.59) NUB:(0.20,0.25,0.32,0.23) VV: (0.32,0.68) O3:(0.47,0.53)	(12,13,14,16)	0.3848	58.61	9.6655
	500	TMP:(0.35,0.65) NUB:(0.38,0.17,0.23,0.22) VV: (0.90,0.10) O3:(0.41,0.59)	(12,13,14,16)	0.4012	57.62	9.3747
	500	TMP:(0.43,0.57) NUB:(0.36,0.24,0.16,0.24) VV: (0.49,0.51) O3:(0.51,0.49)	(12,13,14,16)	0.4061	57.65	9.4469

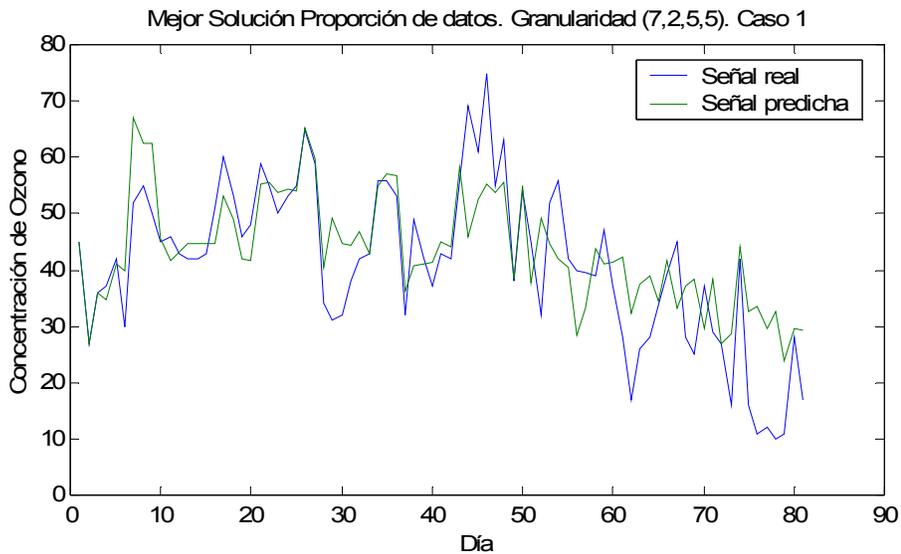
**Tabla 43.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1). **Partición Caso 1.**

Núm. Clases: (7,2,5,5)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	4000	TMP:(0.11,0.14,0.25,0.11,0.14,0.13,0.12) NUB:(0.59,0.41) VV: (0.10,0.18,0.09,0.26,0.37) O3:(0.29,0.15,0.12,0.09,0.35)	(2,13,16)	0.4847	56.10	10.6968
	2000	TMP:(0.11,0.09,0.21,0.35,0.12,0.06,0.06) NUB:(0.45,0.55) VV: (0.21,0.19,0.11,0.24,0.25) O3:(0.23,0.22,0.14,0.18,0.23)	(6,12,14,16)	0.3837	60.56	14.0877
	1000	TMP:(0.09,0.16,0.20,0.14,0.15,0.14,0.12) NUB:(0.47,0.53) VV: (0.29,0.22,0.16,0.16,0.17) O3:(0.24,0.14,0.18,0.21,0.23)	(6,12,14,16)	0.3996	50.26	15.3648
	500	TMP:(0.11,0.14,0.19,0.15,0.20,0.11,0.10) NUB:(0.64,0.36) VV: (0.30,0.13,0.17,0.21,0.19) O3:(0.29,0.15,0.14,0.12,0.30)	(2,13,16)	0.4697	59.95	10.7710
B	4000	TMP:(0.11,0.17,0.13,0.18,0.18,0.14,0.09) NUB:(0.52,0.48) VV: (0.11,0.14,0.22,0.12,0.41) O3:(0.23,0.18,0.17,0.14,0.28)	(2,13,14,16)	0.4431	70.56	9.0861
	4000	TMP:(0.14,0.11,0.26,0.11,0.13,0.16,0.09) NUB:(0.55,0.45) VV: (0.26,0.28,0.16,0.17,0.13) O3:(0.22,0.20,0.19,0.18,0.21)	(2,13,16)	0.3992	70.14	10.1947
	2000	TMP:(0.19,0.16,0.15,0.08,0.11,0.15,0.16) NUB:(0.68,0.32) VV: (0.35,0.14,0.18,0.16,0.17) O3:(0.31,0.12,0.19,0.22,0.16)	(13,14,16)	0.4498	83.34	9.5529
	2000	TMP:(0.19,0.20,0.15,0.17,0.10,0.09,0.10) NUB:(0.41,0.59) VV: (0.16,0.21,0.21,0.30,0.12) O3:(0.23,0.13,0.21,0.20,0.23)	(10,13,16)	0.4296	79.41	10.2082
	1000	TMP:(0.14,0.17,0.12,0.15,0.18,0.09,0.15) NUB:(0.57,0.43) VV: (0.27,0.14,0.22,0.21,0.16) O3:(0.31,0.12,0.18,0.14,0.25)	(13,14,16)	0.5337	75.04	10.1180
	1000	TMP:(0.15,0.15,0.17,0.09,0.10,0.23,0.11) NUB:(0.46,0.54) VV: (0.18,0.21,0.20,0.23,0.18) O3:(0.22,0.21,0.13,0.19,0.25)	(10,13,16)	0.4193	90.59	10.3664
	500	TMP:(0.11,0.19,0.19,0.12,0.15,0.16,0.08) NUB:(0.53,0.47) VV: (0.18,0.30,0.18,0.18,0.16) O3:(0.20,0.24,0.17,0.19,0.20)	(2,13,14,16)	0.4021	78.82	9.2199
	500	TMP:(0.19,0.11,0.11,0.10,0.17,0.19,0.13) NUB:(0.60,0.40) VV: (0.20,0.19,0.23,0.23,0.15) O3:(0.15,0.15,0.34,0.13,0.23)	(10,13,16)	0.4779	82.36	10.1385

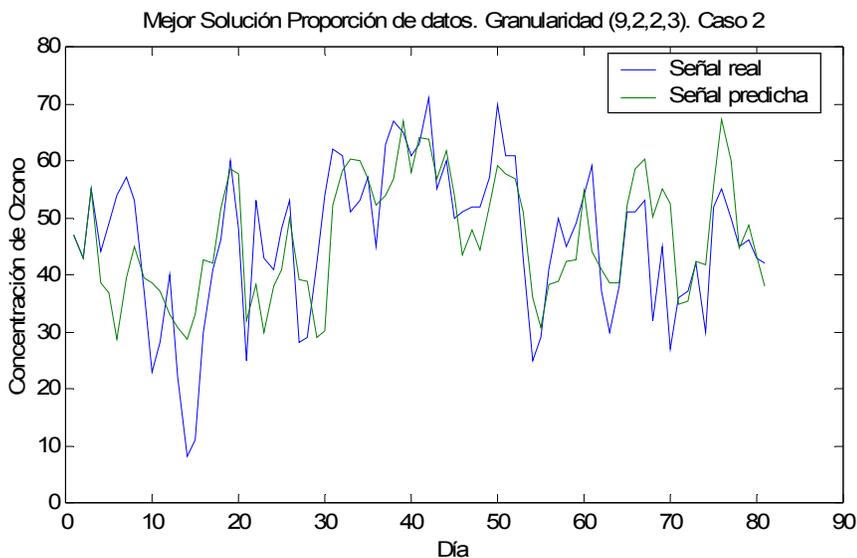
**Tabla 44.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en Austria usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1). **Partición Caso 2.**

Núm. Clases: (9,2,2,3)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	4000	TMP:(0.11,0.12,0.10,0.10,0.15,0.10,0.08,0.14,0.10) NUB:(0.64,0.36) VV: (0.63,0.37) O3:(0.19,0.10,0.71)	(3,13,16)	0.6211	69.85	11.4159
	2000	TMP:(0.11,0.08,0.12,0.12,0.13,0.12,0.12,0.11,0.09) NUB:(0.40,0.60) VV: (0.65,0.35) O3:(0.19,0.12,0.69)	(3,13,16)	0.6334	69.96	11.1952
	1000	TMP:(0.11, 0.2,0.12,0.12,0.11,0.07,0.16,0.06,0.05) NUB:(0.39,0.61) VV: (0.73,0.27) O3:(0.38,0.27,0.35)	(3,13,16)	0.5237	72.09	12.1791
	500	TMP:(0.11,0.09,0.13,0.11,0.10,0.14,0.09,0.09,0.14) NUB:(0.5, 0.5) VV: (0.64,0.36) O3:(0.39, 0.3,0.31)	(3,13,16)	0.5166	72.24	11.4115
B	4000	TMP:(0.12,0.11,0.10,0.11,0.10,0.10,0.16,0.15,0.05) NUB:(0.56,0.44) VV: (0.46,0.54) O3:(0.22,0.25,0.53)	(13,14,16)	0.6111	80.26	10.0261
	4000	TMP:(0.12,0.10,0.10,0.11,0.11,0.11,0.10,0.12,0.13) NUB:(0.57,0.43) VV: (0.58,0.42) O3:(0.22,0.18,0.60)	(13,14,16)	0.5933	78.60	10.1025
	2000	TMP:(0.11,0.09,0.15,0.09,0.11,0.06,0.05,0.16,0.18) NUB:(0.70,0.30) VV: (0.43,0.57) O3:(0.33,0.33,0.34)	(13,14,16)	0.5288	75.53	9.8730
	2000	TMP:(0.12,0.10,0.11,0.10,0.10,0.12,0.11,0.11,0.13) NUB:(0.59,0.41) VV: (0.46,0.54) O3:(0.23,0.10,0.67)	(13,14,16)	0.6312	78.02	10.1087
	1000	TMP:(0.11,0.11,0.12,0.11,0.14,0.09,0.09,0.09,0.14) NUB:(0.68,0.32) VV: (0.56,0.44) O3:(0.33,0.33,0.34)	(13,14,16)	0.5155	75.53	9.7258
	1000	TMP:(0.11,0.11,0.12,0.11,0.08,0.10,0.11,0.11,0.15) NUB:(0.67,0.33) VV: (0.39,0.61) O3:(0.27,0.32,0.41)	(13,14,16)	0.5645	75.53	9.7956
	500	TMP:(0.12,0.09,0.10,0.16,0.06,0.12,0.08,0.22,0.05) NUB:(0.64,0.36) VV: (0.56,0.44) O3:(0.23,0.35,0.42)	(13,14,16)	0.6148	81.30	9.6089
	500	TMP:(0.12,0.09,0.11,0.11,0.13,0.10,0.10,0.10,0.14) NUB:(0.55,0.45) VV: (0.47,0.53) O3:(0.47,0.14,0.39)	(13,14,16)	0.6634	82.04	9.7781

Las *figuras 38* y *39* muestran como los mejores modelos obtenidos por el método **AG1+AG2** usando la función objetivo  $MSE_{entr}$  siguen la señal real del sistema bajo estudio para las distribuciones de datos Caso 1 y Caso 2.



**Figura 38.** Predicción del mejor modelo FIR para la **Partición Caso 1**. Problema concentraciones de ozono en Austria. Función objetivo  $MSE_{entr}$ . Granularidad (7,2,5,5) con  $Q=0.4431$  y distribución de landmarks:  $TMP:(0.11,0.17,0.13, 0.18,0.18,0.14,0.09)$ - $NUB:(0.52,0.48)$ - $VV:(0.11,0.14,0.22,0.12,0.41)$ - $O_3:(0.23,0.18,0.17,0.14,0.28)$ .  $RMS_{prue}=9.0861$ . Método  $AG1+AG2$ .



**Figura 39.** Predicción del mejor modelo FIR para la **Partición Caso 2**. Problema concentraciones de ozono en Austria. Función objetivo  $MSE_{entr}$ . Granularidad (9,2,2,3) con  $Q=0.6148$  y distribución de landmarks:  $TMP:(0.12,0.09, 0.10, 0.16,0.06, 0.12, 0.08, 0.22,0.05)$ - $NUB:(0.64,0.36)$ - $VV:(0.56,0.44)$ - $O_3:(0.23,0.35, 0.42)$ .  $RMS_{prue}=9.6089$ . Método  $AG1+AG2$ .

La *Tabla 45* muestra el tiempo de CPU necesitado por la combinación **AG1+AG2** para realizar las treinta ejecuciones para cada uno de los criterios de parada estipulados y distribuciones de datos (Partición Inicial, Caso 1 y Caso 2) cuando se utilizan las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente

**Tabla 45.** Tiempo consumido por el AG1 + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Distribuciones de Datos (Ozono Austria): **Partición Inicial, Caso 1 y Caso 2.** Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Función objetivo	Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas)		
		Partición Inicial	Caso 1	Caso 2
1-Q	500	0:34	0:57	0:39
	1000	1:09	1:52	1:19
	2000	2:17	3:53	2:36
	4000	4:33	7:43	5:12
$MSE_{entr}$	500	0:52	2:19	1:20
	1000	1:40	4:36	2:38
	2000	3:15	9:03	5:13
	4000	6:26	18:02	10:25

#### 6.3.3.4 Aprendizaje global del número óptimo de clases y de las funciones de pertenencia de las clases (AG3)

A continuación se exponen los resultados para el algoritmo genético **AG3**, que optimiza de manera conjunta los parámetros de fusificación de la metodología FIR, es decir, el número de clases y las funciones de pertenencia que definen su semántica.

Las *Tablas 46, 47 y 48* muestran los resultados obtenidos para la estimación de concentraciones de ozono usando la función objetivo 1-Q para cada una de las distribuciones de datos consideradas (Partición Inicial, Caso 1 y Caso 2). Las tablas están organizadas como se describe a continuación. La primera columna se divide en cinco secciones. La sección *A* corresponde a los mejores resultados de la función objetivo obtenidos por el **AG3**. La sección *B* corresponde a los mejores resultados de predicción sobre el conjunto de prueba obtenidos por el **AG3**. La sección *C* indica los resultados obtenidos cuando se emplean parámetros por defecto en FIR (**FIR DEFECTO**), es decir, todas las variables se discretizan en tres grupos y se obtienen particiones difusas uniformes usando el método EFP para establecer las funciones de pertenencia. La sección *D* y *E* muestran los resultados de las combinaciones **AG1** y **AG1+AG2** mostrados en apartados previos. *C, D* y *E* se usan como referencia para el análisis de los resultados. La segunda columna indica el criterio de parada, es decir, el número de evaluaciones de cromosomas realizadas por el AG. La tercera y cuarta

columnas presentan la solución sugerida por el **AG3** (número de clases y distribución de landmarks) para las variables de entrada (TMP,NUB,VV) y la variable de salida ( $O_3$ ). La quinta columna contiene la máscara óptima (en notación de posición) encontrada por FIR para esa partición. La sexta columna corresponde a la calidad asociada a la máscara óptima. La séptima columna es el valor de la función objetivo 1-Q. La última columna muestra el error de predicción RMS para el conjunto de datos de prueba ( $RMS_{prue}$ ).

Por otro lado, las *Tablas 49, 50 y 51* presentan los resultados del **AG3** cuando se usa la función objetivo  $MSE_{entr}$ . Estas tablas están organizadas como la *Tabla 46*. La diferencia está en que la séptima columna contiene los valores de la función objetivo  $MSE_{entr}$  en vez de la de 1-Q.

La *Tabla 52* muestra el tiempo de CPU que requiere el **AG3** para realizar las treinta ejecuciones para cada uno de los criterios de parada estipulados y distribuciones de datos (Partición Inicial, Caso 1 y Caso 2) cuando se utilizan las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente

Las *figuras 40 y 41* muestran la señal real y las señales de predicción obtenidas por los mejores modelos del método **AG3** (función objetivo 1-Q), para las distribuciones de datos Partición Inicial y Caso 1. Las *figuras 42 y 43* muestran ambas señales cuando se usan los mejores modelos del método **AG3** (función objetivo  $MSE_{entr}$ ) para esas mismas distribuciones de datos.

**Tabla 46.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo 1-Q. **Partición Inicial.**

# eval	Gran.	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>	
A	16000	(2,3,7,2)	TMP:(0.65,0.35) NUB:(0.29,0.40,0.31)	(3,16)	0.9823	0.0177	20.9833
	8000	(3,4,7,2)	VV: (0.14,0.14,0.17,0.16,0.06,0.16,0.17) O3:(0.95,0.05) TMP:(0.28,0.42,0.30) NUB:(0.26,0.27,0.22,0.25)	(3,16)	0.9823	0.0177	20.1400
	4000	(2,3,3,2)	VV: (0.14,0.15,0.15,0.16,0.06,0.16,0.18) ; O3:(0.95,0.05) TMP:(0.68,0.32) NUB:(0.26,0.16,0.58)	(13,14,16)	0.9814	0.0186	11.9828
	2000	(2,3,2,2)	VV: (0.33,0.33,0.34) ; O3:(0.95,0.05) TMP:(0.70,0.30) NUB:(0.36,0.28,0.36)	(6,13,16)	0.9808	0.0192	11.8037
	1000	(9,7,8,2)	VV: (0.56,0.44) ; O3:(0.94,0.06) TMP:(0.07,0.08,0.09,0.05,0.13,0.19,0.05,0.17,0.17) NUB:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) VV: (0.27,0.12,0.06,0.05,0.12,0.15,0.06,0.17) O3:(0.95,0.05)	(5,16)	0.9771	0.0229	16.7516
	500	(2,7,2,2)	TMP:(0.66,0.34) NUB:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) VV: (0.20,0.80) ; O3:(0.93,0.07)	(14,16)	0.9548	0.0452	13.0337
B	16000	(3,7,4,3)	TMP:(0.28,0.37,0.35) NUB:(0.12,0.12,0.24,0.12,0.13,0.13,0.14), VV: (0.28,0.17,0.28,0.27) ; O3:(0.89,0.06,0.05)	(3,13,16)	0.9456	0.0544	10.2188
	16000	(2,3,5,3)	TMP:(0.64,0.36) NUB:(0.20,0.23,0.57)	(10,13,16)	0.9580	0.0420	10.8162
	8000	(4,3,3,4)	VV: (0.16,0.16,0.16,0.38,0.14) ; O3:(0.90,0.05,0.05) TMP:(0.36,0.19,0.16,0.29) NUB:(0.25,0.29,0.46)	(2,13,16)	0.9078	0.0922	10.8603
	8000	(3,3,2,4)	VV: (0.29,0.46,0.25) ; O3:(0.83,0.07,0.05,0.05) TMP:(0.50,0.21,0.29) NUB:(0.30,0.34,0.36)	(2,13,16)	0.9034	0.0966	11.0906
	4000	(6,2,7,2)	VV: (0.50,0.50) ; O3:(0.83,0.05,0.07,0.05) TMP:(0.19,0.30,0.11,0.08,0.07,0.25) NUB:(0.50,0.50)	(2,13,16)	0.8860	0.1140	10.5350
	4000	(2,4,9,3)	VV: (0.11,0.11,0.11,0.12,0.19,0.23,0.13) O3:(0.88,0.12) TMP:(0.64,0.36) NUB:(0.19,0.17,0.17,0.47)	(10,13,16)	0.9567	0.0433	10.8910
	2000	(5,2,8,4)	VV: (0.09,0.05,0.22,0.12,0.20,0.10,0.11,0.05,0.06) O3:(0.90,0.05,0.05) TMP:(0.31,0.22,0.10,0.10,0.27) NUB:(0.43,0.57)	(13,14,16)	0.7596	0.2404	9.7214
	2000	(3,3,2,2)	VV: (0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) O3:(0.68,0.07,0.20,0.05) TMP:(0.30,0.41,0.29) NUB:(0.19,0.20,0.61)	(11,13,16)	0.9452	0.0548	10.5619
	1000	(3,4,2,2)	VV: (0.38,0.62) ; O3:(0.91,0.09) TMP:(0.38,0.27,0.35) NUB:(0.25,0.25,0.25,0.25)	(1,3,13,16)	0.7446	0.2554	9.8303
	1000	(3,5,6,4)	VV: (0.50,0.50) ; O3:(0.76,0.24) TMP:(0.28,0.26,0.46) NUB:(0.20,0.20,0.20,0.20,0.20)	(13,14,16)	0.5910	0.4090	10.1629
	500	(3,7,3,2)	VV: (0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.05,0.05,0.58,0.32) TMP:(0.33,0.33,0.34)	(13,15,16)	0.7593	0.2407	8.8196
	500	(5,7,2,3)	NUB:(0.15,0.15,0.15,0.15,0.15,0.10,0.15) VV: (0.33,0.33,0.34) ; O3:(0.80,0.20) TMP:(0.25,0.26,0.08,0.17,0.24) VV: (0.14,0.14,0.14,0.14,0.14,0.15,0.15) VV: (0.50,0.50) ; O3:(0.64,0.06,0.30)	(3,13,16)	0.6214	0.3786	9.8226
C		(3,3,3,3)	Equal Frequency Partition method	(12,14,15,16)	0.4393	0.5607	14.0723
D		(5,4,2,2)	Equal Frequency Partition method	(13,14,16)	0.4867	0.5133	10.5776
E		(5,4,2,2)	TMP:(0.17,0.11,0.21,0.23,0.28) NUB:(0.20,0.12,0.40,0.28) VV: (0.54,0.46) ; O3:(0.94,0.06)	(3,13,16)	0.9777	0.0223	9.8054

**Tabla 47.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo 1-Q. **Partición Caso 1.**

	# eval	Gran.	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	16000	(3,2,2,2)	TMP:(0.28,0.38,0.34) NUB:(0.82,0.18)	(1,3,14,16)	0.9477	0.0523	12.2658
			VV: (0.52,0.48) ; O3:(0.05,0.95)				
	8000	(3,3,5,2)	TMP:(0.28,0.38,0.34) NUB:(0.33,0.33,0.34)	(7,13,16)	0.9503	0.0497	10.2067
			VV: (0.20,0.20,0.20,0.20,0.20) ; O3:(0.05,0.95)				
	4000	(3,5,5,2)	TMP:(0.28,0.42,0.30) NUB:(0.20,0.20,0.20,0.20,0.20)	(7,13,16)	0.9503	0.0497	10.0988
			VV: (0.20,0.20,0.20,0.20,0.20) ; O3:(0.05,0.95)				
	2000	(3,5,3,2)	TMP:(0.29,0.35,0.36) NUB:(0.12,0.27,0.15,0.27,0.19)	(1,14,16)	0.9233	0.0767	12.8642
			VV: (0.49,0.30,0.21) ; O3:(0.05,0.95)				
	1000	(2,2,2,2)	TMP:(0.45,0.55) NUB:(0.68,0.32)	(1,10,14,16)	0.8835	0.1165	11.685
			VV: (0.50,0.50) ; O3:(0.07,0.93)				
	500	(7,3,6,2)	TMP:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) NUB:(0.33,0.33,0.34)	(13,16)	0.8804	0.1196	10.4194
		VV: (0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.06,0.94)					
B	16000	(3,4,4,2)	TMP:(0.28,0.27,0.45) NUB:(0.29,0.24,0.24,0.23)	(11,13,16)	0.9041	0.0959	9.5778
			VV: (0.21,0.33,0.16,0.30) ; O3:(0.06,0.94)				
	16000	(7,9,5,2)	TMP:(0.09,0.09,0.28,0.16,0.09,0.17,0.12) NUB:(0.13,0.11,0.11,0.11,0.07,0.11,0.12,0.12,0.12)	(13,16)	0.8473	0.1527	9.6790
			VV: (0.22,0.20,0.20,0.15,0.23) ; O3:(0.94,0.06)				
	8000	(2,3,5,2)	TMP:(0.28,0.72) NUB:(0.68,0.24,0.08)	(13,14,16)	0.8674	0.1326	9.3807
			VV: (0.20,0.20,0.20,0.20,0.20) ; O3:(0.09,0.91)				
	8000	(7,4,2,2)	TMP:(0.23,0.28,0.17,0.12,0.06,0.07,0.07) NUB:(0.18,0.20,0.37,0.25)	(3,13,16)	0.8906	0.1094	9.6695
			VV: (0.41,0.59) ; O3:(0.95,0.05)				
	4000	(9,3,9,2)	TMP:(0.11,0.05,0.16,0.12,0.11,0.11,0.10,0.11,0.13) NUB:(0.25,0.16,0.59)	(13,16)	0.8280	0.1720	9.1383
			VV: (0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) O3:(0.94,0.06)				
	4000	(8,3,2,2)	TMP:(0.10,0.10,0.10,0.11,0.12,0.23,0.12,0.12) NUB:(0.16,0.34,0.50)	(13,16)	0.8455	0.1545	9.2037
			VV: (0.40,0.60) ; O3:(0.94,0.06)				
	2000	(9,4,3,2)	TMP:(0.10,0.10,0.10,0.10,0.15,0.10,0.11,0.11,0.13), NUB:(0.25,0.25,0.25,0.25)	(13,16)	0.7869	0.2131	9.5800
			VV: (0.33,0.33,0.34) ; O3:(0.89,0.11)				
	2000	(2,9,4,2)	TMP:(0.29,0.71) NUB:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12)	(13,15,16)	0.8781	0.1219	9.6338
			VV: (0.20,0.27,0.27,0.26) ; O3:(0.07,0.93)				
1000	(7,9,7,2)	TMP:(0.30,0.27,0.05,0.21,0.05,0.07,0.05) NUB:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12)	(13,16)	0.8580	0.1420	9.4220	
		VV: (0.28,0.15,0.07,0.14,0.06,0.25,0.05) O3:(0.95,0.05)					
1000	(2,3,2,2)	TMP:(0.57,0.43) NUB:(0.33,0.33,0.34)	(11,13,14,16)	0.6537	0.3463	9.6851	
		VV: (0.67,0.33) ; O3:(0.42,0.58)					
	500	(8,5,3,2)	TMP:(0.17,0.07,0.24,0.06,0.11,0.22,0.07,0.06) NUB:(0.20,0.20,0.20,0.20,0.20)	(13,16)	0.7512	0.2488	8.8088
		VV: (0.40,0.07,0.53) ; O3:(0.84,0.16)					
	500	(4,3,2,2)	TMP:(0.25,0.21,0.26,0.28) NUB:(0.33,0.33,0.34)	(11,13,15,16)	0.7956	0.2044	8.9897
		VV: (0.47,0.53) ; O3:(0.09,0.91)					
C		(3,3,3,3)	Equal Frequency Partition method	(7,13,15,16)	0.4740	0.5260	No predice
D		(7,2,2,2)	Equal Frequency Partition method	(13,14,16)	0.5350	0.4650	8.6116
E		(7,2,2,2)	TMP:(0.09,0.09,0.25,0.08,0.30,0.11,0.08) NUB:(0.56,0.44) VV:(0.59,0.41) ; O3:(0.95,0.05)	(13,16)	0.8305	0.1695	9.3463

**Tabla 48.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo 1-Q. **Partición Caso 2.**

	# eval	Gran.	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	16000	(5,4,3,2)	TMP:(0.09,0.16,0.31,0.33,0.11) NUB:(0.24,0.39,0.21,0.16) VV: (0.41,0.34,0.25) ; O3:(0.08,0.92)	(3,13,16)	0.8903	0.1097	12.7219
	8000	(4,2,3,2)	TMP:(0.10,0.17,0.19,0.54) NUB:(0.54,0.46) VV: (0.49,0.39,0.12) ; O3:(0.14,0.86)	(3,13,16)	0.8949	0.1051	12.8056
	4000	(5,9,2,2)	TMP:(0.12,0.14,0.40,0.14,0.20) NUB:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) VV: (0.72,0.28) ; O3:(0.05,0.95)	(9,11,16)	0.8802	0.1198	15.3636
	2000	(7,8,2,2)	TMP:(0.10,0.09,0.12,0.16,0.14,0.27,0.12) NUB:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) VV: (0.73,0.27) ; O3:(0.05,0.95)	(11,13,16)	0.8696	0.1304	12.4752
	1000	(9,6,8,2)	TMP:(0.05,0.05,0.05,0.10,0.05,0.15,0.34,0.05,0.16) NUB:(0.16,0.16,0.17,0.17,0.17,0.17) VV: (0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) O3:(0.14,0.86)	(13,16)	0.8448	0.1552	13.0607
	500	(9,2,2,2)	TMP:(0.10,0.19,0.10,0.10,0.10,0.10,0.10,0.10,0.11) NUB:(0.50,0.50) VV: (0.50,0.50) ; O3:(0.14,0.86)	(13,16)	0.8325	0.1675	12.6968
B	16000	(8,2,3,2)	TMP:(0.22,0.12,0.11,0.09,0.10,0.09,0.09,0.18) NUB:(0.50,0.50) VV: (0.35,0.33,0.32) ; O3:(0.60,0.40)	(13,14,16)	0.8358	0.1642	10.2179
	16000	(7,2,4,3)	TMP:(0.26,0.15,0.13,0.10,0.09,0.09,0.18) NUB:(0.47,0.53) VV: (0.26,0.26,0.26,0.22) ; O3:(0.52,0.05,0.43)	(13,14,16)	0.8689	0.1311	10.2701
	8000	(7,2,3,3)	TMP:(0.21,0.15,0.18,0.13,0.07,0.11,0.15) NUB:(0.48,0.52) VV: (0.46,0.28,0.26) ; O3:(0.52,0.05,0.43)	(13,14,16)	0.8713	0.1287	10.2884
	8000	(9,2,5,5)	TMP:(0.13,0.17,0.09,0.07,0.07,0.08,0.09,0.12,0.18) NUB:(0.57,0.43) VV: (0.19,0.18,0.21,0.26,0.16) O3:(0.46,0.05,0.05,0.05,0.39)	(13,14,16)	0.8097	0.1903	10.3932
	4000	(7,2,2,3)	TMP:(0.24,0.17,0.11,0.14,0.07,0.13,0.14) NUB:(0.50,0.50) VV: (0.37,0.63) ; O3:(0.55,0.06,0.39)	(13,14,16)	0.8515	0.1485	10.2828
	4000	(3,2,2,2)	TMP:(0.53,0.18,0.29) NUB:(0.51,0.49) VV: (0.41,0.59) ; O3:(0.62,0.38)	(13,14,15,16)	0.8608	0.1392	10.4954
	2000	(8,2,3,3)	TMP:(0.21,0.10,0.10,0.11,0.12,0.12,0.12,0.12) NUB:(0.54,0.46) VV: (0.39,0.19,0.42) ; O3:(0.52,0.07,0.41)	(13,14,16)	0.8074	0.1926	10.2933
	2000	(7,2,6,3)	TMP:(0.21,0.11,0.13,0.12,0.12,0.17,0.14) NUB:(0.46,0.54) VV: (0.09,0.29,0.09,0.30,0.08,0.15) O3:(0.53,0.05,0.42)	(13,14,16)	0.8450	0.1550	10.3888
	1000	(9,2,3,2)	TMP:(0.23,0.13,0.09,0.08,0.08,0.07,0.10,0.08,0.14), NUB:(0.49,0.51) VV: (0.33,0.33,0.34) O3:(0.6, 0.4)	(13,14,16)	0.8168	0.1832	10.3477
	1000	(7,2,2,2)	TMP:(0.21,0.16,0.15,0.15,0.11,0.13,0.09) NUB:(0.46,0.54) VV: (0.45,0.55) ; O3:(0.57,0.43)	(13,14,16)	0.8094	0.1906	10.3792
	500	(3,4,4,3)	TMP:(0.45,0.19,0.36) NUB:(0.25,0.25,0.25,0.25) VV: (0.10,0.43,0.09,0.38) O3:(0.48,0.05,0.47)	(13,14,16)	0.8049	0.1951	10.3245
	500	(8,2,6,2)	TMP:(0.25,0.11,0.10,0.10,0.11,0.11,0.11,0.11) NUB:(0.50,0.50) VV: (0.20,0.43,0.07,0.18,0.06,0.06) O3:(0.56,0.44)	(13,14,16)	0.7936	0.2064	10.3555
	C	(3,3,3,3)	Equal Frequency Partition method	(2,13,16)	0.5584	0.4416	11.5686
	D	(5,2,9,2)	Equal Frequency Partition method	(13,14,16)	0.7452	0.2548	10.2620
E	(5,2,9,2)	TMP:(0.19,0.25,0.14,0.22,0.20) NUB:(0.57,0.43) VV:(0.09,0.11,0.12,0.11,0.17,0.10,0.12,0.09,0.09) O3:(0.53,0.47)	(13,14,16)	0.7857	0.2143	9.8198	

**Tabla 49.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo  $MSE_{entr}$ . **Partición Inicial.**

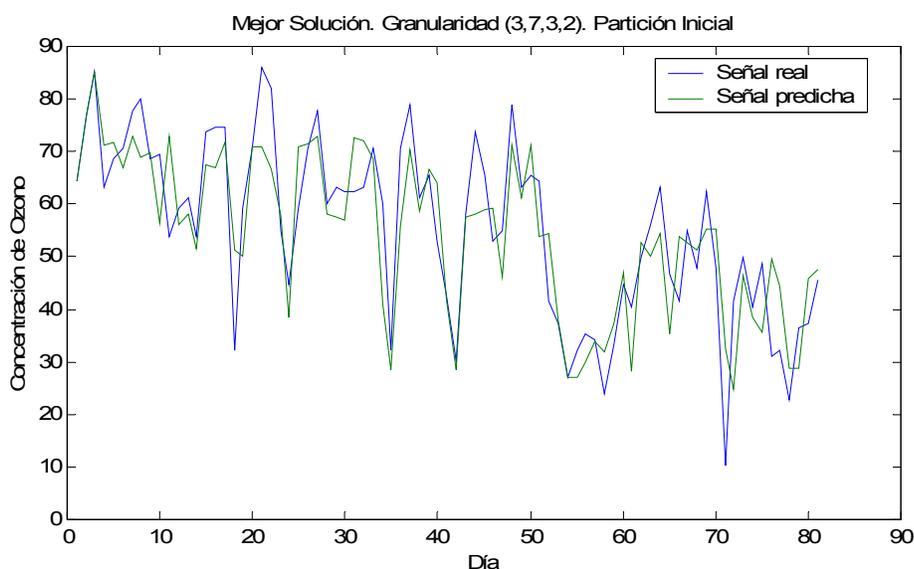
# eval	Gran.	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{pmue}$	
A	16000	(6,4,6,6)	TMP:(0.10,0.15,0.18,0.22,0.27,0.08) NUB:(0.22,0.26,0.27,0.25) VV: (0.09,0.11,0.21,0.14,0.11,0.34) O3:(0.05,0.41,0.07,0.07,0.07,0.33)	(13,14,16)	0.4298	41.44	22.9025
	8000	(2,3,3,3)	TMP:(0.36,0.64) ; NUB:(0.34,0.30,0.36) VV: (0.24,0.32,0.44) ; O3:(0.06,0.48,0.46)	(11,13,14,16)	0.5170	37.98	11.7773
	4000	(4,4,4,4)	TMP:(0.10,0.22,0.19,0.49) NUB:(0.19,0.40,0.11,0.30) VV: (0.09,0.15,0.55,0.21) O3:(0.06,0.41,0.26,0.27)	(13,14,16)	0.3721	46.74	10.0564
	2000	(8,7,5,9)	TMP:(0.11,0.11,0.11,0.11,0.12,0.13,0.13,0.18) NUB:(0.11,0.29,0.21,0.05,0.19,0.08,0.07) VV: (0.08,0.22,0.07,0.06,0.57) O3:(0.09,0.05,0.14,0.17,0.06,0.08,0.08,0.06,0.27)	(13,14,16)	0.2863	48.31	24.5347
	1000	(9,6,7,9)	TMP:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) NUB:(0.23,0.34,0.06,0.16,0.10,0.11) VV: (0.37,0.24,0.12,0.05,0.05,0.08,0.09) O3:(0.07,0.09,0.07,0.19,0.18,0.08,0.13,0.12,0.07)	(13,14,16)	0.3283	50.03	32.8391
	500	(6,6,5,8)	TMP:(0.12,0.10,0.27,0.15,0.12,0.24) NUB:(0.05,0.37,0.19,0.22,0.05,0.12) VV: (0.23,0.11,0.26,0.09,0.31) O3:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13)	(13,14,16)	0.2972	52.60	11.1948
B	16000	(3,3,4,3)	TMP:(0.13,0.23,0.64) ; NUB:(0.47,0.31,0.22) VV: (0.32,0.13,0.27,0.28) ; O3:(0.05,0.66,0.29)	(13,14,16)	0.5847	46.68	8.8591
	16000	(3,2,3,4)	TMP:(0.25,0.10,0.65) ; NUB:(0.58,0.42) VV: (0.17,0.19,0.64) ; O3:(0.05,0.55,0.35,0.05)	(13,14,16)	0.5964	54.85	9.3618
	8000	(3,3,6,3)	TMP:(0.36,0.35,0.29) ; NUB:(0.11,0.50,0.39) VV: (0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.06,0.47,0.47)	(13,14,16)	0.4428	57.13	9.2453
	8000	(3,2,6,4)	TMP:(0.36,0.13,0.51) NUB:(0.55,0.45) VV: (0.14,0.32,0.16,0.11,0.15,0.12) O3:(0.05,0.73,0.17,0.05)	(13,14,16)	0.7417	54.16	9.4123
	4000	(3,3,4,3)	TMP:(0.12,0.26,0.62) NUB:(0.44,0.28,0.28) VV: (0.18,0.18,0.17,0.47) ; O3:(0.05,0.56,0.39)	(13,14,16)	0.4549	48.27	9.3005
	4000	(3,2,6,3)	TMP:(0.36,0.20,0.44) ; NUB:(0.58,0.42) VV: (0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.05,0.71,0.24)	(13,14,16)	0.6945	54.33	9.4444
	2000	(2,3,3,7)	TMP:(0.41,0.59) ; NUB:(0.36,0.58,0.06) VV: (0.56,0.34,0.10) O3:(0.08,0.30,0.21,0.07,0.11,0.17,0.06)	(1,13,14,16)	0.3226	51.68	9.7052
	2000	(3,2,3,3)	TMP:(0.33,0.33,0.34) ; NUB:(0.66,0.34) VV: (0.06,0.54,0.40) ; O3:(0.05,0.53,0.42)	(13,14,16)	0.5365	55.00	9.7414
	1000	(2,3,9,2)	TMP:(0.47,0.53) ; NUB:(0.44,0.28,0.28) VV: (0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) O3:(0.50,0.50)	(12,13,14,16)	0.3688	57.10	9.7278
	1000	(2,4,4,2)	TMP:(0.50,0.50) NUB:(0.25,0.25,0.25,0.25) VV: (0.20,0.40,0.29,0.11) ; O3:(0.50,0.50)	(12,13,14,16)	0.4118	64.84	9.7832
	500	(3,2,4,4)	TMP:(0.53,0.25,0.22) ; NUB:(0.43,0.57) VV: (0.24,0.44,0.26,0.06) O3:(0.05,0.38,0.05,0.52)	(13,14,16)	0.4333	62.84	9.4563
	500	(2,5,2,2)	TMP:(0.47,0.53) NUB:(0.27,0.14,0.19,0.14,0.26) VV: (0.50,0.50) ; O3:(0.50,0.50)	(12,13,14,16)	0.4100	63.38	9.4767
C	(3,3,3,3)	Equal Frequency Partition method	(12,14,15,16)	0.4393	----	14.0723	
D	(2,4,2,2)	Equal Frequency Partition method	(12,13,14,16)	0.4118	64.84	9.7832	
E	(2,4,2,2)	TMP:(0.41,0.59) NUB:(0.34,0.23,0.18,0.25) VV: (0.48,0.52) ; O3:(0.51,0.49)	(12,13,14,16)	0.3936	57.61	9.1685	

**Tabla 50.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo  $MSE_{entr}$ . **Partición Caso 1.**

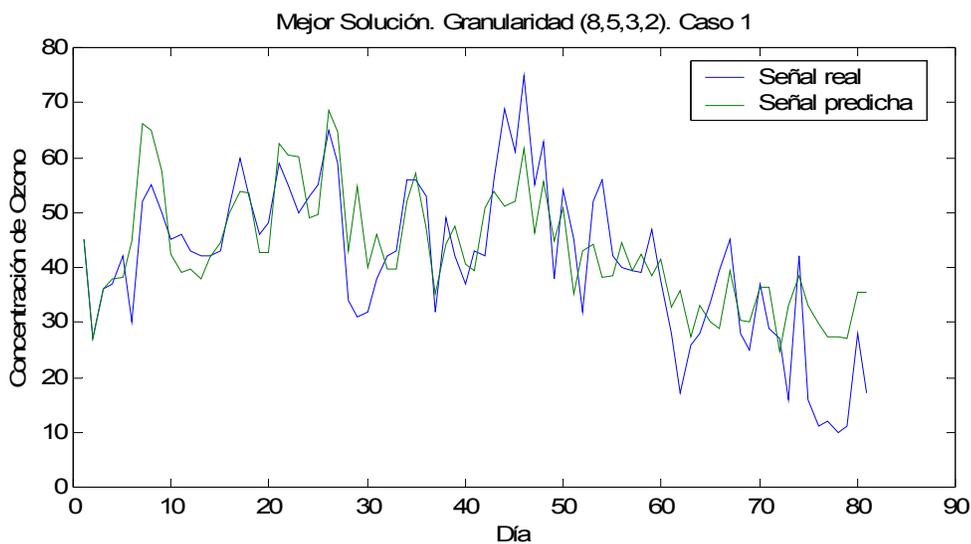
# eval	Gran.	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	16000	(2,8,5,8) TMP:(0.34,0.66) NUB:(0.13,0.12,0.12,0.13,0.13,0.17,0.11,0.09) VV: (0.21,0.13,0.27,0.14,0.25) O3:(0.09,0.10,0.10,0.10,0.11,0.12,0.29,0.09)	(5,12,16)	0.3941	31.87	15.3840
	8000	(2,4,7,8) TMP:(0.32,0.68) NUB:(0.44,0.37,0.11,0.08) VV: (0.16,0.17,0.11,0.31,0.07,0.05,0.13) O3:(0.09,0.10,0.10,0.10,0.13,0.12,0.26,0.10)	(5,12,16)	0.3736	27.30	17.3819
	4000	(5,2,9,9) TMP:(0.15,0.17,0.24,0.23,0.21) NUB:(0.35,0.65) VV: (0.10,0.15,0.10,0.10,0.10,0.10,0.11,0.12,0.12) O3:(0.10,0.10,0.10,0.10,0.10,0.10,0.09,0.09,0.22)	(12,14,16)	0.3973	38.28	11.4624
	2000	(3,2,4,4) TMP:(0.49,0.32,0.19); NUB:(0.53,0.47) VV: (0.18,0.27,0.27,0.28) O3:(0.25,0.25,0.25,0.25)	(9,11,14,16)	0.4636	47.03	10.3813
	1000	(2,2,9,8) TMP:(0.30,0.70) NUB:(0.25,0.75) VV: (0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) O3:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13)	(1,12,16)	0.3180	58.87	14.5104
	500	(4,3,5,2) TMP:(0.60,0.14,0.13,0.13); NUB:(0.42,0.42,0.16) VV: (0.20,0.20,0.20,0.20,0.20) O3:(0.46,0.54)	(8,12,14,16)	0.5890	63.69	12.3891
B	16000	(4,2,8,8) TMP:(0.29,0.24,0.29,0.18); NUB:(0.57,0.43) VV: (0.11,0.12,0.16,0.11,0.14,0.13,0.12,0.11) O3:(0.10,0.21,0.07,0.06,0.09,0.08,0.27,0.12)	(2,13,14,16)	0.4784	42.96	8.7843
	16000	(4,2,9,9) TMP:(0.24,0.29,0.29,0.18); NUB:(0.57,0.43) VV: (0.12,0.10,0.11,0.11,0.11,0.12,0.11,0.11,0.11) O3:(0.32,0.06,0.05,0.07,0.07,0.09,0.24,0.05,0.05)	(2,13,14,16)	0.4800	46.13	8.9138
	8000	(4,2,3,6) TMP:(0.17,0.53,0.13,0.17); NUB:(0.51,0.49) VV: (0.37,0.39,0.24) O3:(0.23,0.08,0.05,0.29,0.29,0.06)	(2,13,14,16)	0.4624	53.68	9.1730
	8000	(3,2,2,8) TMP:(0.72,0.09,0.19); NUB:(0.33,0.67) VV: (0.23,0.77) O3:(0.08,0.15,0.07,0.06,0.27,0.09,0.18,0.10)	(2,13,14,16)	0.3553	52.68	9.2455
	4000	(4,2,7,3) TMP:(0.15,0.15,0.28,0.42); NUB:(0.68,0.32) VV: (0.12,0.22,0.12,0.12,0.12,0.14,0.16) O3:(0.36,0.31,0.33)	(13,14,16)	0.4637	79.17	8.6461
	4000	(3,2,2,6) TMP:(0.41,0.40,0.19); NUB:(0.57,0.43) VV: (0.30,0.70) O3:(0.24,0.11,0.05,0.18,0.37,0.05)	(2,13,14,16)	0.5550	50.42	8.6992
	2000	(4,2,4,7) TMP:(0.41,0.33,0.11,0.15); NUB:(0.61,0.39) VV: (0.14,0.49,0.23,0.14) O3:(0.06,0.24,0.05,0.12,0.21,0.12,0.20)	(2,13,14,16)	0.4360	53.55	9.2200
	2000	(2,2,7,7) TMP:(0.81,0.19); NUB:(0.30,0.70) VV: (0.08,0.05,0.05,0.17,0.31,0.15,0.19) O3:(0.24,0.06,0.05,0.05,0.22,0.13,0.25)	(2,13,14,16)	0.4244	53.68	9.2776
	1000	(3,3,8,5) TMP:(0.41,0.41,0.18); NUB:(0.59,0.07,0.34) VV: (0.11,0.11,0.20,0.11,0.12,0.12,0.12,0.11) O3:(0.20,0.20,0.20,0.20,0.20)	(2,13,14,16)	0.4417	64.26	8.9984
	1000	(3,4,3,4) TMP:(0.30,0.34,0.36); NUB:(0.41,0.05,0.09,0.45) VV: (0.49,0.06,0.45) O3:(0.25,0.25,0.25,0.25)	(4,13,16)	0.3983	79.10	9.1429
	500	(2,2,4,5) TMP:(0.29,0.71); NUB:(0.50,0.50) VV: (0.12,0.26,0.52,0.10) O3:(0.05,0.17,0.18,0.25,0.35)	(10,13,14,16)	0.4440	77.86	8.9412
	500	(2,2,9,3) TMP:(0.29,0.71); NUB:(0.62,0.38) VV: (0.08,0.07,0.21,0.07,0.09,0.08,0.21,0.08,0.11) O3:(0.18,0.20,0.62)	(2,13,14,16)	0.6154	82.46	9.1513
C	(3,3,3,3)	Equal Frequency Partition method	(7,13,15,16)	0.4740	-----	No predice
D	(7,2,5,5)	Equal Frequency Partition method	(10,13,16)	0.4413	104.91	10.2654
E	(7,2,5,5)	TMP:(0.11,0.17,0.13,0.18,0.18,0.14,0.09) NUB:(0.52,0.48) VV: (0.11,0.14,0.22,0.12,0.41) O3:(0.23,0.18,0.17,0.14,0.28)	(2,13,14,16)	0.4431	70.56	9.0861

**Tabla 51.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en Austria usando la función objetivo  $MSE_{entr}$ . **Partición Caso 2.**

# eval	Gran.	Data Proportion	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	16000	(7,7,4,7) TMP:(0.11,0.14,0.15,0.15,0.15,0.16,0.14) NUB:(0.11,0.12,0.19,0.22,0.12,0.12,0.12) VV: (0.16,0.27,0.28,0.29) O3:(0.20,0.06,0.14,0.17,0.14,0.15,0.14)	(3,13,16)	0.4383	52.67	12.0661
	8000	(8,8,2,6) TMP:(0.11,0.13,0.09,0.13,0.16,0.12,0.15,0.11) NUB:(0.11,0.20,0.13,0.11,0.12,0.11,0.11,0.11) VV: (0.72,0.28) O3:(0.18,0.13,0.10,0.26,0.17,0.16)	(3,13,16)	0.4125	69.82	11.9203
	4000	(9,4,3,8) TMP:(0.11,0.13,0.11,0.11,0.11,0.11,0.09,0.11,0.12) NUB:(0.25,0.25,0.25,0.25); VV: (0.43,0.28,0.29) O3:(0.17,0.14,0.10,0.10,0.11,0.11,0.13,0.14)	(3,13,16)	0.4363	53.33	12.6804
	2000	(9,2,2,6) TMP:(0.11,0.10,0.18,0.10,0.10,0.10,0.10,0.10,0.11) NUB:(0.48,0.52); VV: (0.63,0.37) O3:(0.16,0.16,0.17,0.17,0.17,0.17)	(3,13,16)	0.4534	70.30	11.9202
	1000	(9,2,3,5) TMP:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) NUB:(0.76,0.24); VV: (0.35,0.53,0.12) O3:(0.27,0.11,0.30,0.08,0.24)	(13,14,16)	0.4530	75.53	10.5897
	500	(8,8,2,8) TMP:(0.11,0.14,0.08,0.12,0.12,0.18,0.14,0.11) NUB:(0.06,0.22,0.11,0.06,0.19,0.08,0.15,0.13) VV: (0.64,0.36) O3:(0.18,0.18,0.07,0.16,0.06,0.20,0.10,0.05)	(3,13,16)	0.4406	70.23	11.6585
B	16000	(7,2,4,3) TMP:(0.11,0.11,0.24,0.11,0.11,0.12,0.20) NUB:(0.73,0.27); VV: (0.30,0.16,0.22,0.32) O3:(0.33,0.33,0.34)	(13,14,16)	0.5341	75.53	10.1107
	16000	(8,2,2,3) TMP:(0.12,0.11,0.13,0.11,0.13,0.12,0.11,0.17) NUB:(0.57,0.43); VV: (0.37,0.63) O3:(0.05,0.60,0.35)	(13,14,16)	0.7087	72.63	10.1591
	8000	(7,2,9,4) TMP:(0.12,0.11,0.11,0.11,0.11,0.12,0.32) NUB:(0.57,0.43) VV:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) O3:(0.25,0.25,0.25,0.25)	(13,14,16)	0.4809	80.90	10.2190
	8000	(9,2,5,3) TMP:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) NUB:(0.68,0.32); VV: (0.22,0.18,0.25,0.2,0.15) O3:(0.29,0.31,0.40)	(13,14,16)	0.5696	75.53	10.3061
	4000	(9,2,4,3) TMP:(0.11,0.11,0.11,0.10,0.10,0.10,0.12,0.10,0.15) NUB:(0.67,0.33); VV: (0.34,0.21,0.19,0.26) O3:(0.33,0.33,0.34)	(13,14,16)	0.5087	75.53	9.8778
	4000	(9,2,9,3) TMP:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) NUB:(0.67,0.33) VV: (0.11,0.11,0.11,0.11,0.11,0.12,0.11,0.11,0.11) O3:(0.29,0.35,0.36)	(13,14,16)	0.5110	75.53	10.3282
	2000	(8,2,2,3) TMP:(0.12,0.08,0.12,0.13,0.14,0.12,0.13,0.16), NUB:(0.58,0.42); VV: (0.73,0.27) O3:(0.27,0.17,0.56)	(13,14,16)	0.6031	80.41	10.0350
	2000	(7,2,6,8) TMP:(0.12,0.11,0.14,0.14,0.13,0.17,0.19) NUB:(0.66,0.34); VV:(0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.05,0.37,0.05,0.05,0.05,0.33,0.05,0.05)	(13,14,16)	0.5977	76.07	10.2007
	1000	(7,2,6,8) TMP:(0.12,0.10,0.21,0.20,0.09,0.11,0.17) NUB:(0.65,0.35); VV:(0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.05,0.22,0.24,0.24,0.05,0.05,0.05,0.10)	(13,14,16)	0.4548	80.39	10.0056
	1000	(8,2,7,7) TMP:(0.12,0.12,0.13,0.13,0.14,0.14,0.08,0.14) NUB:(0.55,0.45) VV: (0.26,0.07,0.06,0.05,0.05,0.41,0.10) O3:(0.05,0.35,0.23,0.18,0.09,0.05,0.05)	(13,14,16)	0.5146	76.89	10.5629
	500	(9,2,6,5) TMP:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) NUB:(0.69,0.31); VV:(0.16,0.16,0.17,0.17,0.17,0.17) O3:(0.20,0.20,0.20,0.20,0.20)	(13,14,16)	0.4695	76.50	10.4185
	500	(8,2,6,6) TMP:(0.10,0.16,0.08,0.10,0.13,0.12,0.19,0.12) NUB:(0.50,0.50) VV: (0.10,0.21,0.28,0.09,0.17,0.15) O3:(0.05,0.30,0.22,0.05,0.30,0.08)	(13,14,16)	0.5601	91.65	10.466
C	(3,3,3,3)	Equal Frequency Partition method	(2,13,16)	0.5584	-----	11.5686
D	(9,2,2,3)	Equal Frequency Partition method	(13,14,16)	0.5661	98.29	10.3630
E	(9,2,2,3)	TMP:(0.12,0.09,0.10,0.16,0.06,0.12,0.08,0.22,0.05) NUB:(0.64,0.36); VV: (0.56,0.44) O3:(0.23,0.35,0.42)	(13,14,16)	0.6148	81.30	9.6089



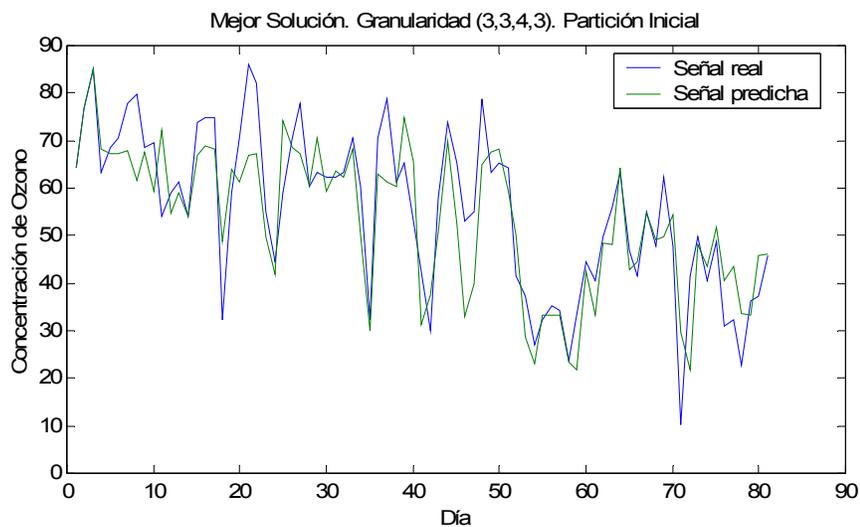
**Figura 40.** Predicción del mejor modelo FIR para la **Partición Inicial**. Problema concentraciones de ozono en Austria. Función objetivo  $1-Q$ . Granularidad  $(3,7,3,2)$  con  $Q=0.7593$  y distribución de landmarks:  $TMP:(0.33,0.33,0.34)$ - $NUB:(0.15, 0.15, 0.15, 0.15, 0.15,0.10,0.15)$ - $VV:(0.33,0.33,0.34)$ - $O_3:(0.80,0.20)$ .  $RMS_{prue}=8.8196$ . Método AG3.



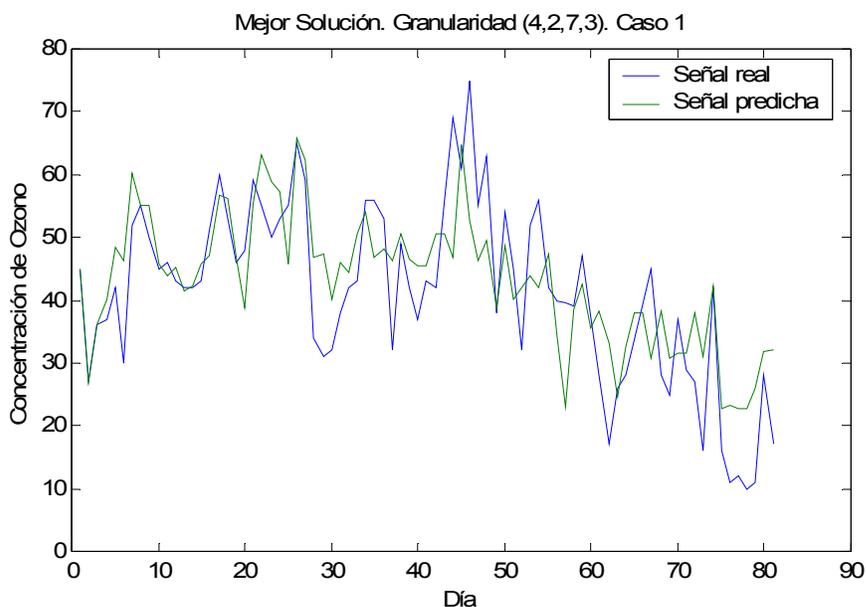
**Figura 41.** Predicción del mejor modelo FIR para la **Partición Caso 1**. Problema concentraciones de ozono en Austria. Función objetivo  $1-Q$ . Granularidad  $(8,5,3,2)$  con  $Q=0.7512$  y distribución de landmarks:  $TMP:(0.17,0.07,0.24,0.06,0.11, 0.22,0.07,0.06)$ - $NUB:(0.20,0.20,0.20,0.20,0.20)$ - $VV:(0.40,0.07,0.53)$ - $O_3:(0.84,0.16)$ .  $RMS_{prue}=8.8088$ . Método AG3.

**Tabla 52.** Tiempo consumido por el AG3 para realizar las treinta ejecuciones por cada criterio de parada. Distribuciones de Datos (Ozono Austria): **Partición Inicial, Caso 1 y Caso 2.** Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Función objetivo	Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas)		
		Partición Inicial	Caso 1	Caso 2
1-Q	500	2:49	3:16	3:17
	1000	4:51	5:40	5:07
	2000	6:44	10:26	9:04
	4000	10:40	16:53	14:33
	8000	17:13	32:25	35:23
	16000	41:19	60:21	60:06
$MSE_{entr}$	500	3:32	4:01	5:08
	1000	5:29	6:58	8:26
	2000	8:24	11:12	16:36
	4000	15:43	18:58	31:38
	8000	28:32	40:00	60:10
	16000	58:05	89:48	122:05



**Figura 42.** Predicción del mejor modelo FIR para la **Partición Inicial**. Problema concentraciones de ozono en Austria. Función objetivo  $MSE_{entr}$ . Granularidad (3,3,4,3) con  $Q=0.5847$  y distribución de landmarks:  $TMP:(0.13,0.23, 0.64)$ - $NUB:(0.47,0.31,0.22)$ - $VV:(0.32,0.13,0.27,0.28)$ - $O_3:(0.05,0.66,0.29)$ .  $RMS_{prue}=8.8591$ . Método AG3.



**Figura 43.** Predicción del mejor modelo FIR para la **Partición Caso 1**. Problema concentraciones de ozono en Austria. Función objetivo  $MSE_{entr}$  Granularidad (4,2,7,3) con  $Q=0.4637$  y distribución de landmarks:  $TMP:(0.15,0.15,0.28,0.42)$ - $NUB:(0.68,0.32)$ - $VV:(0.12,0.22,0.12,0.12,0.12,0.14,0.16)$ - $O_3:(0.36,0.31,0.33)$ .  $RMS_{prue}=8.6461$ . Método AG3.

### 6.3.4 Análisis de resultados y conclusiones

Esta sección comprende un análisis de los resultados obtenidos por los métodos **AG1**, **AG2** y **AG3** propuestos en esta memoria para la resolución del problema de estimación de concentraciones locales máximas de ozono en Austria usando las distribuciones de datos: Partición Inicial, Caso 1 y Caso 2. Se señala que las dos funciones objetivo establecidas serán evaluadas y comparadas desde la perspectiva de su eficiencia y tiempo computacional. Por último, se darán algunas conclusiones.

#### 6.3.4.1 Partición datos: Inicial

##### A) Método AG1

Si se observa la *Tabla 36* se encuentra que la solución óptima para la función objetivo 1-Q corresponde a la granularidad (5,4,X,2) con una calidad de 0.4867. Vemos que las calidades de las máscaras son en general bajas. La solución óptima cuando se usa la función objetivo  $MSE_{entr}$  es la partición (2,4,X,2) con un  $MSE_{entr}$  de

64.84%. Es necesario señalar, que las granularidades sugeridas por el AG tienen en común que siempre definen 2 clases para la variable  $O_3$ .

Como ya sucedió en anteriores aplicaciones, los modelos FIR obtenidos con la función objetivo  $MSE_{entr}$  tienen mayor eficiencia en la predicción del conjunto de prueba. El mejor  $RMS_{prue}$  fue de 9.7832, siendo éste del mismo orden de magnitud que el error del mejor modelo neuronal mostrado en la *Tabla 24* ( $RMS_{prue}=9.9579$ ) y el error del mejor modelo FIR manual mostrado en la *Tabla 25* ( $RMS_{prue}=9.7577$ ). Cuando se utilizan los parámetros de fusificación por defecto (sección *C* de la *Tabla 46*) el error aumenta hasta el 14.0723. El menor error se obtiene con las redes modificadas de Elman con función de activación tangente hiperbólica (*Tabla 27*) que consiguen un  $RMS_{prue}$  de 9.1930.

#### *B) Método AG2 con granularidad obtenida manualmente (MANUAL + AG2)*

En la *Tabla 29* se observa que todas las soluciones de la función objetivo 1-Q obtenidas por el **AG2** son mejores que los resultados de referencia (sección *C* y *D*).

En cuanto a la precisión de cada modelo (última columna) se observa que no siempre el  $RMS_{prue}$  es inferior cuando se usa el **AG2**. Como ya se ha explicado anteriormente esto es debido a que la máscara óptima se identifica a partir de los datos de entrenamiento y por lo tanto no se puede garantizar el mínimo error de predicción cuando se utiliza el modelo para predecir datos que no forman parte del conjunto de entrenamiento. Los  $RMS_{prue}$  obtenidos en la sección *A* son peores incluso a aquellos obtenidos en los estudios previos usando FIR clásico, es decir a los resultados de referencia.

Con respecto a la función objetivo  $MSE_{entr}$ , a partir de la *Tabla 33* se puede señalar que la mayoría de los resultados del **AG2** poseen errores  $RMS_{prue}$  inferiores o del mismo orden de magnitud a los resultados de referencia (*C*) y (*D*) considerados. Los mejores  $RMS_{prue}$  fueron conseguidos nuevamente por soluciones subóptimas. El menor  $RMS_{prue}$  fue de 8.8040 siendo este inferior no solo a los resultados de referencia sino también a los obtenidos por las metodologías mostradas en las *Tablas 24* y *27*.

De la *Tabla 32* se puede observar que a medida que se aumenta el número de evaluaciones, el tiempo de CPU se duplica. Además cabe señalar que los tiempos de ejecución de la función objetivo  $MSE_{entr}$  son muy parecidos a los de la función objetivo 1-Q usando la misma granularidad. Esto es consecuencia del reducido tamaño del conjunto de datos disponible. En este caso la granularidad será determinante en la duración de las ejecuciones de la función objetivo  $MSE_{entr}$ .

*C) Método AG2 con granularidad suministrada por AG1 (AG1 + AG2)*

Observando la *Tabla 39*, se aprecia que todas las soluciones sugeridas por el AG para la función objetivo 1-Q son mejores que aquellas obtenidas cuando se usa por defecto el método EFP, es decir, el resultado de referencia de la sección *D* de la *Tabla 46*. Sin embargo, los  $RMS_{prue}$  de la sección *A* han presentado, de nuevo, las peores predicciones. El mejor  $RMS_{prue}$  fue de 9.8054.

Para la función objetivo  $MSE_{entr}$  se puede indicar a partir de la *Tabla 42*, que todas las soluciones sugeridas por el **AG2** para la función objetivo  $MSE_{entr}$  son mejores al resultado de referencia de la sección *D* de la *Tabla 49*. Sin embargo, las mejores predicciones son obtenidas sólo por las soluciones subóptimas.

Los mismos comentarios que se hicieron para la combinación **MANUAL+AG2** respecto los tiempos de ejecución para las funciones objetivo  $MSE_{entr}$  y 1-Q, son válidos aquí.

Analizando los  $RMS_{prue}$  de las *Tablas 39* y *42*, se observa, una vez, más que la función objetivo  $MSE_{entr}$  tiene mayor eficiencia en la predicción del conjunto de datos de prueba que la función objetivo 1-Q. El mejor resultado logrado fue de 9.1685, siendo éste del mismo orden de magnitud que el mejor resultado obtenido en estudios previos para el mismo problema.

*D) Método AG3*

Todas las soluciones del **AG3** presentadas en las *Tablas 46* y *49* tienen una máscara de alta calidad respecto a los modelos FIR manual de la *Tabla 25* y **FIR DEFECTO** (sección *C* de cada tabla).

Nótese que aunque los  $RMS_{prue}$  son usualmente más pequeños cuando se usa el **AG3**, esto no sucede siempre debido a que el conjunto de datos de prueba no interviene en el proceso de identificación del modelo. Solamente las soluciones subóptimas (fila *B* de cada función objetivo) presentan errores más bajos a los mostrados en las *Tablas 24* y *25*.

El modelo FIR del **AG3** que tiene un mejor comportamiento en la predicción del conjunto de prueba usando la función objetivo 1-Q, tiene un valor Q de 0.7593 y la máscara óptima es: (13,15,16), que incluye una relación causal de las variables de entrada TMP y VV con la variable de salida  $O_3$ . Por su parte para la otra función objetivo, el mejor modelo encontrado tiene un valor  $MSE_{entr}$  de 46.68 y la máscara óptima es: (13,14,16), que incluye una relación causal de las variables de entrada TMP y NUB con la variable de salida  $O_3$ .

### E) Resumen

La *Tabla 53* resume los mejores resultados de predicción del conjunto de datos de prueba de la Partición Inicial logrados por las diferentes metodologías mencionadas en esta aplicación.

**Tabla 53.**  $ECM_{prue}$  de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para la **Partición Inicial** usando: AG1+EFP, MANUAL+AG2, AG1+AG2, AG3, FIR DEFECTO, FIR MANUAL y otros métodos basados en RN.

AG1+EFP	MAN.+AG2	AG1+AG2	AG3	FIR DEF.	FIR MAN.	RN
9.7832	8.8040	9.1685	8.8196	14.0723	9.7577	9.1930

Como se observa en la *Tabla 53* el menor error se logra con el método **MANUAL+AG2** seguido por el método **AG3** y **AG1+AG2**. A corta distancia se sitúan las redes neuronales MEN-HYP. Con un error medio punto mayor se encuentran los métodos **AG1+EFP** y **FIR MANUAL**. Finalmente el **FIR DEFECTO** es el que obtiene peores resultados. Es decir, los métodos que combinan FIR con AGs obtienen buenos resultados si se comparan con los demás métodos presentados.

Los mejores modelos FIR obtenidos por los métodos genéticos usando la función objetivo  $MSE_{entr}$  tienen en común una relación causal de las variables de entrada TMP y NUB con la variable de salida  $O_3$ . En cambio para la función 1-Q la mayoría presentan una relación causal de las variables de entrada TMP y VV con la variable de salida  $O_3$ . Indiscutiblemente, la temperatura juega un papel muy importante en la predicción de la concentración de ozono.

#### 6.3.4.2 Partición datos: Caso 1

##### A) Método AG1

Se observa a partir de la *Tabla 37*, para la función objetivo 1-Q, que la solución óptima sugerida corresponde a la granularidad (3,X,7,2) con la más alta calidad ( $Q=0.5912$ ), sin embargo esta solución no tiene la mejor habilidad de predicción. Por otro lado, el AG sugiere granularidades que tienen en común una partición difusa siempre de 2 clases para la variable de salida  $O_3$ .

Con respecto a la función objetivo  $MSE_{entr}$ , la partición (2,X,X,2) es la solución óptima sugerida por el **AG1** con un  $MSE_{entr}$  de 62.33%.

Por otro lado, se detecta que las soluciones subóptimas sugeridas son las que presentan mayor habilidad de predicción. En este sentido, los modelos FIR obtenidos por la función objetivo 1-Q han sido superiores a la función objetivo  $MSE_{entr}$ . El mejor  $RMS_{prue}$  fue de 8.6116 y lo obtuvo la granularidad (7,2,2,2), siendo este error inferior a los obtenidos en [GNRA+03] mostrados en la *Tabla 26*.

#### *B) Método AG2 con granularidad obtenida manualmente (MANUAL + AG2)*

En este caso la eficiencia de los modelos FIR cuando se usa la función objetivo 1-Q (Véase *Tabla 30*) es superior a la eficiencia de la función objetivo  $MSE_{entr}$  (Véase *Tabla 34*), con errores  $RMS_{prue}=8.1732$  vs  $RMS_{prue}=8.8942$ . El menor error de predicción cuando se utiliza el **AG2** es significativamente inferior a los obtenidos por el resultado de referencia (C) en cada tabla, es decir, al mejor resultado obtenido cuando se utiliza FIR de la forma clásica usando máscaras subóptimas (Véase *Tabla 26*). Sin embargo, no se puede decir lo mismo de las mejores soluciones por el **AG2** (fila A de cada tabla) ya que presentan peores  $RMS_{prue}$ , como ya hemos visto anteriormente.

De la *Tabla 32*, nótese que los tiempos de ejecución de la función objetivo  $MSE_{entr}$  son muy parecidos a los de la función objetivo 1-Q usando la misma granularidad. Como ya se ha indicado, la duración de las ejecuciones para **AG2** está influenciada por la granularidad, asimismo el proceso de predicción de FIR que se lleva a cabo para la evaluación de cada partición difusa se ejecuta más rápido por la poca cantidad de datos disponibles.

#### *C) Método AG2 con granularidad suministrada por AG1 (AG1 + AG2)*

De las *Tablas 40 y 43* se concluye que la función objetivo  $MSE_{entr}$  tuvo mejor comportamiento que la función objetivo 1-Q ( $RMS_{prue}=9.0861$  vs  $RMS_{prue}=9.3463$ ), mejorando, también, los resultados obtenidos usando FIR clásico con máscaras subóptimas (Véase *Tabla 26*).

De la *Tabla 43*, para la función objetivo  $MSE_{entr}$ , se deriva que las soluciones subóptimas que poseen una máscara: (2,13,14,16), la cual incluye relaciones causales de las variables de entrada TMP y NUB con la variable de salida  $O_3$ , son las que tienen mayor habilidad de predicción.

#### *D) Método AG3*

Es interesante comentar la *Tabla 52* donde se muestran los tiempos de cómputo del **AG3** para las 3 particiones estudiadas. Como puede observarse, en la Partición

Inicial, el tiempo que necesita la función objetivo  $MSE_{entr}$  es moderadamente superior al que requiere la función objetivo 1-Q. Esa diferencia aumenta en el Caso 1 y en el Caso 2. El tiempo requerido por la función  $MSE_{entr}$ , en estos casos, duplica el necesitado por la función objetivo 1-Q. Como ya se ha mencionado anteriormente, en esta aplicación para granularidades bajas los tiempos de cómputo de las dos funciones objetivo son prácticamente equivalentes puesto que se dispone de un número reducido de datos. Sin embargo, cuando la granularidad aumenta, el tiempo requerido para evaluar las dos funciones objetivo se resiente. Esto es justo lo que sucede en los Casos 1 y 2, con respecto a la Partición Inicial como se puede comprobar en las *Tablas 49, 50 y 51*. Si sumamos el número de clases de todas las granularidades sugeridas para cada partición tenemos 263 clases para la Partición Inicial, 319 para el Caso 1 y 379 para el Caso 2. Estas diferencias justifican el costo computacional asociado.

#### E) Resumen

La *Tabla 54* resume los mejores resultados de predicción del conjunto de datos de prueba del Caso 1 logrados por las diferentes metodologías mencionadas en esta aplicación.

**Tabla 54.**  $ECM_{prue}$  de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para la **Partición Caso 1** usando: AG1+EFP, MANUAL+AG2, AG1+AG2, AG3, FIR DEFECTO, FIR MANUAL.

AG1+EFP	MAN.+AG2	AG1+AG2	AG3	FIR DEF.	FIR MAN.
8.6116	8.1732	9.0861	8,6461	No predice	9.3165

Como se observa en la *Tabla 54*, el mejor resultado se obtiene con el método **MANUAL+AG2**, seguido por **AG1+EFP** y **AG3**. En este caso las predicciones obtenidas por todas las combinaciones que hacen uso de los AGs desarrollados son mejores a las obtenidas por el **FIR MANUAL** o el **FIR DEFECTO**. Para esta partición de datos no se dispone de resultados de otras metodologías como las RNs.

Los mejores modelos FIR obtenidos por los SDEs usando la función objetivo  $MSE_{entr}$  tienen en común una relación causal de las variables de entrada TMP y NUB con la variable de salida  $O_3$ . En cambio para la 1-Q, la mayoría presentan una relación causal entre las variables de entrada TMP y VV ó solamente TMP y la variable de salida  $O_3$ . Se reitera que la temperatura es una variable esencial para la predicción de las concentraciones de ozono en esta aplicación, teniendo en cuenta las variables disponibles.

### 6.3.4.2 *Partición datos: Caso 2*

#### *A) Método AG1*

Se observa en la *Tabla 38* que para la función objetivo 1-Q, la solución óptima sugerida corresponde a la granularidad (3,2,2,2) con la más alta calidad ( $Q=0.7745$ ). Sin embargo, como ya ha sucedido otras veces, ésta no tiene la mejor habilidad de predicción. Este AG sugiere granularidades que tienen en común una partición difusa siempre de 2 clases para la variable de salida  $O_3$ .

Es interesante remarcar aquí que las calidades de las máscaras encontradas cuando se trabaja con la función objetivo  $MSE_{entr}$  son claramente inferiores a las calidades obtenidas por la función 1-Q. Esto sugiere que la medida de calidad  $Q$  definida en FIR debe ser mejorada para que represente de manera precisa la capacidad de predicción de la máscara.

#### *B) Método AG2 con granularidad obtenida manualmente (MANUAL + AG2)*

Analizando los resultados de predicción de las *Tablas 31* y *35* se observa que la eficiencia de los modelos FIR cuando se usa la función objetivo 1-Q es superior a la de la función objetivo  $MSE_{entr}$  ( $RMS_{prue}=9.6190$  vs  $RMS_{prue}=10.0253$ ). Sin embargo, ninguna de las soluciones subóptimas presenta un  $RMS_{prue}$  inferior al resultado de referencia (C) (mejor resultado de la *Tabla 26*).

Obsérvese que los tiempos de ejecución del **AG2** para ambas funciones objetivo son muy parecidos puesto que se está trabajando con una granularidad baja de (3,2,2,2) (Véase *Tabla 32*).

#### *C) Método AG2 con granularidad suministrada por AG1 (AG1 + AG2)*

Analizando las *Tablas 41* y *44*, se observa que los mejores resultados de predicción del **AG1+AG2** se obtienen usando la función objetivo  $MSE_{entr}$  ( $RMS_{prue}=9.6089$  vs  $RMS_{prue}=9.8198$ ). Sin embargo, en ninguno de los casos se logra un  $RMS_{prue}$  inferior a los mejores resultados obtenidos usando FIR clásico con máscaras subóptimas (Véase *Tabla 26*). Asimismo, siguiendo el mismo patrón observado en todos los métodos genéticos propuestos, las mejores soluciones sugeridas por el **AG2** (fila *A* de cada tabla) presentan peores  $RMS_{prue}$ .

Los mejores modelos que presentaron los  $RMS_{prue}$  más bajos en ambas funciones objetivo incluyen en la máscara óptima una relación causal de las variables de entrada TMP y NUB con la variable de salida  $O_3$ .

#### D) Método AG3

De las *Tablas 48 y 51* se observa que los mejores  $RMS_{prue}$  fueron conseguidos por la función objetivo  $MSE_{entr}$  ( $RMS_{prue}=9.8778$  vs  $10.2179$ ). Sin embargo, ambas funciones objetivo no lograron soluciones de mayor eficiencia en la predicción que las soluciones logradas en los estudios previos descritos para el mismo problema (Véase *Tabla 26*). En este sentido, la eficiencia del **AG3** no fue superior a **AG1+AG2**.

Los mejores modelos del **AG3** que presentaron los menores  $RMS_{prue}$  en ambas funciones objetivo poseen la misma máscara óptima que en **AG1+AG2**, (13,14,16), con una relación causal de las variables de entrada TMP y NUB con la variable de salida  $O_3$ .

#### E) Resumen

La *Tabla 55* resume los mejores resultados de predicción del conjunto de datos de prueba del Caso 2 logrados por las diferentes metodologías mencionadas en esta aplicación.

**Tabla 55.**  $ECM_{prue}$  de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para la **Partición Caso 2** usando: AG1+EFP, MANUAL+AG2, AG1+AG2, AG3, FIR DEFECTO, FIR MANUAL.

<b>AG1+EFP</b>	<b>MAN.+AG2</b>	<b>AG1+AG2</b>	<b>AG3</b>	<b>FIR DEF.</b>	<b>FIR MAN.</b>
10.2620	9.6190	9.6089	9.8778	11.5686	9.3343

En la *Tabla 55* se observa que los errores de predicción son superiores a los de las particiones de datos anteriores. Aquí el **FIR MANUAL** es el que logra mejor desempeño, seguido por el método **AG1+AG2** y **MANUAL+AG2**. **FIR DEFECTO** es el que obtiene peores resultados al igual que en la Partición Inicial y en el Caso 1. Asimismo, no se dispone de resultados de otras metodologías como las redes neuronales para esta partición de datos.

Los mejores modelos FIR obtenidos tienen en su mayoría una relación causal de las variables de entrada TMP y NUB con la variable de salida  $O_3$ . La temperatura sigue siendo esencial en la predicción de las concentraciones de ozono en esta aplicación.

#### 6.3.4.2 Conclusiones Generales

Para la Partición Inicial y el Caso 1 los SDEs propuestos en esta investigación han demostrado ser capaces de capturar el comportamiento dinámico del sistema bajo estudio de forma más precisa que otros métodos clásicos tales como redes neuronales (ENs, MENs, MLP y LSTM) [WW99, GNRA+03], algoritmos numéricos y aproximaciones estadísticas [Loi96], modelos IMPO [SWKK96] y modelos Santiago [AJP96]. Asimismo, han mejorado los resultados obtenidos por los modelos FIR clásicos (que no usan AGs) [GNRA+03]. Sin embargo, los errores de predicción ( $RMS_{prue}$ ) obtenidos usando los mejores modelos de los diferentes métodos mencionados son todavía bastante altos. Este es debido a que los datos disponibles son pobres tanto en cantidad como en calidad (pocas observaciones, mediciones con ruido, uso de modelos de predicción en vez de datos reales para las variables atmosféricas) y/o no se tomaron en cuenta algunas variables importantes para predecir las concentraciones de ozono.

Para la Partición Caso 2 el mejor resultado se obtiene con el **FIR MANUAL** seguido por los SDEs desarrollados en esta investigación. Cabe recordar que el **FIR MANUAL** corresponde a la opción en que la granularidad se determina a base de prueba y error y las funciones de pertenencia se definen mediante el método EFP. Por lo tanto, esta opción no es muy recomendable ni práctica. Disponer de métodos automáticos que determinen los parámetros de discretización en FIR es de vital importancia para que esta metodología pueda ser usada por usuarios inexpertos en lógica difusa que requieran modelar un sistema de complejidad elevada.

Por otro lado, es necesario señalar que los mejores modelos FIR obtenidos usando los SDEs han determinado siempre que la variable temperatura es esencial para la estimación de las concentraciones de ozono en Austria.

Finalmente es interesante remarcar que las predicciones realizadas con los mejores modelos FIR encontrados usando los SDEs propuestos en esta investigación siguen la tendencia de la señal real. Si se observan las *figuras 32 a 43* vemos que los modelos FIR son capaces de predecir de manera adecuada los valores máximos del ozono. El objetivo fundamental del modelado de contaminantes ambientales en general y del ozono en particular es predecir posibles contingencias que puedan producir consecuencias nocivas para la salud. Estas contingencias ambientales suceden cuando los contaminantes superan ciertos valores máximos, siendo pues fundamental predecir correctamente estos picos de la señal.

Las *figuras 35, 40 y 42* corresponden a la Partición Inicial. Si se observan atentamente estas figuras, se advierte que los picos más altos de la primera parte de la señal (desde julio a mediados de agosto) no son alcanzados por la señal de predicción. En otras palabras, los modelos FIR fallan básicamente en los primeros 45

puntos del conjunto de datos de prueba donde se encuentran los niveles de contaminación de ozono más dañinos.

La razón de la pobre habilidad de los modelos FIR para la predicción de la primera parte de la señal es clara. FIR es una metodología inductiva que captura la base de reglas patrón a partir de los datos observados (conjunto de identificación) y lleva a cabo la predicción desde el conocimiento que ésta contiene. Por consiguiente, FIR no es capaz de predecir comportamientos de los sistemas que no aparecen en el conjunto de identificación (entrenamiento), es decir, datos que nunca ha visto anteriormente. Este es el motivo por el cual los primeros valores de la Partición Inicial no se logran predecir correctamente, puesto que son niveles máximos de ozono que no aparecen en el conjunto de entrenamiento.

Analizando los datos del conjunto de identificación (Véase *Figura 31*), se observa que los valores de concentraciones de ozono nunca alcanzan valores por encima de 75 ppb. Sin embargo, desde principios de julio a mediados de agosto del conjunto de prueba todos los picos altos de ozono van más allá de este nivel. De este modo, FIR no es capaz de predecir este comportamiento porque no está incluido en el conjunto de identificación. Es por ello, que se realizó también un estudio adicional con los SDEs desarrollados considerando dos nuevas particiones (Caso 1 y Caso 2) que no presentan este problema.

## **CAPITULO VII:**

---

### **Problema de Predicción de Concentraciones de Ozono en la Ciudad de México**



## Capítulo VII: Problema de Predicción de Concentraciones de Ozono en la Ciudad de México

En el capítulo anterior, se analizó un caso de estudio en donde se aplicaron los métodos desarrollados en esta memoria a un problema medioambiental de gran impacto, incluso mundial, en el que intervienen muchos factores fuertemente relacionados. Dicha problemática había sido tratada por [WW99] y tenía como fin la predicción a corto plazo (un día) de las máximas concentraciones de ozono en la región Oriental Austriaca. En [WW99], los autores usaron varias topologías de redes perceptrónicas multicapas (MLPs), *Elman Networks* (EN) y *Modified Elman Networks* (MEN). Más adelante en [GNRA+03], la habilidad de la metodología FIR (usándola de la manera convencional) fue probada para predecir concentraciones de ozono usando el mismo conjunto de datos. Este estudio mostró que la eficiencia de los modelos FIR obtenidos era comparable con las redes neuronales, proporcionando resultados levemente mejores.

Por su parte, los métodos propuestos **AG1**, **AG2** y **AG3** fueron también evaluados para el mismo problema, obteniéndose una cierta mejora con respecto a los resultados obtenidos en los trabajos previos. En estos estudios se utilizó un período de muestreo de un día (valores promedios) y sólo se consideraron como variables de entrada la temperatura, la cobertura de nubes (nivel de nubosidad) y la velocidad del viento.

En este capítulo se presenta un problema de modelado medioambiental que tiene como objetivo la predicción a largo plazo de las concentraciones de ozono en la zona centro del Área Metropolitana de la Ciudad de México (AMCM)<sup>16</sup> y se estudia el comportamiento de los métodos propuestos en esta memoria para su resolución. Los datos disponibles para este estudio fueron registrados cada hora. Las señales de entrada consideradas fueron las principales variables atmosféricas: la humedad relativa, la temperatura del medio ambiente, la velocidad del viento y la dirección del viento además de la variable hora del día. El número de variables analizadas, la alta frecuencia de las señales y el hecho de que se pretenden abordar predicciones a largo plazo, incrementa la complejidad de la aplicación. Una dificultad adicional estriba en

---

<sup>16</sup> El área metropolitana de la Ciudad de México (AMCM) comprende el territorio de las 16 demarcaciones territoriales del Distrito Federal y de los municipios conurbanos del Estado de México. Está constituida por cinco zonas: Dos en el norte (Xalostoc y Tlanepilantla), una en el centro (Merced) y dos en el Sur (Pedregal y Cerro de la Estrella).

la presencia de valores perdidos (*missing values*) tanto en los datos meteorológicos como en los del ozono.

La estructura de este capítulo es la siguiente: se comenzará con una descripción del problema, estudiando después las variables que se van a considerar, el proceso de generación de los datos y la división de éstos en los conjuntos de entrenamiento y de prueba. A continuación, se mostrarán los resultados obtenidos por los métodos propuestos **AG1**, **AG2** y **AG3**. Finalmente, se analizarán los mismos comparándolos con los de otros métodos tal como se ha hecho en el capítulo anterior y se darán algunas conclusiones.

## 7.1 Descripción del Problema

El contaminante ambiental es una sustancia del aire, agua o suelo que cuando está presente en exceso disminuye la calidad o capacidad de la vida. Están presentes en el ambiente en cantidades que no dañan a plantas o animales. Sin embargo, cuando se excede determinado nivel constituyen un peligro real o potencial. Este nivel, por debajo del cual no tienen efectos nocivos, se determina con fundamento en antecedentes científicos.

Los contaminantes del aire se pueden clasificar de acuerdo a su tamaño, origen, estado físico o naturaleza química. Pueden ser átomos; moléculas o compuestos orgánicos o inorgánicos; iones o radicales libres; mezclas o agregados; partículas o emulsiones físicas, químicas o biológicas de origen natural o artificial generados por fuente móvil o estacionaria, primarias o secundarias, extramuros o intramuros.

Los volcanes, por ejemplo, son fuentes naturales que emiten óxidos de azufre ( $\text{SO}_2$ ) y partículas; las tormentas eléctricas que generan ozono ( $\text{O}_3$ ) y óxidos de nitrógeno ( $\text{NO}_2$ ); los animales que producen bióxido de carbono ( $\text{CO}_2$ ) e hidrogeniones ( $\text{H}^+$ ) y la erosión, deforestación y sequías que generan partículas esparcidas por el viento.

Entre los ejemplos de fuentes artificiales están algunas industrias que emiten  $\text{NO}_2$  y  $\text{SO}_2$ , así como partículas. Los motores de combustión interna que emiten monóxido de carbono ( $\text{CO}$ ), bióxido de nitrógeno ( $\text{NO}_2$ ) e hidrocarburos ( $\text{HC}$ ) y producen ruido. La combustión de carbón, leña o basura que genera  $\text{CO}$ , partículas y  $\text{SO}_2$ . Otros contaminantes son los sulfatos ( $\text{SO}_4^-$ ), los nitratos ( $\text{NO}_3^-$ ), el ácido nítrico ( $\text{HNO}_3$ ) y el ácido sulfúrico ( $\text{HSO}_4$ ) que resultan principalmente de la combustión de carbón y biomasa.

Por otro lado, el humo producido por la combustión del tabaco es el contaminante más importante intramuros y genera partículas suspendidas de naftaleno, fenantreno, antraceno, pyreno, fenol, nitrosamina, anilina, varios ácidos grasos e hidrocarburos aromáticos policíclicos y de metales pesados como el plomo. Algunos se han comprobado científicamente que son cancerígenos como los bencenos, las nitrosaminas, el formaldehído y el etil-carbamato. También genera gases, algunos de ellos dañinos como el monóxido de carbono, el cianuro de hidrógeno, el amonio, el acetileno, el ácido fórmico, el formaldehído, el ácido acético, la acetona y el acetaldehído, entre otros [Hub89, RSP00, RRT05].

No es común encontrar un contaminante aislado, generalmente se observan en forma de mezclas o agregados, además es importante considerar las interacciones de los contaminantes del aire con otros factores como por ejemplo, la humedad o la radiación solar. Es por ello, que los contaminantes del aire también se pueden clasificar en primarios y secundarios. Los primarios son los que permanecen en la atmósfera tal y como fueron emitidos por la fuente. Para fines de la evaluación de la calidad de aire se consideran: óxidos de azufre ( $\text{SO}_x$ ), óxido de carbono (CO), óxidos de nitrógeno ( $\text{NO}_x$ ), hidrocarburos y partículas. Los secundarios son los que han estado sujetos a cambios químicos, o bien, son el producto de la reacción de dos o más contaminantes primarios en la atmósfera. Entre ellos destacan los oxidantes fotoquímicos y algunos radicales de corta existencia como el  $\text{O}_3$ .

El ozono llamado también "smog" fotoquímico, se forma en la baja atmósfera mediante la reacción química de compuestos volátiles orgánicos (VOC) y bióxido de nitrógeno ( $\text{NO}_2$ ) en presencia de la luz solar. Es por ello que las temperaturas elevadas estimulan la reacción fotoquímica. El ozono es un gas compuesto de tres átomos de oxígeno y es el resultado de la cantidad de combustible que consumen los vehículos, industrias y servicios de cualquier centro urbano. La creación del ozono depende de la cantidad y tipo de fuentes emisoras, la localización geográfica del área y de los asentamientos humanos y fabriles. El ozono produce efectos tóxicos porque fácilmente quita electrones a otras moléculas, iniciando reacciones en cadena y perturbando estructuras vitales en las células ocasionando daños que afectan la salud de la población en general, tales como dolores de cabeza, problemas respiratorios, irritaciones en los ojos, nariz y garganta, así como el empeoramiento de las enfermedades cardiovasculares.

Como ya se comentó en la sección 6.3, existen distintos modelos disponibles para la predicción de la calidad del aire a escala local y para la previsión de niveles de ozono. Muchos de ellos se basan en aproximaciones estadísticas [CJ01, SKM02] mientras que otros se sustentan en técnicas que se enmarcan en el área de la inteligencia artificial, tales como las redes neuronales [Com97, AA02, WW99, LHP+04, YB06, Ghi05, MV03, WL06], los sistemas difusos [LHP+04, Ghi05,

HK04, PDP+00, ODH04, MV03, MYS05, GNMW01, GNRA+03], los árboles de decisión o las reglas de asociación [WW01, RHBF03].

### 7.1.1 La Ciudad de México

México es uno de los países más contaminado del mundo, el 65% de las emisiones de contaminantes es de origen vehicular. Su capital, Ciudad de México, concentra casi la mitad de las industrias del país y es considerada como una de las ciudades que posee mayor población a nivel mundial.

La Ciudad de México se encuentra ubicada en una cuenca a 2240 m de altura sobre el nivel del mar. El clima se caracteriza por madrugadas frías y medio días calurosos. Como sucede en valles similares, durante la noche la temperatura en las laderas de las montañas aumenta a medida que se asciende, lo cual guarda una relación inversa al comportamiento habitual. Este fenómeno determina la inmovilidad de la capa de aire frío próxima al suelo, la cual concentra los contaminantes generados por la actividad industrial y el tránsito vehicular del día anterior. Esta capa se empieza a movilizar cuando el sol la calienta; se eleva, desaparece la inversión térmica y se dispersan los contaminantes.

Es por ello, que el aire aparece hoy día como el principal objeto de preocupación entre los distintos problemas ambientales que se enfrentan en la Ciudad de México [Lez00]. Este problema comenzó a ser relevante a partir de 1940, cuando el carácter centralista del proceso de desarrollo desencadenó una aglomeración industrial y poblacional. Esta última variable llegó a presentar tasas de crecimiento superiores al 3% anual [Her90, CCG99].

Actualmente, se calcula que en un año la ciudad arroja al aire alrededor de cinco millones de toneladas de contaminantes, entre los que destacan los NO<sub>2</sub>, los SO<sub>2</sub>, los óxidos de carbono, los hidrocarburos aromáticos como el benceno y una gran variedad de partículas de origen, naturaleza y tamaño diverso, algunas de las cuales, menores de 10 µg de diámetro, permanecen suspendidas en el aire y son factibles de ser respiradas y depositadas en las vías aéreas [Per92, MP03, MM02].

La elevada radiación ultravioleta sobre el valle de México que, se explica por el frecuente número de días soleados y la menor densidad y grosor de la capa protectora de aire, determina la transformación de las moléculas de bióxido de nitrógeno (NO<sub>2</sub>) y el HC, en O<sub>3</sub>. Es por ello, que esto hace que el problema de contaminación sea muy complejo aún cuando se han implementado medidas para tratar de aminorarla en estos últimos años, tales como cambios en el uso de los carburantes del parque automotor e

implementación de diversos programas: "Hoy no circula"<sup>17</sup>, "Contingencias ambientales"<sup>18</sup>, "Verificación Vehicular"<sup>19</sup>, entre otros; y se estén haciendo esfuerzos institucionales para concienciar a la ciudadanía y fomentar la investigación en este campo.

De todo lo anterior se desprende que el principal problema de contaminación del aire que ha sido identificado en el Área Metropolitana de la Ciudad de México (AMCM) es la formación del "smog" fotoquímico. Cabe destacar que los niveles de ozono en la Ciudad de México siguen siendo altos actualmente aunque se ha reducido en los últimos años. Distinto de las mayorías de las ciudades en el hemisferio norte donde el fenómeno ozono de la troposfera es solamente presentado durante los días de verano, la AMCM presenta condiciones favorables para la formación de O<sub>3</sub> a lo largo de todo el año [Mau94]. Por lo tanto, para que las autoridades de ésta o cualquier centro urbano puedan tomar decisiones y salvaguardar la salud de su ciudadanía, requieren de modelos capaces de capturar el comportamiento de cualquier contaminante con la finalidad de predecir posibles contingencias ambientales<sup>20</sup> con la debida anticipación.

Hay que tener en cuenta, pues, que el problema más grande de la calidad del aire en los centros urbanos es la contaminación por ozono, ya que a su vez, resulta ser uno de los contaminantes de más difícil predicción. Es por ello que para poder advertir a tiempo y de manera adecuada a la población, es importante tener previsiones lo más precisas y fiables posibles de los niveles altos futuros de ozono. Por consiguiente, la construcción de modelos de ozono que capturen de la manera más precisa posible el comportamiento de éste contaminante en la atmósfera y permitan generar un pronóstico fiable de los niveles máximos de ozono en el aire, resultan ser el principal interés no sólo de los científicos medioambientales sino de las agencias gubernamentales.

---

<sup>17</sup> De acuerdo al último dígito de la placa permanente de circulación del vehículo automotor no puede transitar ciertos días de la semana

<sup>18</sup> A partir del 01/09/2006 se pone en vigencia el "Decreto por el que se reforma y adiciona el Programa para Contingencias Ambientales Atmosféricas (PCAA)". Para más información consulte la Gaceta Oficial del Distrito Federal: [http://www.sma.df.gob.mx/simat/pdf/gaceta\\_df\\_300806.pdf](http://www.sma.df.gob.mx/simat/pdf/gaceta_df_300806.pdf), publicada el 30/08/2006.

<sup>19</sup> Se encarga de establecer el calendario y los lineamientos conforme a los cuales todos los vehículos deberán ser verificados, en sus emisiones de contaminantes.

<sup>20</sup> Entiéndase por contingencia ambiental a la situación eventual y transitoria, declarada por las autoridades competentes, en que la concentración de contaminantes en la atmósfera alcanza niveles dañinos para la salud de la población en general.

### 7.1.2. Sistema de Monitoreo Atmosférico (SIMAT)

En diciembre del año 2000 la Secretaría del Medio Ambiente del Gobierno del Distrito Federal<sup>21</sup>, integró el Sistema de Monitoreo Atmosférico de la Ciudad de México (SIMAT), con la finalidad de unificar los esfuerzos destinados a la medición de contaminantes atmosféricos y parámetros meteorológicos.

El objetivo del SIMAT es vigilar y evaluar el estado de la calidad del aire en la Ciudad de México, como medida de prevención para protección de la salud de sus habitantes. Asimismo, busca consolidarse ante los habitantes de la Ciudad de México, como un organismo confiable, eficaz, honesto, innovador y transparente en los servicios de información de la calidad del aire. Es por ello, que continúan trabajando con la aplicación de procedimientos de gestión de mejora continua.

**Tabla 56.** Relación de parámetros registrados por las estaciones del SIMAT, 2006.

RED	EQUIPOS	PARÁMETROS
RAMA	36	Ozono (O <sub>3</sub> ), dióxido de azufre (SO <sub>2</sub> ), óxidos de nitrógeno (NO <sub>2</sub> , NO <sub>x</sub> y NO), monóxido de carbono (CO), partículas menores a 10 micrómetros (PM <sub>10</sub> ), partículas menores a 2.5 micrómetros (PM <sub>2.5</sub> ) y ácido sulfhídrico (H <sub>2</sub> S).
REDMET	15	Temperatura ambiente (TMP), velocidad de viento (VV), dirección del viento (DV), humedad relativa (HR) y radiación solar UV-B (UVb).
REDMA	14	Partículas suspendidas totales (PST), PM <sub>10</sub> , PM <sub>2.5</sub> , nitratos (NO <sub>3</sub> <sup>-</sup> ), sulfatos (SO <sub>4</sub> <sup>2-</sup> ) y metales pesados (Pb, Cd, Cu, Fe, V, Mn, Ni).
REDDA	16	pH, conductividad específica, aniones (NO <sub>3</sub> <sup>-</sup> , SO <sub>4</sub> <sup>2-</sup> , Cl <sup>-</sup> y CO <sub>3</sub> <sup>2-</sup> ) y cationes (NH <sub>4</sub> <sup>+</sup> , Na <sup>+</sup> , Ca <sup>2+</sup> , Mg <sup>2+</sup> y K <sup>+</sup> ).

Para lograr su propósito, el SIMAT cuenta con varias estaciones de monitoreo por medio de las cuales se registran diferentes parámetros en el Área Metropolitana de la Ciudad de México. Dichas estaciones se clasifican de acuerdo a la forma en que operan y al parámetro o parámetros que miden, conformando de ese modo las cuatro redes de monitoreo o subsistemas operativos especializados que integran el SIMAT (Véase *Tabla 56*).

<sup>21</sup> <http://www.sma.df.gob.mx/sma/>

### 7.1.3 El Índice Metropolitano de la Calidad del Aire (IMECA)

En México, el Índice Metropolitano de la Calidad del Aire (IMECA) es un valor de referencia que permite a la población conocer los niveles de contaminación prevalentes en su zona de residencia, de manera precisa y oportuna, pudiendo de este modo tomar las medidas pertinentes de protección. Se ha asignado una escala de 0 a 500 para cada uno de los contaminantes. La *Tabla 57* presenta la interpretación del IMECA, se observa que 100 puntos equivalen al nivel permitido, tanto en concentración como en duración. Por debajo de 100, en un tiempo determinado, no existe riesgo para la integridad de la vida vegetal y animal; en cambio, concentraciones superiores a 100 en un tiempo determinado representan un riesgo cada vez mayor.

**Tabla 57.** Interpretación del IMECA.

IMECA	Condición	Efectos a la salud
0-100	Condición dentro de la norma	Ninguno
101-200	Condición no satisfactoria	Molestias en ojos, nariz y garganta en personas sensibles
201-300	Condición mala	Evitar actividades al aire libre. Posibles problemas respiratorios
301-500	Condición muy mala	Se agudizan los síntomas anteriores en personas sensibles y quienes fuman o padecen de enfermedades crónicas

De todos los contaminantes del aire, tanto las investigaciones nacionales como internacionales han mostrado fehacientemente que las principales partículas que ejercen un efecto nocivo sobre la salud respiratoria son el O<sub>3</sub> y las PM<sub>10</sub>.

## 7.2 Trabajos previos

En [NM06] se realizó una investigación usando la plataforma *Visual-FIR* basada en la metodología FIR para la identificación de modelos de concentraciones de ozono a largo plazo en la zona Centro del AMCM. Los datos usados para este estudio fueron tomados del Sistema de Monitoreo Atmosférico (SIMAT), que mide contaminantes y variables atmosféricas desde varias estaciones distribuidas a lo largo de las cinco zonas del AMCM. Las variables registradas son los contaminantes

atmosféricos: O<sub>3</sub>, SO<sub>2</sub>, NO<sub>2</sub>, CO y PM<sub>10</sub>, así como las variables meteorológicas: temperatura, humedad relativa, velocidad del viento y dirección del viento. Estas variables se muestrean las 24 horas del día, cada día del año. Siempre que por alguna razón, un instrumento de medición de una de las estaciones falle, se produce un conjunto de valores perdidos (*missing values*).

Esta investigación está dirigida al modelado de un sólo contaminante en la zona centro (MERCED) del AMCM. El contaminante elegido fue el O<sub>3</sub> por ser el problema de contaminación más severo, aunque resulta más complicado de tratar por ser un contaminante secundario. La zona elegida fue la central debido a que allí se concentra el mayor asentamiento poblacional, y por ende, cuenta con la mayor actividad humana durante los días y las horas laborables, siendo más propensa a un elevado índice de contaminación. De este modo, el ozono (O<sub>3</sub>) es la variable de salida del sistema y está medido en partes por mil millón (ppm). Las variables de entrada son: a) la temperatura (TMP), medida en °C; b) la humedad relativa (HR), establecida en porcentaje (0% a 100%); c) la velocidad del viento (VV), determinada en metros por segundo (m/s); d) la dirección del viento (DV), medida en grados (desde 0° a 359°); y e) la hora del día (HD), desde 1 a 24. La página web del SIMAT<sup>22</sup> ofrece una base de datos de registros meteorológicos y de contaminantes desde 1986 hasta la actualidad. Para este estudio, se utilizaron datos medidos desde Enero 2000 hasta Febrero 2006.

La raíz del error cuadrático medio, *Root Mean Square Error* (RMS), se usó para determinar la validez de cada uno de los modelos y viene dado por:

$$RMS = \sqrt{\frac{\sum_{i=1}^N (y_i(t) - \hat{y}_i(t))^2}{N}} \quad (51)$$

donde:  $y(t)$  es la salida del sistema,  $\hat{y}(t)$  es la salida predicha y  $N$  denota el número de observaciones.

En este trabajo se definieron dos perspectivas de modelado. El primero buscaba el modelado del comportamiento del ozono durante un mes específico del año. El segundo estaba centrado en el modelado del ozono en cierta estación de tiempo del año. Existen dos períodos estacionales bien diferenciados en la Ciudad de México: a) Estación seca: está comprendida aproximadamente de noviembre a abril, y b) Estación lluviosa: la cual va desde mayo a octubre. Todas las variables, excepto HD y O<sub>3</sub>, fueron discretizadas en dos clases usando el método EFP. Las variables HD y O<sub>3</sub> fueron discretizadas manualmente en tres clases siguiendo las recomendaciones de los expertos ambientales. Los landmarks que se obtuvieron para cada variable se muestran en la *Tabla 58* y fueron usados en ambas perspectivas de modelado.

<sup>22</sup> <http://www.sma.df.gob.mx/simat/>

**Tabla 58.** Landmarks usados por [NM06] en el proceso de fusificación para las variables de entrada y de salida.

	<b>HD</b>	<b>HR</b>	<b>TMP</b>	<b>DV</b>	<b>VV</b>	<b>O<sub>3</sub></b>
Clase 1	1..12	6..64	10.7..16.7	0..124	0.06..1.40	0.0..0.05
Clase 2	12..17	64..93	16.7..28.4	124..360	1.40..4.82	0.05..0.10
Clase 3	17..24	-----	-----	-----	-----	0.10..0.20

En esta memoria sólo serán evaluados los métodos **AG1**, **AG2** y **AG3** usando la primera perspectiva, es decir, el modelado del O<sub>3</sub> durante un mes específico del año. En [NM06] fueron identificados dos modelos mensuales, el primero de ellos “Enero” corresponde a la estación seca y el segundo “Agosto” correspondiente a la estación lluviosa. En nuestra investigación realizaremos, también, ambas identificaciones usando los SDEs desarrollados.

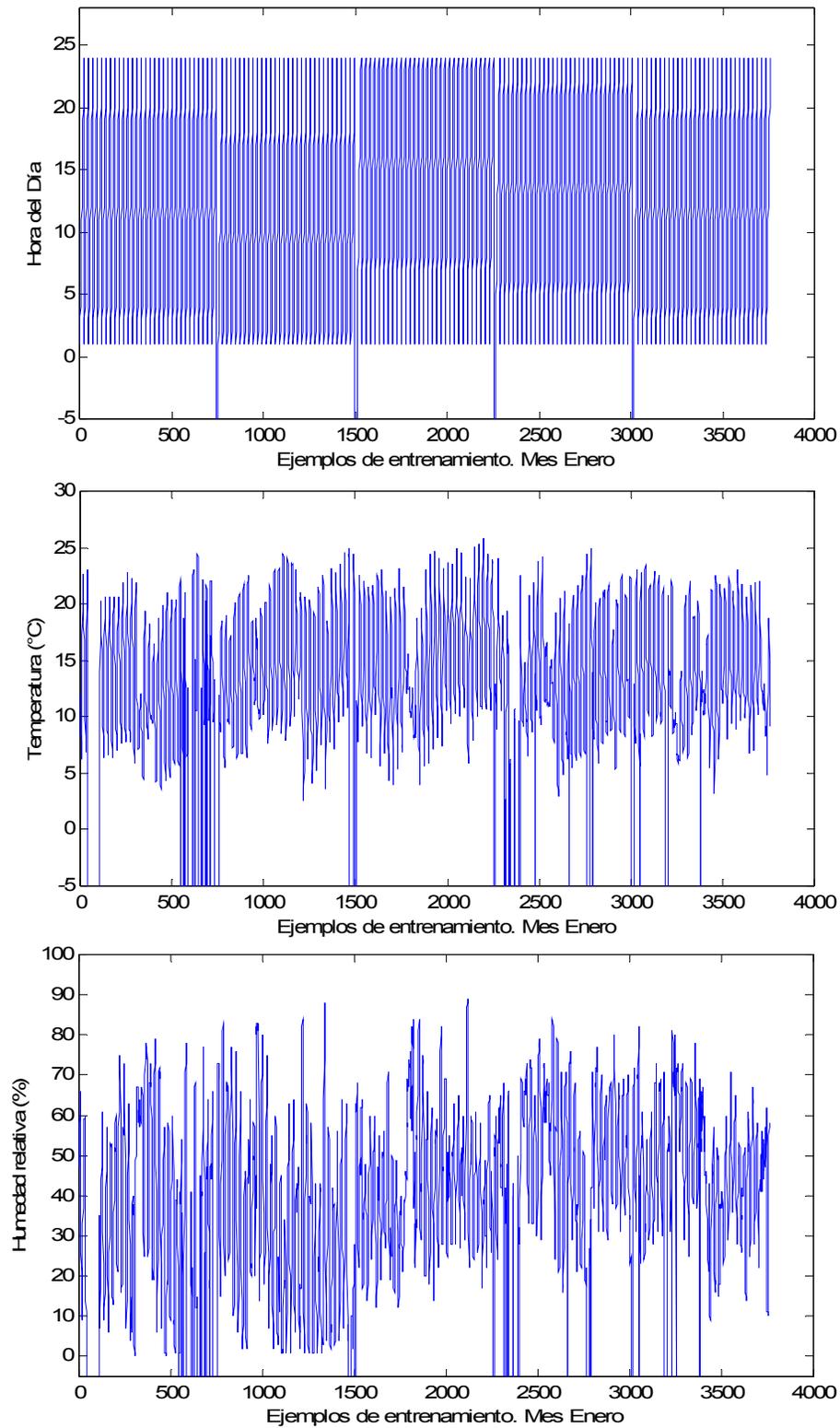
### 7.2.1 Modelo del mes Enero

En esta sección se muestran los modelos FIR publicados en [NM06] que capturan el patrón de comportamiento del contaminante O<sub>3</sub> en el mes de Enero. Este mes corresponde a la estación seca en la Ciudad de México y los niveles de O<sub>3</sub> son usualmente altos. Los datos registrados en Enero de los años 2001 a 2005 (ambos inclusive) se usaron como conjunto de entrenamiento mientras que Enero del año 2006 se usó como conjunto de prueba para mostrar la eficiencia del modelo identificado. El conjunto de datos disponible contiene en total 3720 registros, en los cuales hay 318 valores perdidos. Los datos perdidos están distribuidos a lo largo de todas las variables a excepción de la hora del día. Para evitar la generación de relaciones inexistentes entre las variables, se añadió un conjunto de datos abarcando cuatro instantes de tiempo con valores perdidos en la concatenación de los meses de Enero de los diferentes años. Las *figuras 44* y *45* muestran las señales de entrenamiento (con valores perdidos) de las variables de entrada (HD, HR, TMP, VV y DV) y de la variable de salida O<sub>3</sub> para el mes de Enero.

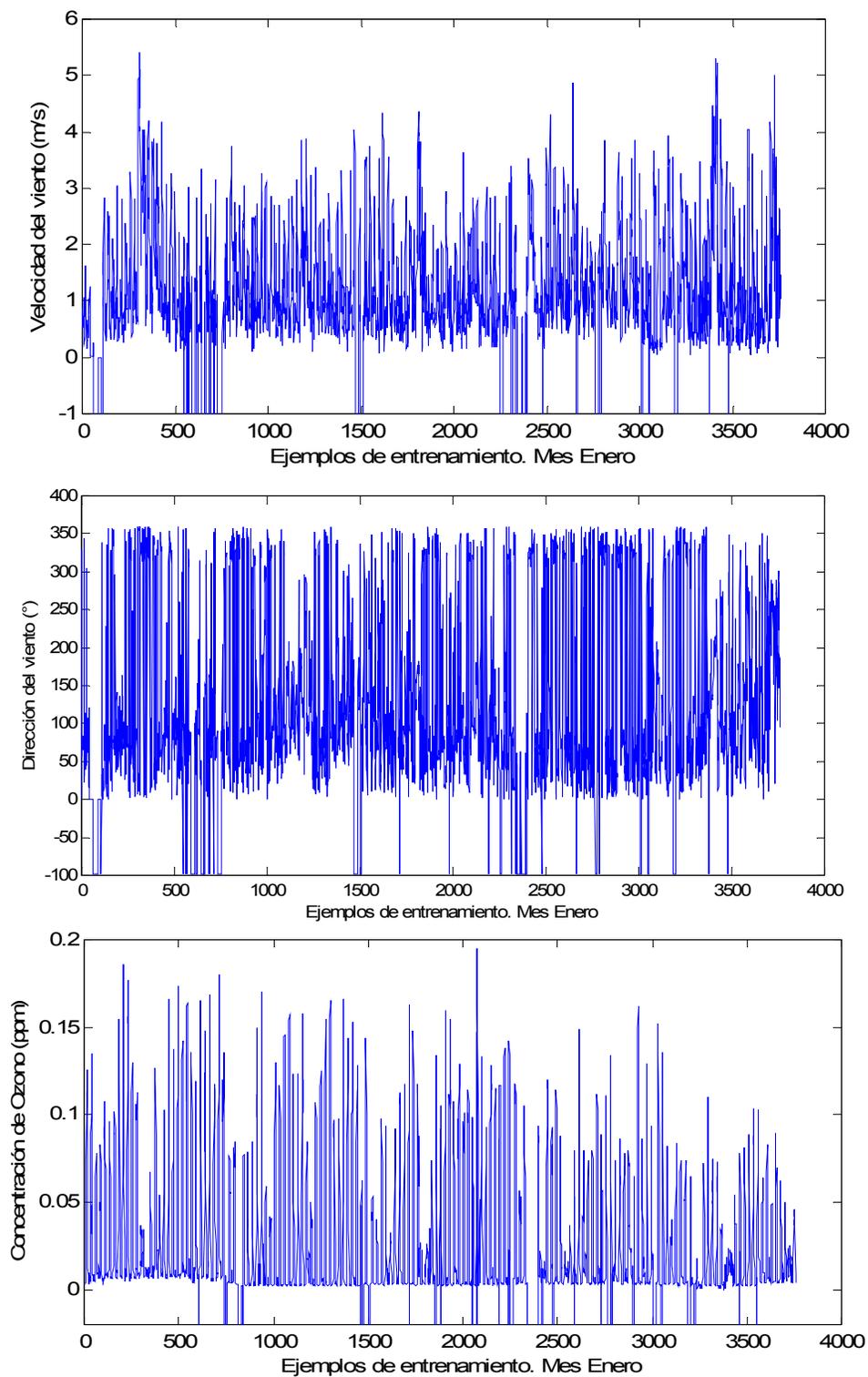
**Tabla 59.** Máscaras obtenidas por [NM06] para el modelo mensual **Enero**.

<b>Máscara (en posición de notación)</b>	<b>Complejidad</b>	<b>Q</b>	<b>RMS<sub>prue</sub></b>
(1,6,10,11,18)	5	0.5850	0.0176
(1,10,12,18)	4	0.7390	0.0185

La *Tabla 59* presenta el conjunto de las mejores máscaras obtenidas para diferentes complejidades usando *Visual-FIR*.



**Figura 44.** Señales de entrenamiento para las variables de entrada HD, TMP y HR. Mes Enero.  
Nota: Los valores negativos corresponden a valores perdidos.



**Figura 45.** Señales de entrenamiento para las variables de entrada  $VV$ ,  $DR$  y la variable de salida  $O_3$ . **Mes Enero.** Nota: Los valores negativos corresponden a valores perdidos.

La primera columna de la tabla contiene las máscaras en notación de posición. La segunda y tercera columnas muestran la complejidad y la calidad asociada a esa máscara. La última columna indica el error RMS cuando esa máscara se usa para predecir el conjunto de datos de prueba. Se notó que las variables hora del día, dirección del viento y ozono eran seleccionadas en todas las máscaras obtenidas, lo cual significa que estas variables son muy relevantes. La variable velocidad del viento se consideró también importante para incrementar la precisión de la predicción. Por otro lado, se observó que la temperatura y la humedad relativa no eran seleccionadas, lo cual indicaba que no tienen influencia significativa en el comportamiento del ozono durante la estación seca.

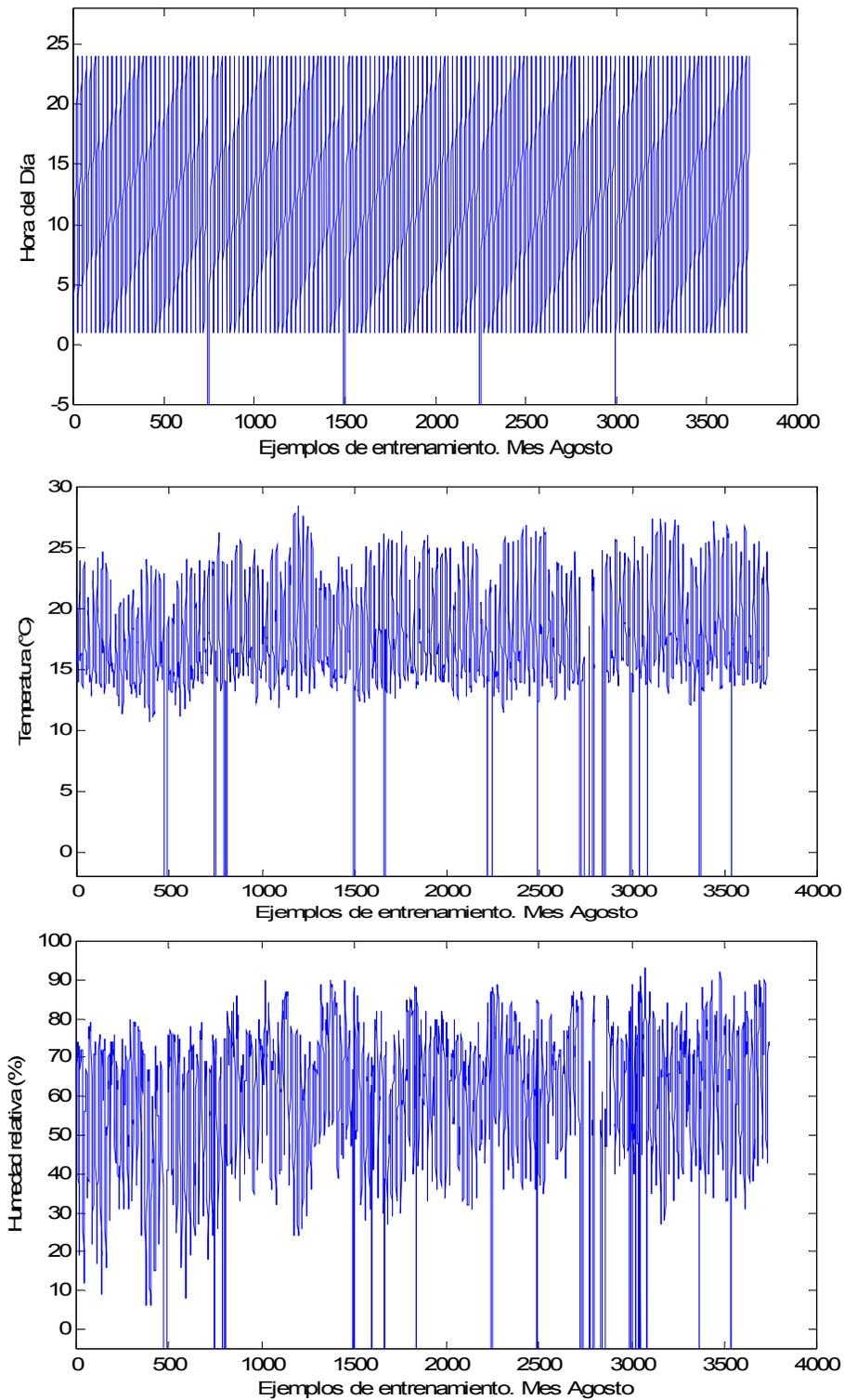
### 7.2.2 Modelo del mes Agosto

En esta otra sección se muestran los modelos FIR publicados en [NM06] que capturan el patrón de comportamiento del contaminante  $O_3$  en el mes de Agosto, el cual corresponde a la estación lluviosa de la Ciudad de México. Los niveles de  $O_3$  no son tan altos como en la temporada seca pero están aún por encima de 100 ppm. Los datos registrados en Agosto de los años 2000 a 2004 (ambos inclusive) se usaron como conjunto de entrenamiento mientras que Agosto del año 2005 se usó como conjunto de prueba para mostrar la eficiencia del modelo identificado. El conjunto de datos disponible contiene en total 3720 registros, en los cuales hay 153 valores perdidos. Como antes, un conjunto de valores perdidos abarcando cuatro instantes de tiempo se incluyó para separar los datos de dos meses de Agosto de años consecutivos con el fin de evitar relaciones no reales entre las variables. Las *figuras 46 y 47* muestran las señales de entrenamiento (con valores perdidos) de las variables de entrada (HD, HR, TMP, VV y DV) y de la variable de salida  $O_3$  para el mes de Agosto.

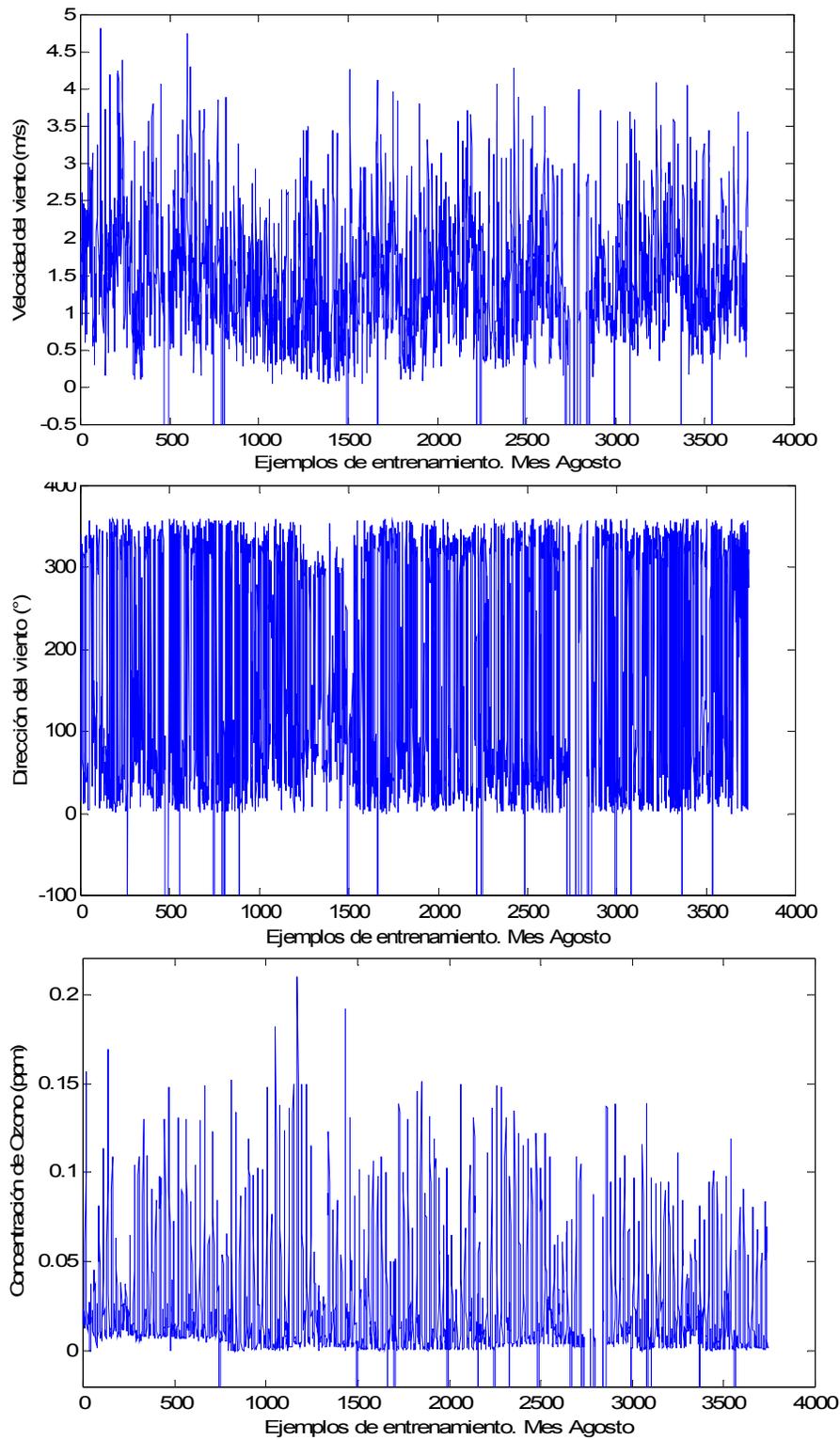
**Tabla 60.** Máscaras obtenidas por [NM06] para el modelo mensual Agosto.

Máscara (en posición de notación)	Complejidad	Q	RMS <sub>prue</sub>
(1,5,8,12,18)	5	0.1960	0.0180
(1,6,8,18)	4	0.3260	0.0189

La *Tabla 60* presenta el conjunto de mejores máscaras obtenidas para diferentes complejidades en el proceso de la máscara óptima de *Visual-FIR*. La tabla fue estructurada como la *Tabla 59*. Como sucedió en los modelos de Enero, instantes previos de las variables ozono, hora del día y velocidad del viento llegan a ser cruciales para la predicción. Sin embargo, ahora la humedad relativa apareció como un componente relevante en vez de la dirección del viento. Esto resultó ser un interesante hallazgo que tiene mucho sentido debido a que Agosto está dentro de la



**Figura 46.** Señales de entrenamiento para las variables de entrada HD, TMP y HR. Mes Agosto. Nota: Los valores negativos corresponden a valores perdidos.



**Figura 47.** Señales de entrenamiento para las variables de entrada  $VV$ ,  $DR$  y la variable de salida  $O_3$ . **Mes Agosto.** Nota: Los valores negativos corresponden a valores perdidos.

estación lluviosa, y por lo tanto, la presencia de la humedad relativa en el aire influye directamente en los niveles de ozono. Los errores de predicción de los modelos de Agosto fueron similares a los obtenidos por los modelos de Enero, siendo ambos realmente bajos.

### 7.3 Modelos obtenidos por los métodos *AG1*, *AG2* y *AG3*

Del mismo modo que en las experimentaciones del capítulo anterior, en este caso de estudio se han realizado 30 ejecuciones (con distinto valor de la semilla) para cada función objetivo y criterio de parada en cada método genético planteado en esta memoria. Del mismo modo, se estudian dos funciones objetivo diferentes: a) 1-Q, y b) el error de predicción MSE de una parte del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). En esta aplicación medioambiental el último 25% de la señal de entrenamiento se usa para la evaluación de la función objetivo  $MSE_{entr}$  y el 75% inicial de la señal se usa para la obtención de los modelos FIR (máscaras y base de reglas patrón).

**Tabla 61.** Parámetros de los AGs para el problema de la estimación de concentraciones locales máximas de ozono en México DF.

Parámetro	Tipo AG	Valor
<i>Tamaño de la población</i> (# individuos)	AG1	9
	AG2, AG3	50
<i>Probabilidad de cruce</i>	Todos	0.6
<i>Probabilidad de mutación</i>	Todos	0.1
<i>Criterio de parada</i> (# evaluaciones de cromosomas)	AG1	{100, 200, 400, 800}
	AG2	{1000, 2000, 4000, 8000}
	AG3	{1000, 2000, 4000, 8000}

Como ya se ha mencionado anteriormente se ha trabajado con la perspectiva de modelado del comportamiento del ozono durante un mes específico del año considerada por [NM06]. Es así como se seleccionaron dos meses de períodos estacionales diferentes, el mes Enero que pertenece a la época seca y el mes de Agosto a la lluviosa.

$t \setminus x$	HD	HR	TMP	DV	VV	03
$t - 2\delta t$	-1	-1	-1	-1	-1	-1
$t - \delta t$	-1	-1	-1	-1	-1	-1
$t$	-1	-1	-1	-1	-1	+1

(52)

En la ecuación 52 se muestra la máscara candidata de profundidad 3 (tres filas) utilizada, la cual permite encontrar las relaciones causales y temporales entre las variables de la aplicación bajo estudio. Los elementos -1 de la máscara candidata representan las relaciones causales y temporales potenciales con la variable de salida (elemento +1).

Los valores establecidos para los parámetros de los AGs se presentan en la *Tabla 61*. Las probabilidades de cruce y de mutación se establecen de acuerdo a [Gre86]. Se empleó un cluster del Laboratorio de Cálculo LSI que contiene ocho procesadores Pentium IV (3.2 GHz) para realizar las simulaciones correspondientes a los métodos **AG2** y **AG3**, mientras que para **AG1** se utilizó un PC Pentium Centrino (1.60HZ). En [ANF06b] se muestran todos los resultados obtenidos para los métodos **AG1**, **AG1+AG2**, **EXPERTOS+AG2** y **AG3**. Los mejores resultados obtenidos por los métodos propuestos se presentan en las *Tablas 62 a 79*.

A continuación se muestran los resultados en cuatro sub-apartados denominados en base al método empleado para la resolución del problema.

### **7.3.1 Aprendizaje de las funciones de pertenencia de las clases (AG2). Granularidad (3,2,2,2,2,3) suministrada por expertos en estudios previos**

En éste bloque se presentan los resultados obtenidos para el problema de estimación de concentraciones de ozono en México por el método **AG2**. Se considera la granularidad (3,2,2,2,2,3) suministrada por expertos en [NM06], por lo tanto, se tiene la combinación **EXPERTOS+AG2**.

Las *Tablas 62 y 63* revelan los resultados del **AG2** para la granularidad (3,2,2,2,2,3) cuando se usa la función objetivo 1-Q para los meses de Enero y Agosto, respectivamente. Las tablas están organizadas como se especifica a continuación. La primera columna se divide en tres secciones. La sección *A* corresponde a los mejores resultados obtenidos por la combinación **EXPERTOS+AG2**, es decir mínimo valor de la función objetivo. La sección *B* corresponde a los mejores resultados de predicción del conjunto de datos de prueba obtenidos por **EXPERTOS+AG2**. La sección *C* muestra los mejores resultados de predicción obtenidos por FIR en [NM06] cuando éste AG no es usado, es decir, cuando se usa el método EFP para discretizar todas las variables (a excepción de HD y O<sub>3</sub>) y se usan máscaras subóptimas (Véanse *Tablas 59 y 60*). De esta forma, *C* se usa como referencia para el análisis de los resultados. La segunda columna indica el criterio de parada, es decir, el número de evaluaciones de cromosomas realizadas por el AG. La tercera columna muestra la proporción de datos para las variables de entrada (HD,HR,TMP,DV,VV) y la variable de salida (O<sub>3</sub>). El número de elementos de la proporción de datos corresponde al número de clases para cada variable (mostradas en la cabecera de cada tabla). La

proporción de datos es la salida del **AG2** y el segundo parámetro de entrada del proceso de fusificación de la metodología FIR. La cuarta columna presenta la máscara óptima (en notación de posición) encontrada por FIR cuando se usa la proporción de datos obtenida para determinar los landmarks. La quinta columna corresponde a la calidad asociada a la máscara óptima. La sexta columna es el valor de la función objetivo 1-Q. La última columna muestra el valor del error  $RMS_{prue}$ .

**Tabla 64.** Tiempo consumido por el EXPERTOS + AG2 para realizar las treinta ejecuciones para cada criterio de parada. Caso Estudio: Ozono México-Mes Enero. Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas) por función objetivo	
	1-Q	$MSE_{entr}$
1000	39:17	47:44
2000	77:09	94:15
4000	152:53	186:38
8000	315:13	371:40

En la *Tabla 64* se presenta el tiempo necesario de CPU para alcanzar las treinta ejecuciones en el modelado de concentraciones de ozono en México durante el mes de Enero haciendo uso de la combinación **EXPERTOS + AG2** y considerando cada uno de los criterios de parada estipulados y funciones objetivo. Hay que señalar que para el mes de Agosto el tiempo consumido ha sido del mismo orden de magnitud.

Por otro lado, las *Tablas 65 y 66* muestran los resultados para la función objetivo  $MSE_{entr}$ . Estas tablas están organizadas como la *Tabla 62*. La diferencia estriba en que la sexta columna representa los valores de la función objetivo  $MSE_{entr}$  en vez de 1-Q.

**Tabla 62.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable proporcionado por expertos en [NM06]). **Mes Enero.**

Núm. Clases: (3,2,2,2,2,3)						
# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>	
A	8000	HD:(0.51,0.23,0.26) HR:(0.18,0.82) ; TMP: (0.49,0.51) DV:(0.46,0.54) ; VV:(0.64,0.36) 03:(0.07,0.88,0.05)	(5,7,14,17,18)	0.8911	0.1089	0.0176
	4000	HD:(0.52,0.42,0.06) HR:(0.47,0.53) ; TMP:(0.23,0.77) DV:(0.40,0.60) ; VV:(0.68,0.32) 03:(0.09,0.86,0.05)	(12,17,18)	0.8793	0.1207	0.0659
	2000	HD:(0.57,0.17,0.26) HR:(0.47,0.53) ; TMP: (0.88,0.12) DV:(0.52,0.48) ; VV:(0.65,0.35) 03:(0.09,0.86,0.05)	(5,8,13,17,18)	0.8808	0.1192	0.0175
	1000	HD:(0.18,0.63,0.19) HR:(0.86,0.14) ; TMP: (0.28,0.72) DV:(0.59,0.41) ; VV:(0.49,0.51) 03:(0.09,0.11,0.80)	(1,12,18)	0.8667	0.1333	0.0427
B	8000	HD:(0.51,0.23,0.26) HR:(0.18,0.82) ; TMP: (0.49,0.51) DV:(0.46,0.54) ; VV:(0.64,0.36) 03:(0.07,0.88,0.05)	(5,7,14,17,18)	0.8911	0.1089	0.0176
	8000	HD:(0.27,0.48,0.25) HR:(0.52,0.48) ; TMP: (0.56,0.44) DV:(0.46,0.54) ; VV:(0.65,0.35) 03:(0.08,0.87,0.05)	(2,3,13,17,18)	0.8801	0.1199	0.0176
	4000	HD:(0.38,0.34,0.28) HR:(0.44,0.56) ; TMP: (0.49,0.51) DV:(0.69,0.31) ; VV:(0.45,0.55) 03:(0.64,0.06,0.30)	(7,12,16,18)	0.8395	0.1605	0.0179
	4000	HD:(0.39,0.3,0.31) HR:(0.45,0.55) ; TMP: (0.58,0.42) DV:(0.37,0.63) ; VV:(0.53,0.47) 03:(0.05,0.89,0.06)	(1,5,14,17,18)	0.8617	0.1383	0.0184
	2000	HD:(0.57,0.17,0.26) HR:(0.47,0.53) ; TMP: (0.88,0.12) DV:(0.52,0.48) ; VV:(0.65,0.35) 03:(0.09,0.86,0.05)	(5,8,13,17,18)	0.8808	0.1192	0.0175
	2000	HD:(0.51,0.20,0.29) HR:(0.41,0.59) ; TMP: (0.26,0.74) DV:(0.50,0.50) ; VV:(0.55,0.45) 03:(0.09,0.81,0.10)	(5,13,14,17,18)	0.8404	0.1596	0.0180
	1000	HD:(0.21,0.48,0.31) HR:(0.38,0.62) ; TMP: (0.24,0.76) DV:(0.64,0.36) ; VV:(0.61,0.39) 03:(0.06,0.18,0.76)	(12,13,18)	0.8573	0.1427	0.0160
	1000	HD:(0.34,0.23,0.43) HR:(0.56,0.44) ; TMP: (0.50,0.50) DV:(0.47,0.53) ; VV:(0.45,0.55) 03:(0.64,0.06,0.30)	(1,12,18)	0.8079	0.1921	0.0177
	C	Método EFP en todas las variables, excepto en HD y 03. Máscara subóptima	(1,6,10,11,18)	0.5850	0.4150	0.0176

**Tabla 63.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable proporcionado por expertos en [NM06]). **Mes Agosto.**

		Núm. Clases: (3,2,2,2,2,3)				
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>pme</sub>
A	8000	HD:(0.32,0.50,0.18) HR:(0.25,0.75) ; TMP: (0.33,0.67) DV:(0.20,0.80) ; VV:(0.39,0.61) O3:(0.05,0.06,0.89)	(12,14,17,18)	0.9261	0.0739	0.0275
	4000	HD:(0.27,0.37,0.36) HR:(0.36,0.64) ; TMP: (0.51,0.49) DV:(0.54,0.46) ; VV:(0.43,0.57) O3:(0.05,0.05,0.90)	(12,14,17,18)	0.9243	0.0757	2.7981
	2000	HD:(0.52,0.26,0.22) HR:(0.42,0.58) ; TMP: (0.58,0.42) DV:(0.39,0.61) ; VV:(0.45,0.55) O3:(0.05,0.89,0.06)	(12,14,18)	0.8474	0.1526	0.0258
	1000	HD:(0.32,0.35,0.33) HR:(0.76,0.24) ; TMP: (0.57,0.43) DV:(0.56,0.44) ; VV:(0.45,0.55) O3:(0.05,0.05,0.90)	(9,12,17,18)	0.9201	0.0799	3.9616
B	8000	HD:(0.32,0.50,0.18) HR:(0.25,0.75) ; TMP: (0.33,0.67) DV:(0.20,0.80) ; VV:(0.39,0.61) O3:(0.05,0.06,0.89)	(12,14,17,18)	0.9261	0.0739	0.0275
	8000	HD:(0.21,0.43,0.36) HR:(0.45,0.55) ; TMP: (0.51,0.49) DV:(0.67,0.33) ; VV:(0.08,0.92) O3:(0.05,0.06,0.89)	(12,17,18)	0.9257	0.0743	0.0405
	4000	HD:(0.55,0.15,0.30) HR:(0.67,0.33) ; TMP: (0.71,0.29) DV:(0.54,0.46) ; VV:(0.57,0.43) O3:(0.82,0.07,0.11)	(1,12,15,18)	0.7857	0.2143	4.5962
	4000	HD:(0.38,0.26,0.36) HR:(0.50,0.50) ; TMP: (0.52,0.48) DV:(0.57,0.43) ; VV:(0.51,0.49) O3:(0.10,0.65,0.25)	(7,12,16,18)	0.7976	0.2024	9.5603
	2000	HD:(0.34,0.25,0.41) HR:(0.44,0.56) ; TMP: (0.46,0.54) DV:(0.46,0.54) ; VV:(0.62,0.38) O3:(0.10,0.66,0.24)	(1,12,18)	0.7946	0.2054	0.0187
	2000	HD:(0.36,0.26,0.38) HR:(0.53,0.47) ; TMP: (0.47,0.53) DV:(0.54,0.46) ; VV:(0.31,0.69) O3:(0.05,0.69,0.26)	(1,12,18)	0.8168	0.1832	0.0194
	1000	HD:(0.41,0.31,0.28) HR:(0.29,0.71) ; TMP: (0.51,0.49) DV:(0.53,0.47) ; VV:(0.47,0.53) O3:(0.05,0.61,0.34)	(12,13,18)	0.7623	0.2377	0.0185
	1000	HD:(0.30,0.35,0.35) HR:(0.55,0.45) ; TMP: (0.49,0.51) DV:(0.59,0.41) ; VV:(0.49,0.51) O3:(0.10,0.65,0.25)	(1,12,18)	0.7707	0.2293	0.0219
C		Método EFP en todas las variables, excepto en HD y O3. Máscara subóptima	(1,5,8,12,18)	0.1960	0.8040	0.0180

**Tabla 65.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo  $MSE_{entr}$  (Número de clases por variable proporcionado por expertos en [NM06]). **Mes Enero.**

Núm. Clases: (3,2,2,2,2,3)						
# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$	
A	8000	HD:(0.38,0.33,0.29) HR:(0.59,0.41); TMP: (0.50,0.50) DV:(0.57,0.43); VV:(0.53,0.47) O3:(0.57,0.29,0.14)	(4,12,13,16,18)	0.6370	29.3040	0.0180
	4000	HD:(0.38,0.34,0.28) HR:(0.48,0.52); TMP: (0.33,0.67) DV:(0.57,0.43); VV:(0.49,0.51) O3:(0.46,0.40,0.14)	(4,12,13,16,18)	0.6048	32.6062	0.0184
	2000	HD:(0.40,0.36,0.24) HR:(0.64,0.36); TMP: (0.63,0.37) DV:(0.57,0.43); VV:(0.48,0.52) O3:(0.46,0.36,0.18)	(4,12,13,16,18)	0.6090	32.7024	0.0185
	1000	HD:(0.16,0.28,0.56) HR:(0.85,0.15); TMP: (0.43,0.57) DV:(0.46,0.54); VV:(0.89,0.11) O3:(0.09,0.41,0.50)	(3,12,15,18)	0.5636	29.6646	0.0223
B	8000	HD:(0.40,0.30,0.30) HR:(0.59,0.41); TMP: (0.55,0.45) DV:(0.53,0.47); VV:(0.61,0.39) O3:(0.33,0.21,0.46)	(8,12,13,17,18)	0.6364	34.8826	0.0167
	8000	HD:(0.41,0.34,0.25) HR:(0.60,0.40); TMP: (0.11,0.89) DV:(0.67,0.33); VV:(0.59,0.41) O3:(0.56,0.19,0.25)	(4,12,13,16,18)	0.6684	33.4619	0.0168
	4000	HD:(0.38,0.40,0.22) HR:(0.57,0.43); TMP: (0.44,0.56) DV:(0.72,0.28); VV:(0.60,0.40) O3:(0.30,0.27,0.43)	(8,12,13,17,18)	0.6638	35.1395	0.0168
	4000	HD:(0.36,0.32,0.32) HR:(0.65,0.35); TMP: (0.76,0.24) DV:(0.58,0.42); VV:(0.89,0.11) O3:(0.57,0.29,0.14)	(4,7,12,16,18)	0.6397	33.4943	0.0170
	2000	HD:(0.41,0.34,0.25) HR:(0.42,0.58); TMP: (0.51,0.49) DV:(0.57,0.43); VV:(0.47,0.53) O3:(0.39,0.28,0.33)	(4,12,13,16,18)	0.7057	37.5814	0.0174
	2000	HD:(0.35,0.40,0.25) HR:(0.49,0.51); TMP: (0.58,0.42) DV:(0.55,0.45); VV:(0.50,0.50) O3:(0.41,0.23,0.36)	(7,8,12,17,18)	0.7057	37.3764	0.0176
	1000	HD:(0.36,0.43,0.21) HR:(0.58,0.42); TMP: (0.43,0.57) DV:(0.46,0.54); VV:(0.60,0.40) O3:(0.31,0.22,0.47)	(7,8,12,17,18)	0.6548	36.1851	0.0164
	1000	HD:(0.40,0.32,0.28) HR:(0.64,0.36); TMP: (0.77,0.23) DV:(0.46,0.54); VV:(0.61,0.39) O3:(0.29,0.24,0.47)	(8,12,13,17,18)	0.6443	37.4077	0.0167
C	Método EFP en todas las variables, excepto en HD y O3. Máscara subóptima	(1,6,10,11,18)	0.5850	---	0.0176	

**Tabla 66.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo  $MSE_{entr}$  (Número de clases por variable proporcionado por expertos en [NM06]). **Mes Agosto.**

Núm. Clases: (3,2,2,2,2,3)						
	# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A	8000	HD:(0.38,0.32,0.30) HR:(0.57,0.43) ; TMP: (0.51,0.49) DV:(0.57,0.43) ; VV:(0.35,0.65) O3:(0.13,0.17,0.70)	(12,14,15,16,18)	0.6671	18.9244	1.3820
	4000	HD:(0.18,0.44,0.38) HR:(0.64,0.36) ; TMP: (0.38,0.62) DV:(0.71,0.29) ; VV:(0.53,0.47) O3:(0.15,0.15, 0.70)	(12,14,15,16,18)	0.6673	19.8691	0.0256
	2000	HD:(0.30,0.21,0.49) HR:(0.43,0.57) ; TMP: (0.30,0.70) DV:(0.48,0.52) ; VV:(0.54,0.46) O3:(0.20,0.56,0.24)	(11,12,14,17,18)	0.6587	19.1904	1.9417
	1000	HD:(0.24,0.57,0.19) HR:(0.43,0.57) ; TMP: (0.53,0.47) DV:(0.23,0.77) ; VV:(0.53,0.47) O3:(0.18,0.60,0.22)	(11,12,14,17,18)	0.6658	19.1217	2.0859
B	8000	HD:(0.20,0.46,0.34) HR:(0.54,0.46) ; TMP: (0.35,0.65) DV:(0.44,0.56) ; VV:(0.58,0.42) O3:(0.18,0.34,0.48)	(7,8,12,15,18)	0.5934	21.5514	0.0176
	8000	HD:(0.37,0.34,0.29) HR:(0.50,0.50) ; TMP: (0.41,0.59) DV:(0.81,0.19) ; VV:(0.50,0.50) O3:(0.27,0.39,0.34)	(5,12,13,14,18)	0.6083	21.3091	0.0181
	4000	HD:(0.37,0.40,0.23) HR:(0.49,0.51) ; TMP: (0.52,0.48) DV:(0.51,0.49) ; VV:(0.52,0.48) O3:(0.23,0.44,0.33)	(5,12,13,14,18)	0.6124	21.1435	0.0171
	4000	HD:(0.25,0.27,0.48) HR:(0.49,0.51) ; TMP: (0.47,0.53) DV:(0.74,0.26) ; VV:(0.52,0.48) O3:(0.31,0.34,0.35)	(5,12,13,14,18)	0.5441	20.4738	0.0185
	2000	HD:(0.36,0.37,0.27) HR:(0.50,0.50) ; TMP: (0.57,0.43) DV:(0.60,0.40) ; VV:(0.52,0.48) O3:(0.32,0.33,0.35)	(5,12,13,14,18)	0.5772	22.1061	0.0164
	2000	HD:(0.37,0.39,0.24) HR:(0.51,0.49) ; TMP: (0.42,0.58) DV:(0.36,0.64) ; VV:(0.52,0.48) O3:(0.26,0.41,0.33)	(5,12,13,14,18)	0.6120	20.7821	0.0165
	1000	HD:(0.24,0.33,0.43) HR:(0.51,0.49) ; TMP: (0.36,0.64) DV:(0.81,0.19) ; VV:(0.48,0.52) O3:(0.23,0.22,0.55)	(7,8,12,15,18)	0.5645	22.4100	0.0171
	1000	HD:(0.35,0.46,0.19) HR:(0.58,0.42) ; TMP: (0.52,0.48) DV:(0.40,0.60) ; VV:(0.53,0.47) O3:(0.20,0.47,0.33)	(5,12,13,14,18)	0.6415	21.4275	0.0178
	C	Método EFP en todas las variables, excepto en HD y O3. Máscara subóptima	(1,5,8,12,18)	0.1960	---	0.0180

### 7.3.2 Aprendizaje del número óptimo de clases para cada variable del sistema (AG1)

En esta sección, el **AG1** se ejecuta para obtener el número óptimo de clases tanto para las variables de entrada como para la variable de salida. Aquí se presentan los resultados obtenidos para el problema de predicción de concentraciones de ozono en México. El método EFP se emplea para encontrar las funciones de pertenencia de las clases.

**Tabla 67.** Resultados de granularidad para la estimación de la concentración de ozono en México usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 25% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. **Mes Enero.** (X significa cualquier valor entre [ 2..9]).

1-Q	# eval.	Gran.	Opt. Mask	Q	1-Q	$RMS_{prue}$
	400, 800	(5,X,X,X,X,2)	(12,13,18)	0.6999	0.3001	0.0406
	400, 800	(3,X,X,3,3,2)	(4,7,12,17,18)	0.6826	0.3174	0.0179
	800	(3,4,5,3,8,2)	(7,12,17,18)	0.6894	0.3106	0.0185
	400	(3,X,X,3,X,3)	(4,7,12,18)	0.6816	0.3184	0.0194
	400	(6,3,8,3,3,2)	(7,12,18)	0.6855	0.3145	0.0360
	400, 800	(8,X,X,X,X,2)	(12,13,18)	0.6970	0.3030	0.0416
	100,200	(5,X,X,X,X,2)	(12,13,18)	0.6999	0.3001	0.0406
	100,200	(3,X,X,3,3,2)	(4,7,12,17,18)	0.6826	0.3174	0.0179
	200	(3,3,7,3,6,2)	(7,12,17,18)	0.6864	0.3136	0.0188
	100, 200	(3,X,X,3,X,3)	(4,7,12,18)	0.6816	0.3184	0.0194
	100, 200	(8,X,X,X,X,2)	(12,13,18)	0.6970	0.3030	0.0416
	100	(9,9,9,6,3,2)	(1,12,18)	0.6886	0.3114	0.0416
	Solución óptima sugerida: Partición = (5,X,X,X,X,2); Q=0,6999					
$MSE_{entr}$	# eval.	Gran.	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
	400, 800	(2,2,X,X,4,2)	(1,12,14,17,18)	0.5812	47.94	0.0229
	800	(3,3,6,9,3,2)	(7,12,16,18)	0.6695	56.73	0.0184
	400, 800	(2,X,X,X,4,4)	(1,12,17,18)	0.5587	57.91	0.0193
	400, 800	(2,2,X,X,2,2)	(1,12,14,17,18)	0.5687	49.12	0.0221
	400, 800	(2,2,X,X,3,3)	(1,12,14,17,18)	0.6275	52.85	0.0244
	200	(2,2,5,7,4,2)	(1,12,14,17,18)	0.5812	47.94	0.0229
	100	(3,8,8,8,2,2)	(7,12,16,18)	0.6683	60.89	0.0180
	100	(3,4,3,9,2,2)	(7,12,16,18)	0.6695	56.73	0.0184
	100, 200	(2,X,X,X,4,4)	(1,12,17,18)	0.5587	57.91	0.0193
	200	(3,4,2,4,6,4)	(4,7,12,18)	0.5729	58.49	0.0215
	100, 200	(2,2,X,X,2,2)	(1,12,14,17,18)	0.5687	49.12	0.0221
	100	(2,2,3,2,3,2)	(1,11,12,14,18)	0.5794	57.87	0.0235
	100	(2,2,X,X,3,3)	(1,12,14,17,18)	0.6275	52.85	0.0244
Solución óptima sugerida: Partición = (2,2,X,X,4,2); $MSE_{entr}$ =47.94						

**Tabla 68.** Resultados de granularidad para la estimación de la concentración de ozono en México usando las funciones objetivo: a) 1-Q y, b) Error de predicción MSE del último 25% de datos del conjunto de entrenamiento ( $MSE_{entr}$ ). Método EFP. **Mes Agosto.** (X significa cualquier valor entre [ 2..9]).

1-Q	# eval.	Gran.	Opt. Mask	Q	1-Q	$RMS_{prue}$
	800	(3,4,X,X,X,2)	(1,12,14,18)	0.5846	0.4154	0.5984
	400	(3,4,X,X,X,2)	(1,12,14,18)	0.5846	0.4154	0.5984
	400	(3,3,2,6,4,3)	(1,12,14,18)	0.5795	0.4205	0.6336
	200	(3,4,X,X,X,2)	(1,12,14,18)	0.5846	0.4154	0.5984
	200	(3,3,X,X,X,3)	(1,12,14,18)	0.5795	0.4205	0.6336
	100	(3,4,X,X,X,2)	(1,12,14,18)	0.5846	0.4154	0.5984
	100	(3,3,X,X,X,3)	(1,12,14,18)	0.5795	0.4205	0.6336
	Solución óptima sugerida: Partición = (3,4,X,X,X,2); Q=0.5846					
$MSE_{entr}$	# eval.	Gran.	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
	800	(X,3,2,X,2,3)	(5,12,14,15,18)	0.5526	23.44	0.3499
	400	(X,3,2,X,2,3)	(5,12,14,15,18)	0.5526	23.44	0.3499
	400	(2,2,2,2,3,3)	(1,8,12,15,18)	0.5481	25.34	0.0191
	400	(4,2,4,2,3,2)	(1,10,12,14,18)	0.5379	26.72	3.7964
	200	(2,2,2,X,X,3)	(1,8,12,15,18)	0.5481	25.34	0.0191
	200	(3,3,5,2,8,7)	(7,12,17,18)	0.4828	26.24	2.1143
	100	(2,2,2,3,3,3)	(1,8,12,15,18)	0.5481	25.34	0.0191
	100	(3,4,4,6,8,7)	(7,12,17,18)	0.4828	26.24	2.1143
	100	(3,3,2,2,7,7)	(7,12,17,18)	0.4820	26.33	3.0671
	Solución óptima sugerida: Partición = (X,3,2,X,2,3); $MSE_{entr}=23.44$					

Las *Tablas 67 y 68* presentan los resultados obtenidos por el **AG1** para la estimación de concentraciones locales máximas de ozono en México usando las funciones objetivo: a) 1-Q, y b) Error de predicción MSE del último 25% del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). Como se ha indicado en otros casos, el resto de la señal de entrenamiento (el primer 75% de los datos), en la función objetivo  $MSE_{entr}$  se utiliza para obtener los modelos FIR (máscaras y base de reglas patrón). Cada tabla se refiere a uno de los meses considerados para el estudio de modelado. Las tablas han sido distribuidas de la siguiente manera:

- *# Eval.* : indica el número de evaluaciones de cromosomas hechas por el AG.
- *Gran.*: señala la granularidad sugerida por el **AG1** cuando finaliza su ejecución. Nótese que aquí la granularidad sugerida es el parámetro de entrada al proceso de fusificación de la metodología FIR.
- *Opt. Mask* y *Q*: contienen la máscara óptima (en notación de posición) identificada por FIR para esta granularidad específica y la calidad asociada a esta máscara, respectivamente.
- *1-Q*: corresponde al valor de la función objetivo de la calidad.
- $MSE_{entr}$ : corresponde al valor de la función objetivo del error de predicción

MSE del último 25% de los datos del conjunto de entrenamiento.

- $RMS_{prue}$ : indica al error de predicción del conjunto de datos de prueba (Véase ecuación 51).

En la *Tabla 69* se muestra el tiempo de CPU que necesita el **AG1** para realizar las treinta ejecuciones para cada uno de los criterios de parada estipulados y mes considerado cuando se utilizan las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente.

**Tabla 69.** Tiempo consumido por el AG1 para realizar las treinta ejecuciones por cada criterio de parada. Caso de Estudio: Ozono México-Meses (**Enero y Agosto**). Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Función objetivo	Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas) por mes considerado	
		Enero	Agosto
1-Q	100	11:22	14:43
	200	15:43	23:56
	400	23:17	36:29
	800	44:28	61:55
$MSE_{entr}$	100	16:59	16:02
	200	26:13	24:16
	400	42:16	38:10
	800	86:15	76:04

### 7.3.3 Aprendizaje de las funciones de pertenencia de las clases (AG2). Granularidad suministrada por el AG1

A continuación se exponen los resultados obtenidos del **AG2** utilizando el número óptimo de clases obtenido por el **AG1**. De este modo, se logra la combinación **AG1 + AG2**.

Las *Tablas 70 y 71* muestran los resultados obtenidos por el **AG2** para cada mes usando la función objetivo 1-Q tomando en cuenta las granularidades suministradas por el **AG1**. Las tablas están distribuidas como la *Tabla 62*. La diferencia radica en el contenido de las secciones de la primera columna. La sección *A* corresponde al mejor resultado obtenido (valor 1-Q) por la combinación **AG1 + AG2**. En cambio, la sección *B* pertenece a los mejores resultados de predicción de los datos de prueba obtenidos por la combinación **AG1 + AG2**. Por su parte, las *Tablas 72 y 73* presentan los resultados obtenidos por el **AG2** para la función objetivo  $MSE_{entr}$  y están organizadas como la *Tabla 65* pero se contemplan las dos secciones del mismo modo que la *Tabla 70*.

**Tabla 70.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). **Mes Enero.**

Núm. Clases: (3,2,2,3,3,2)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	8000	HD:(0.26,0.28,0.46) HR:(0.93,0.07) ; TMP: (0.49,0.51) DV:(0.20,0.74,0.06) ; VV:(0.26,0.13,0.61) O3:(0.09,0.91)	(12,14,18)	0.9818	0.0182	0.0328
	4000	HD:(0.25,0.29,0.46) HR:(0.76,0.24) ; TMP: (0.17,0.83) DV:(0.05,0.38,0.57) ; VV:(0.14,0.25,0.61) O3:(0.09,0.91)	(5,14,15,17,18)	0.9821	0.0179	0.0198
	2000	HD:(0.44,0.33,0.23) HR:(0.68,0.32) ; TMP: (0.52,0.48) DV:(0.33,0.38,0.29) ; VV:(0.14,0.29,0.57) O3:(0.09,0.91)	(5,10,14,17,18)	0.9814	0.0186	0.0288
	1000	HD:(0.35,0.44,0.21) HR:(0.60,0.40) ; TMP: (0.43,0.57) DV:(0.35,0.34,0.31) ; VV:(0.28,0.30,0.42) O3:(0.07,0.93)	(5,14,16,17,18)	0.9801	0.0199	0.0285
B	8000	HD:(0.38,0.34,0.28) HR:(0.45,0.55) ; TMP: (0.56,0.44) DV:(0.34,0.26,0.40) ; VV:(0.62,0.22,0.16) O3:(0.69,0.31)	(7,12,17,18)	0.8689	0.1311	0.0183
	8000	HD:(0.39,0.33,0.28) HR:(0.50,0.50) ; TMP: (0.47,0.53) DV:(0.41,0.29,0.30) ; VV:(0.58,0.25,0.17) O3:(0.70,0.30)	(7,12,17,18)	0.8648	0.1352	0.0195
	4000	HD:(0.38,0.34,0.28) HR:(0.40,0.60) ; TMP: (0.47,0.53) DV:(0.46,0.26,0.28) ; VV:(0.24,0.31,0.45) O3:(0.69,0.31)	(7,12,16,18)	0.8575	0.1425	0.0181
	4000	HD:(0.27,0.39,0.34) HR:(0.73,0.27) ; TMP: (0.53,0.47) DV:(0.27,0.46,0.27) ; VV:(0.65,0.17,0.18) O3:(0.09,0.91)	(5,7,14,18)	0.9741	0.0259	0.0187
	2000	HD:(0.39,0.33,0.28) HR:(0.52,0.48) ; TMP: (0.55,0.45) DV:(0.41,0.34,0.25) ; VV:(0.66,0.16,0.18) O3:(0.69,0.31)	(7,12,17,18)	0.8677	0.1323	0.0189
	2000	HD:(0.39,0.32,0.29) HR:(0.61,0.39) ; TMP: (0.50,0.50) DV:(0.34,0.39,0.27) ; VV:(0.57,0.26,0.17) O3:(0.69,0.31)	(7,12,17,18)	0.8678	0.1322	0.0190
	1000	HD:(0.34,0.36,0.30) HR:(0.48,0.52) ; TMP: (0.52,0.48) DV:(0.39,0.31,0.30) ; VV:(0.36,0.35,0.29) O3:(0.70,0.30)	(1,12,18)	0.8501	0.1499	0.0176
	1000	HD:(0.37,0.32,0.31) HR:(0.45,0.55) ; TMP: (0.54,0.46) DV:(0.26,0.31,0.43) ; VV:(0.30,0.43,0.27) O3:(0.70,0.30)	(1,12,18)	0.8501	0.1499	0.0176

**Tabla 71.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo 1-Q (Número de clases por variable suministrado por el AG1). **Mes Agosto.**

Núm. Clases: (3,4,2,2,2,3)						
	# eval	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A	8000	HD:(0.26,0.37,0.37) HR:(0.26,0.11,0.57,0.06) ; TMP: (0.70,0.30) DV:(0.56,0.44) ; VV:(0.16,0.84) O3:(0.05,0.06,0.89)	(5,12,17,18)	0.9240	0.0760	9.5272
	4000	HD:(0.26,0.41,0.33) HR:(0.17,0.41,0.29,0.13) ; TMP: (0.36,0.64) DV:(0.63,0.37) ; VV:(0.09,0.91) O3:(0.05,0.06,0.89)	(12,17,18)	0.9241	0.0759	2.8256
	2000	HD:(0.70,0.24,0.06) HR:(0.15,0.12,0.16,0.57) ; TMP: (0.34,0.66) DV:(0.39,0.61) ; VV:(0.80,0.20) O3:(0.05,0.05,0.90)	(3,12,15,18)	0.9203	0.0797	8.8160
	1000	HD:(0.42,0.36,0.22) HR:(0.20,0.13,0.32,0.35) ; TMP: (0.46,0.54) DV:(0.46,0.54) ; VV:(0.40,0.60) O3:(0.05,0.90,0.05)	(5,12,17,18)	0.8518	0.1482	5.4974
B	8000	HD:(0.30,0.18,0.52) HR:(0.10,0.22,0.11,0.57) ; TMP: (0.47,0.53) DV:(0.51,0.49) ; VV:(0.36,0.64) O3:(0.05,0.89,0.06)	(12,14,18)	0.8587	0.1413	0.7417
	8000	HD:(0.37,0.25,0.38) HR:(0.22,0.41,0.10,0.27) ; TMP: (0.69,0.31) DV:(0.31,0.69) ; VV:(0.35,0.65) O3:(0.05,0.89,0.06)	(5,12,15,18)	0.8583	0.1417	1.4687
	4000	HD:(0.68,0.14,0.18) HR:(0.33,0.26,0.13,0.28) ; TMP: (0.70,0.30) DV:(0.42,0.58) ; VV:(0.37,0.63) O3:( 0.05,0.88,0.07)	(5,12,15,18)	0.8519	0.1481	0.0246
	4000	HD:(0.30,0.38,0.32) HR:(0.06,0.25,0.50,0.19) ; TMP: (0.54,0.46) DV:(0.63,0.37) ; VV:(0.56,0.44) O3:(0.05,0.10,0.85)	(12,14,18)	0.8294	0.1706	0.0287
	2000	HD:(0.27,0.34,0.39) HR:(0.27,0.23,0.23,0.27) ; TMP: (0.55,0.45) DV:(0.35,0.65) ; VV:(0.39,0.61) O3:(0.05,0.71,0.24)	(1,12,18)	0.7985	0.2015	0.0204
	2000	HD:(0.24,0.58,0.18) HR:(0.15,0.36,0.28,0.21) ; TMP: (0.79,0.21) DV:(0.51,0.49) ; VV:(0.47,0.53) O3:(0.07,0.08,0.85)	(1,12,16,18)	0.8216	0.1784	0.1057
	1000	HD:(0.34,0.28,0.38) HR:(0.19,0.34,0.23,0.24) ; TMP: (0.52,0.48) DV:(0.47,0.53) ; VV:(0.53,0.47) O3:(0.10,0.66,0.24)	(1,12,18)	0.7946	0.2054	0.0187
	1000	HD:(0.34,0.32,0.34) HR:(0.27,0.19,0.16,0.38) ; TMP: (0.60,0.40) DV:(0.48,0.52) ; VV:(0.59,0.41) O3:(0.09,0.66,0.25)	(1,12,18)	0.7843	0.2157	0.0196

**Tabla 72.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1). **Mes Enero.**

Núm. Clases: (3,2,2,8,2,2)					
# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A 8000	HD:(0.19,0.26,0.55) HR:(0.28,0.72) ; TMP: (0.41,0.59) DV:(0.13,0.13,0.17,0.12,0.12,0.12,0.11,0.10) VV:(0.55,0.45) ; O3:(0.52,0.48)	(3,11,12,15,18)	0.5836	29.6107	0.0195
4000	HD:(0.24,0.68,0.08) HR:(0.59,0.41) ; TMP: (0.42,0.58) DV:(0.09,0.15,0.09,0.13,0.09,0.13,0.27,0.05) VV:(0.56,0.44) ; O3:(0.52,0.48)	(3,11,12,15,18)	0.5873	31.2674	0.0196
2000	HD:(0.24,0.68,0.08) HR:(0.57,0.43) ; TMP: (0.53,0.47) DV:(0.16,0.12,0.15,0.13,0.13,0.13,0.09,0.09) VV:(0.64,0.36) ; O3:(0.53,0.47)	(3,11,12,15,18)	0.6051	32.5828	0.0179
1000	HD:(0.29,0.58,0.13) HR:(0.35,0.65) ; TMP: (0.42,0.58) DV:(0.09,0.09,0.24,0.07,0.10,0.08,0.10,0.23) VV:(0.67,0.33) ; O3:(0.41,0.59)	(11,12,13,15,18)	0.6612	36.4081	0.0208
B 8000	HD:(0.23,0.53,0.24) HR:(0.49,0.51) ; TMP: (0.54,0.46) DV:(0.11,0.14,0.13,0.15,0.14,0.13,0.11,0.09) VV:(0.46,0.54) ; O3:(0.79,0.21)	(2,13,15,17,18)	0.7268	37.7292	0.0169
8000	HD:(0.39,0.27,0.34) HR:(0.42,0.58) ; TMP: (0.48,0.52) DV:(0.16,0.14,0.07,0.12,0.21,0.12,0.11,0.07) VV:(0.60,0.40) ; O3:(0.49,0.51)	(12,13,14,17,18)	0.6489	40.0706	0.0186
4000	HD:(0.38,0.33,0.29) HR:(0.52,0.48) ; TMP: (0.34,0.66) DV:(0.13,0.10,0.11,0.15,0.10,0.11,0.14,0.16) VV:(0.59,0.41) ; O3:(0.64,0.36)	(7,12,14,17,18)	0.8437	40.8856	0.0185
4000	HD:(0.32,0.48,0.20) HR:(0.51,0.49) ; TMP: (0.44,0.56) DV:(0.09,0.26,0.18,0.15,0.12,0.07,0.07,0.06) VV:(0.67,0.33) ; O3:(0.58,0.42)	(12,13,14,17,18)	0.7129	39.2474	0.0191
2000	HD:(0.41,0.32,0.27) HR:(0.45,0.55) ; TMP: (0.64,0.36) DV:(0.12,0.12,0.12,0.10,0.19,0.14,0.12,0.09) VV:(0.57,0.43) ; O3:(0.41,0.59)	(12,13,14,17,18)	0.7246	42.044	0.0184
2000	HD:(0.38,0.36,0.26) HR:(0.47,0.53) ; TMP: (0.41,0.59) DV:(0.13,0.13,0.13,0.12,0.12,0.14,0.11,0.12) VV:(0.56,0.44) ; O3:(0.41,0.59)	(12,13,14,17,18)	0.7240	42.7117	0.0187
1000	HD:(0.45,0.31,0.24) HR:(0.52,0.48) ; TMP: (0.40,0.60) DV:(0.13,0.11,0.11,0.16,0.14,0.08,0.09,0.18) VV:(0.58,0.42) ; O3:(0.64,0.36)	(12,13,14,17,18)	0.8429	39.9073	0.0183
1000	HD:(0.39,0.27,0.34) HR:(0.40,0.60) ; TMP: (0.35,0.65) DV:(0.08,0.11,0.18,0.09,0.18,0.13,0.15,0.08) VV:(0.59,0.41) ; O3:(0.41,0.59)	(12,13,14,17,18)	0.7099	41.9458	0.0185

**Tabla 73.** Resultados de las funciones de pertenencia para la estimación de concentraciones de ozono en México usando la función objetivo  $MSE_{entr}$  (Número de clases por variable suministrado por el AG1). **Mes Agosto.**

Núm. Clases: (2,2,2,2,2,3)						
# eval	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$	
A 8000	HD:(0.55,0.45) HR:(0.50,0.50) ; TMP: (0.44,0.56) DV:(0.61,0.39) ; VV:(0.57,0.43) O3:(0.30,0.42,0.28)	(1,8,12,15,18)	0.5822	18.1559	1.6349	
4000	HD:(0.50,0.50) HR:(0.43,0.57) ; TMP: (0.34,0.66) DV:(0.57,0.43) ; VV:(0.53,0.47) O3:(0.20,0.58,0.22)	(11,12,14,17,18)	0.6658	19.1217	2.0859	
2000	HD:(0.56,0.44) HR:(0.49,0.51) ; TMP: (0.52,0.48) DV:(0.75,0.25) ; VV:(0.49,0.51) O3:(0.29,0.43,0.28)	(1,8,12,15,18)	0.5902	19.5133	1.8932	
1000	HD:(0.55,0.45) HR:(0.49,0.51) ; TMP: (0.52,0.48) DV:(0.78,0.22) ; VV:(0.16,0.84) O3:(0.30,0.42,0.28)	(1,8,12,15,18)	0.5902	19.5133	1.8932	
B 8000	HD:(0.48,0.52) HR:(0.58,0.42) ; TMP: (0.55,0.45) DV:(0.71,0.29) ; VV:(0.51,0.49) O3:(0.31,0.36,0.33)	(1,5,12,15,18)	0.5421	20.6821	0.0193	
8000	HD:(0.55,0.45) HR:(0.49,0.51) ; TMP: (0.51,0.49) DV:(0.35,0.65) ; VV:(0.63,0.37) O3:(0.29,0.37,0.34)	(1,8,12,15,18)	0.5612	21.0545	0.0209	
4000	HD:(0.43,0.57) HR:(0.59,0.41) ; TMP: (0.60,0.40) DV:(0.30,0.70) ; VV:(0.68,0.32) O3:(0.27,0.42,0.31)	(1,5,12,14,18)	0.5865	20.7490	0.0182	
4000	HD:(0.46,0.54) HR:(0.60,0.40) ; TMP: (0.42,0.58) DV:(0.14,0.86) ; VV:(0.52,0.48) O3:(0.18,0.48,0.34)	(1,5,12,14,18)	0.6059	21.3079	0.0194	
2000	HD:(0.44,0.56) HR:(0.58,0.42) ; TMP: (0.62,0.38) DV:(0.41,0.59) ; VV:(0.70,0.30) O3:(0.26,0.37,0.37)	(1,5,12,14,18)	0.5568	20.8942	0.0189	
2000	HD:(0.55,0.45) HR:(0.49,0.51) ; TMP: (0.51,0.49) DV:(0.61,0.39) ; VV:(0.90,0.10) O3:(0.29,0.37,0.34)	(1,8,12,15,18)	0.5612	21.0545	0.0209	
1000	HD:(0.51,0.49) HR:(0.51,0.49) ; TMP: (0.59,0.41) DV:(0.05,0.95) ; VV:(0.58,0.42) O3:(0.31,0.23,0.46)	(7,8,12,15,18)	0.5200	23.6769	0.0174	
1000	HD:(0.44,0.56) HR:(0.49,0.51) ; TMP: (0.51,0.49) DV:(0.69,0.31) ; VV:(0.53,0.47) O3:(0.31,0.34,0.35)	(1,8,12,15,18)	0.5421	23.231	0.0201	

La *Tabla 74* muestra el tiempo necesario de CPU del **AG1 + AG2** para alcanzar las treinta ejecuciones para cada uno de los criterios de parada estipulados durante el mes de Enero cuando se utilizan las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente. El tiempo consumido para el mes de Agosto es de similar magnitud.

**Tabla 74.** Tiempo consumido por el AG1 + AG2 para realizar las treinta ejecuciones por cada criterio de parada. Caso Estudio: Ozono México-**Mes Enero**. Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas) por función objetivo	
	1-Q	$MSE_{entr}$
1000	50:37	52:19
2000	100:43	102:24
4000	194:54	202:39
8000	390:31	404:52

#### 7.3.4 Aprendizaje global del número óptimo de clases y de las funciones de pertenencia de las clases (AG3)

En esta sección se visualizan los resultados obtenidos por el método propuesto en el Capítulo V sobre el problema de modelado de concentraciones de ozono en México. Las *Tablas 75 y 76* muestran los resultados obtenidos cuando se usa la función objetivo 1-Q para cada uno de los meses considerados. Las tablas están organizadas como se describe a continuación. La primera columna se divide en tres secciones. La sección *A* corresponde al mejor resultado obtenido por el **AG3** (valor mínimo de las funciones objetivo). La sección *B* corresponde a los mejores resultados de predicción obtenidos por el **AG3**. La sección *C* indica los resultados obtenidos cuando se emplean parámetros por defecto en FIR (**FIR DEFECTO**), es decir, todas las variables se discretizan en tres clases y se obtienen particiones difusas uniformes usando el método EFP para establecer las funciones de pertenencia. La sección *D* y *E* muestran los resultados cuando se emplean **AG1** y **AG1+AG2** que fueron implementados en apartados previos. De esta forma, *C*, *D* y *E* se usan como referencia para el análisis de los resultados. La segunda columna indica el criterio de parada, es decir, el número de evaluaciones de cromosomas realizadas por el AG. La tercera y cuarta columnas presentan la solución sugerida por el **AG3** (número de clases y distribución de landmarks) para las variables de entrada (HD, HR, TMP, DV, VV) y la variable de salida ( $O_3$ ) cuando se cumple el criterio de parada estipulado. El número de elementos de la proporción de datos corresponde al número de clases para esa variable. La quinta columna contienen la máscara óptima (en

**Tabla 75.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo 1-Q. **Mes Enero.**

# eval	Gran.	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>prue</sub>
A 8000	(8,9,3,2,3,2)	HD:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) HR:(0.11,0.11,0.14,0.11,0.11,0.14,0.09,0.09,0.10) TMP:(0.45,0.29,0.26) ; DV:(0.50,0.50) VV:(0.24,0.26,0.50) ; O3:(0.09,0.91)	(4,14,16,17,18)	0.9858	0.0142	0.0299
4000	(5,7,3,3,2,2)	HD:(0.20,0.20,0.20,0.20,0.20) HR:(0.19,0.10,0.33,0.10,0.10,0.11,0.07) TMP:(0.33,0.33,0.34) ; DV:(0.25,0.25,0.50) VV:(0.36,0.64) ; O3:(0.08,0.92)	(5,10,14,17,18)	0.9854	0.0146	0.0291
2000	(8,8,3,2,2,2)	HD:(0.17,0.05,0.19,0.12,0.15,0.06,0.11,0.15) HR:(0.11,0.24,0.13,0.20,0.07,0.09,0.09,0.07) TMP:(0.33,0.33,0.34) ; DV:(0.50,0.50) VV:(0.45,0.55) ; O3:(0.06,0.94)	(5,14,16,17,18)	0.9857	0.0143	0.0297
1000	(8,9,5,2,2,2)	HD:(0.06,0.06,0.23,0.08,0.19,0.14,0.14,0.10) HR:(0.10,0.16,0.10,0.09,0.17,0.09,0.09,0.09,0.11) TMP:(0.26,0.07,0.43,0.09,0.15) ; DV:(0.50,0.50) VV:(0.40,0.60) ; O3:(0.08,0.92)	(5,10,14,17,18)	0.9842	0.0158	0.0292
B 8000	(5,5,3,8,2,2)	HD:(0.38,0.08,0.10,0.17,0.27) HR:(0.20,0.20,0.20,0.20,0.20) ; TMP:(0.43,0.42,0.15) DV:(0.09,0.12,0.09,0.09,0.23,0.18,0.10,0.10) VV:(0.83,0.17) ; O3:(0.69,0.31)	(7,12,17,18)	0.8737	0.1263	0.0179
8000	(6,3,4,3,5,3)	HD:(0.11,0.14,0.15,0.32,0.13,0.15) ; HR:(0.51,0.26,0.23) TMP:(0.20,0.24,0.26,0.30) DV:(0.46,0.26,0.28) ; VV:(0.20,0.25,0.21,0.13,0.21) O3:(0.05,0.65,0.30)	(7,14,16,18)	0.8598	0.1402	0.0195
4000	(5,2,4,4,2,2)	HD:(0.35,0.13,0.22,0.12,0.18) HR:(0.50,0.50) ; TMP:(0.25,0.25,0.25,0.25) DV:(0.37,0.36,0.14,0.13) ; VV:(0.83,0.17) O3:(0.67,0.33)	(1,12,17,18)	0.8609	0.1391	0.0177
4000	(3,3,4,2,2,2)	HD:(0.34,0.32,0.34) HR:(0.38,0.24,0.38) ; TMP:(0.17,0.17,0.16,0.50) DV:(0.68,0.32) ; VV:(0.63,0.37) O3:(0.69,0.31)	(1,12,16,18)	0.8534	0.1466	0.0183
2000	(2,6,8,6,2,2)	HD:(0.64,0.36) ; HR:(0.18,0.05,0.06,0.15,0.51,0.05) TMP:(0.32,0.11,0.05,0.05,0.21,0.05,0.16,0.05) DV:(0.10,0.12,0.20,0.05,0.09,0.44) VV:(0.63,0.37) ; O3:(0.95,0.05)	(1,6,12,18)	0.8795	0.1205	0.0182
2000	(4,7,2,8,2,2)	HD:(0.34,0.24,0.11,0.31) HR:(0.14,0.14,0.14,0.14,0.15,0.15) ; TMP:(0.50,0.50) DV:(0.09,0.10,0.10,0.27,0.11,0.11,0.11,0.11) VV:(0.72,0.28) ; O3:(0.70,0.30)	(1,12,17,18)	0.8746	0.1254	0.0185
1000	(3,2,8,4,2,2)	HD:(0.39,0.31,0.30) ; HR:(0.24,0.76) TMP:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) DV:(0.46,0.19,0.3,0.05) VV:(0.72,0.28) ; O3:(0.7,0.30)	(7,12,17,18)	0.8534	0.1466	0.0178
1000	(4,3,3,5,4,3)	HD:(0.25,0.25,0.25,0.25) HR:(0.33,0.33,0.34) ; TMP:(0.33,0.33,0.34) DV:(0.20,0.20,0.20,0.20,0.20) VV:(0.25,0.25,0.25,0.25) ; O3:(0.06,0.80,0.14)	(8,13,17,18)	0.8278	0.1722	0.0178
C	(3,3,3,3,3,3)	Equal Frequency Partition method	(4,7,12,18)	0.6816	0.3184	0.0194
D	(3,2,2,3,3,2)	Equal Frequency Partition method	(4,7,12,17,18)	0.6826	0.3174	0.0179
E	(3,2,2,3,3,2)	HD:(0.34,0.36,0.30) HR:(0.48,0.52) ; TMP:(0.52,0.48) DV:(0.39,0.31,0.30) ; VV:(0.36,0.35,0.29) ; O3:(0.70,0.30)	(1,12,18)	0.8501	0.1499	0.0176

notación de posición) encontrada por FIR para esa partición (número de clases) y proporción de datos (distribución de landmarks). La sexta columna corresponde a la calidad asociada a la máscara óptima. La séptima columna es el valor de la función objetivo 1-Q. La última columna muestra el error de predicción RMS para el conjunto de datos de prueba ( $RMS_{prue}$ ).

Por otra parte, las *Tablas 77 y 78* presentan los resultados del **AG3** cuando se usa la función objetivo del error de predicción MSE del último 25% del conjunto de datos de entrenamiento ( $MSE_{entr}$ ). Estas tablas están organizadas como la *Tabla 75*. La diferencia está en que la séptima columna contiene los valores de la función objetivo  $MSE_{entr}$  en vez de la de 1-Q.

La *Tabla 79* muestra el tiempo de CPU que necesita el **AG3** para realizar las treinta ejecuciones por cada uno de los criterios de parada estipulados durante el mes de Enero, cuando se utilizan las funciones objetivo 1-Q y  $MSE_{entr}$ , respectivamente. Es necesario señalar que el tiempo de CPU para el mes de Agosto es del mismo orden de magnitud.

**Tabla 79.** Tiempo consumido por el AG3 para realizar las treinta ejecuciones para cada criterio de parada. Caso Estudio: Ozono México - **Mes Enero**. Funciones objetivo: 1-Q y  $MSE_{entr}$ .

Criterio de Parada (# evaluaciones de cromosomas)	Tiempo (horas) por función objetivo	
	1-Q	$MSE_{entr}$
1000	147:23	166:25
2000	268:05	285:50
4000	512:14	491:04
8000	919:16	945:55

**Tabla 76.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo 1-Q. **Mes Agosto.**

# eval	Gran.	Proporción de datos	Opt. Mask	Q	1-Q	RMS <sub>punc</sub>
A	8000 (9,7,2,4,2,2)	HD:(0.07,0.07, 0.2,0.09,0.09,0.09,0.09,0.21) HR:(0.11,0.11,0.11,0.11,0.11,0.33,0.12) TMP: (0.50,0.50)	(12,16,18)	0.9496	0.0504	22.0957
	4000 (2,2,9,2,4,2)	DV:(0.06,0.32,0.31,0.31) ; VV:(0.54,0.46) ; O3:(0.05,0.95) HD:(0.50,0.50) ; HR:(0.07,0.93) TMP: (0.09,0.10,0.10,0.20,0.10,0.10,0.10,0.10,0.11) DV:(0.36,0.64) ; VV:(0.54,0.15,0.15,0.16) O3:(0.05,0.95)	(2,7,12,14,18)	0.9569	0.0431	0.8812
	2000 (2,5,3,3,8,2)	HD:(0.50,0.50) ; HR:(0.05,0.19,0.12,0.41,0.23) TMP:(0.33,0.33,0.34) ; DV:(0.05,0.42,0.53) VV:(0.05,0.05,0.19,0.18,0.05,0.38,0.05,0.05) O3:(0.07,0.93)	(1,12,13,16,18)	0.9015	0.0985	4.9265
	1000 (7,8,3,5,2,2)	HD:(0.06,0.23,0.06,0.18,0.21,0.21,0.05) HR:(0.08,0.05,0.20,0.05,0.22,0.27,0.08,0.05) TMP: (0.33,0.33,0.34) ; DV:(0.17,0.26,0.26,0.24,0.07) VV:(0.07,0.93) ; O3:(0.08,0.92)	(7,12,17,18)	0.8986	0.1014	9.4481
B	8000 (4,8,3,4,3,3)	HD:(0.15,0.15,0.49,0.21) HR:(0.07,0.08,0.12,0.07,0.15,0.18,0.28,0.05) TMP: (0.13,0.36,0.51) ; DV:(0.24,0.28,0.24,0.24) VV:(0.33,0.33,0.34) ; O3:(0.05,0.13,0.82)	(7,12,18)	0.7965	0.2035	0.0208
	8000 (5,2,3,9,5,2)	HD:(0.20,0.20,0.20,0.20,0.20) TMP: (0.33,0.33,0.34) DV:(0.05,0.06,0.05,0.15,0.07,0.10,0.18,0.26,0.08) VV:(0.20,0.20,0.20,0.20,0.20) ; O3:(0.10,0.90)	(7,12,14,18)	0.8987	0.1013	1.4699
	4000 (4,3,5,3,3,3)	HD:(0.36,0.23,0.10,0.31) HR:(0.05,0.50,0.45) ; TMP: (0.20,0.20,0.20,0.20,0.20,0.20) DV:(0.34,0.33,0.33) ; VV:(0.53,0.39,0.08) O3:(0.05,0.72,0.23)	(1,12,18)	0.8157	0.1843	0.0218
	4000 (6,9,5,6,9,4)	HD:(0.17,0.14,0.20,0.12,0.29,0.08) HR:(0.05,0.11,0.11,0.12,0.12,0.12,0.12,0.12,0.13) TMP: (0.13,0.15,0.42,0.15,0.15) DV:(0.50,0.09,0.11,0.10,0.10,0.10) VV:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) O3:(0.05,0.60,0.05,0.30)	(1,12,18)	0.7690	0.2310	0.0229
	2000 (4,6,2,8,4,3)	HD:(0.39,0.25,0.12,0.24) HR:(0.16,0.16,0.17,0.17,0.17,0.17) ; TMP:(0.69,0.31) DV:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) VV:(0.08,0.13,0.65,0.14) ; O3:(0.05,0.66,0.29)	(7,12,18)	0.7984	0.2016	0.0194
	2000 (7,7,3,9,2,3)	HD:(0.19,0.05,0.16,0.07,0.24,0.24,0.05) HR:(0.17,0.08,0.09,0.20,0.05,0.30,0.11) TMP:(0.33,0.33,0.34) DV:(0.13,0.06,0.15,0.08,0.07,0.08,0.14,0.11,0.18) VV:(0.36,0.64) ; O3:(0.05,0.08,0.87)	(1,12,18)	0.8370	0.1630	0.0219
	1000 (6,3,3,9,3,3)	HD:(0.15,0.30,0.17,0.08,0.08,0.22) HR:(0.33,0.20,0.47) ; TMP: (0.14,0.76,0.10) DV:(0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.11,0.12) VV:(0.40,0.10,0.50) ; O3:(0.05,0.69,0.26)	(12,13,18)	0.8322	0.1678	0.0189
	1000 (6,3,6,3,5,4)	HD:(0.12,0.12,0.13,0.13,0.13,0.37) HR:(0.33,0.33,0.34) ; TMP:(0.27,0.12,0.10,0.23,0.21,0.07) DV:(0.47,0.53) ; VV:(0.53,0.47) ; O3:(0.10,0.66,0.24)	(1,12,18)	0.8003	0.1997	0.0190
C	(3,3,3,3,3,3)	Equal Frequency Partition method	(1,12,14,18)	0.5795	0.4205	0.6336
D	(3,4,2,2,2,3)	Equal Frequency Partition method	(1,12,14,18)	0.5846	0.4154	0.5984
E	(3,4,2,2,2,3)	HD:(0.34,0.28,0.38) HR:(0.19,0.34,0.23,0.24) ; TMP: (0.52,0.48) DV:(0.47,0.53) ; VV:(0.53,0.47) ; O3:(0.10,0.66,0.24)	(1,12,18)	0.7946	0.2054	0.0187

**Tabla 77.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo  $MSE_{entr}$ . **Mes Enero.**

# eval	Gran.	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A 8000	(2,4,2,4,3,2)	HD:(0.22,0.78) ; HR:(0.09,0.09,0.20,0.62) TMP:(0.47,0.53) DV:(0.22,0.22,0.21,0.35) ; VV:(0.33,0.33,0.34) O3:(0.45,0.55)	(3,12,15,16,18)	0.5977	30.1209	0.0198
4000	(2,3,4,2,9,3)	HD:(0.74,0.26) ; HR:(0.70,0.13,0.17) TMP:(0.25,0.25,0.25,0.25) ; DV:(0.50,0.50) VV:(0.10,0.10,0.10,0.10,0.19,0.10,0.10,0.10,0.11) O3:(0.68,0.14,0.18)	(10,12,13,16,18)	0.7485	26.2010	0.0197
2000	(2,6,2,2,3,3)	HD:(0.74,0.26) HR:(0.05,0.25,0.09,0.27,0.07,0.27) TMP:(0.50,0.50) ; DV:(0.50,0.50) VV:(0.09,0.31,0.60) ; O3:(0.71,0.13,0.16)	(4,12,13,16,18)	0.7622	31.8716	0.0190
1000	(2,5,2,3,3,3)	HD:(0.46,0.54) ; HR:(0.27,0.06,0.26,0.06,0.35) TMP:(0.50,0.50) ; DV:(0.11,0.52,0.37) VV:(0.34,0.10,0.56) ; O3:(0.29,0.16,0.55)	(3,12,15,18)	0.6025	38.0939	0.0217
B 8000	(3,7,7,2,8,8)	HD:(0.38,0.33,0.29) HR:(0.17,0.11,0.13,0.25,0.11,0.11,0.12) TMP:(0.13,0.15,0.13,0.17,0.13,0.14,0.15) DV:(0.37,0.63) VV:(0.07,0.07,0.25,0.05,0.05,0.20,0.21,0.10) O3:(0.07,0.22,0.11,0.07,0.08,0.29,0.08,0.08)	(4,12,13,16,18)	0.5311	30.3373	0.0180
8000	(3,2,2,5,3,2)	HD:(0.51,0.24,0.25) ; HR:(0.50,0.50) TMP:(0.39,0.61) ; DV:(0.17,0.36,0.21,0.07,0.19) VV:(0.32,0.37,0.31) ; O3:(0.83,0.17)	(8,13,17,18)	0.7711	38.0946	0.0180
4000	(3,4,4,2,3,2)	HD:(0.33,0.33,0.34) ; HR:(0.37,0.32,0.16,0.15) TMP:(0.26,0.25,0.24,0.25) ; DV:(0.75,0.25) VV:(0.66,0.16,0.18) ; O3:(0.53,0.47)	(1,4,12,16,18)	0.7141	40.5327	0.0176
4000	(3,2,8,2,4,2)	HD:(0.33,0.33,0.34) ; HR:(0.73,0.27) TMP:(0.11,0.25,0.10,0.10,0.11,0.11,0.11,0.11) DV:(0.76,0.24) ; VV:(0.25,0.25,0.25,0.25) O3:(0.56,0.44)	(1,4,12,16,18)	0.7379	40.2915	0.0178
2000	(4,5,6,2,3,2)	HD:(0.25,0.25,0.25,0.25) HR:(0.18,0.07,0.30,0.13,0.32) TMP:(0.19,0.06,0.60,0.05,0.05,0.05) DV:(0.45,0.55) ; VV:(0.33,0.33,0.34) O3:(0.83,0.17)	(8,13,17,18)	0.7545	42.5093	0.0169
2000	(3,3,7,2,5,5)	HD:(0.31,0.31,0.38) ; HR:(0.33,0.33,0.34) TMP:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) DV:(0.50,0.50) ; VV:(0.41,0.05,0.05,0.39,0.10) O3:(0.07,0.34,0.18,0.21,0.20)	(1,10,12,16,18)	0.6093	39.5302	0.0180
1000	(4,2,6,9,3,2)	HD:(0.25,0.25,0.25,0.25) ; HR:(0.50,0.50) TMP:(0.09,0.23,0.05,0.32,0.05,0.26) DV:(0.07,0.11,0.2,0.06,0.05,0.06,0.16,0.15,0.14) VV:(0.33,0.33,0.34) ; O3:(0.87,0.13)	(8,13,17,18)	0.7808	40.0559	0.0170
1000	(2,4,4,2,4,4)	HD:(0.71,0.29) ; HR:(0.32,0.09,0.31,0.28) TMP:(0.58,0.14,0.07,0.21) ; DV:(0.76,0.24) VV:(0.32,0.06,0.57,0.05) ; O3:(0.51,0.18,0.11,0.20)	(10,12,13,16,18)	0.6587	40.5973	0.0181
C	(3,3,3,3,3,3)	Equal Frequency Partition method	(4,7,12,18)	0.6816	-----	0.0194
D	(3,2,2,8,2,2)	Equal Frequency Partition method	(7,12,16,18)	0.6683	60.89	0.0180
E	(3,2,2,8,2,2)	HD:(0.23,0.53,0.24) HR:(0.49,0.51) ; TMP:(0.54,0.46) DV:(0.11,0.14,0.13,0.15,0.14,0.13,0.11,0.09) VV:(0.46,0.54) ; O3:(0.79,0.21)	(2,13,15,17,18)	0.7268	37.7292	0.0169

**Tabla 78.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando la función objetivo  $MSE_{entr}$ . **Mes Agosto.**

# eval	Gran.	Proporción de datos	Opt. Mask	Q	$MSE_{entr}$	$RMS_{prue}$
A 8000	(3,3,9,3,2,3)	HD:(0.30,0.27,0.43) ; HR:(0.33,0.33,0.34) TMP:(0.07,0.23,0.07,0.19,0.08,0.12,0.07,0.09,0.08) DV:(0.33,0.33,0.34) ; VV:(0.53,0.47) O3:(0.30,0.45,0.25)	(11,12,14,17,18)	0.5867	18.1724	0.9152
4000	(3,2,2,4,2)	HD:(0.33,0.33,0.34) HR:(0.42,0.58) ; TMP: (0.49,0.51) DV:(0.21,0.79) ; VV:(0.25,0.25,0.25,0.25) O3:(0.40,0.60)	(1,8,12,16,18)	0.5562	19.7815	2.7428
2000	(2,2,2,3,2,2)	HD:(0.68,0.32) HR:(0.50,0.50) ; TMP: (0.30,0.70) DV:(0.33,0.33,0.34) ; VV:(0.50,0.50) O3:(0.24,0.76)	(3,12,14,15,18)	0.6702	21.0331	0.1255
1000	(8,4,4,7,4,2)	HD:(0.12,0.12,0.11,0.24,0.12,0.05,0.12,0.12) HR:(0.21,0.40,0.07,0.32) TMP:(0.25,0.25,0.25,0.25) DV:(0.08,0.32,0.10,0.17,0.11,0.13,0.09) VV:(0.21,0.27,0.31,0.21) ; O3:(0.18,0.82)	(6,12,14,18)	0.7368	22.0196	1.3107
B 8000	(2,3,8,8,2,4)	HD:(0.64,0.36) ; HR:(0.34,0.33,0.33) TMP:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) DV:(0.10,0.10,0.10,0.22,0.10,0.11,0.11,0.16) VV:(0.56,0.44) ; O3:(0.26,0.21,0.26,0.27)	(1,5,12,14,18)	0.5433	18.8495	0.0199
8000	(2,3,3,3,2,3)	HD:(0.75,0.25) HR:(0.33,0.33,0.34) ; TMP: (0.33,0.33,0.34) DV:(0.36,0.31,0.33) ; VV:(0.49,0.51) O3:(0.26,0.46,0.28)	(5,12,13,14,18)	0.6215	19.0894	0.0200
4000	(3,5,2,3,5,5)	HD:(0.33,0.33,0.34) HR:(0.20,0.20,0.20,0.20,0.20) ; TMP: (0.37,0.63) DV:(0.38,0.30,0.32) ; VV:(0.14,0.20,0.20,0.20,0.26) O3:(0.20,0.20,0.20,0.20,0.20)	(7,12,17,18)	0.5263	24.4285	0.0205
4000	(2,5,2,5,4,3)	HD:(0.50,0.50) HR:(0.12,0.13,0.48,0.13,0.14) ; TMP:(0.40,0.60) DV:(0.15,0.21,0.22,0.21,0.21) VV:(0.25,0.25,0.25,0.25) ; O3:(0.33,0.33,0.34)	(1,5,12,15,18)	0.5449	21.7027	0.0216
2000	(3,7,8,3,4,4)	HD:(0.33,0.33,0.34) HR:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) TMP:(0.17,0.05,0.05,0.17,0.05,0.05,0.05,0.41) DV:(0.23,0.35,0.42) ; VV:(0.25,0.25,0.25,0.25) O3:(0.23,0.33,0.22,0.22)	(4,7,12,16,18)	0.5344	23.0381	0.0175
2000	(3,2,8,7,6,5)	HD:(0.27,0.36,0.37) ; HR:(0.50,0.50) TMP:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) DV:(0.21,0.06,0.30,0.08,0.25,0.05,0.05) VV:(0.16,0.16,0.17,0.18,0.17,0.16) O3:(0.20,0.20,0.20,0.20,0.20)	(1,12,17,18)	0.5203	24.7734	0.0204
1000	(3,4,6,3,5,5)	HD:(0.33,0.33,0.34) ; HR:(0.25,0.25,0.25,0.25) TMP:(0.16,0.16,0.17,0.17,0.17,0.17) DV:(0.33,0.33,0.34) ; VV:(0.20,0.20,0.20,0.20,0.20) O3:( 0.2, 0.2, 0.2, 0.2, 0.2)	(4,7,12,16,18)	0.5150	25.9782	0.0174
1000	(3,2,3,7,4,4)	HD:( 0.33,0.33,0.34) HR:(0.37,0.63) ; TMP: (0.33,0.33,0.34) DV:(0.05,0.06,0.14,0.05,0.07,0.26,0.37) VV:(0.25,0.25,0.25,0.25) ; O3:(0.25,0.25,0.25,0.25)	(7,12,14,18)	0.5246	25.4962	0.0208
C	(3,3,3,3,3,3)	Equal Frequency Partition method	(1,12,14,18)	0.5795	-----	0.6336
D	(2,2,2,2,2,3)	Equal Frequency Partition method	(1,8,12,15,18)	0.5481	25.3400	0.0191
E	(2,2,2,2,2,3)	HD:(0.51,0.49) HR:(0.51,0.49) ; TMP: (0.59,0.41) DV:(0.05,0.95) ; VV:(0.58,0.42) O3:(0.31,0.23,0.46)	(7,8,12,15,18)	0.5200	23.6769	0.0174

## 7.4 Análisis de resultados y conclusiones

Este apartado comprende un análisis de los resultados obtenidos por los métodos **AG1**, **AG2** y **AG3** propuestos en esta memoria para la resolución del problema de estimación de concentraciones de ozono en la zona centro del Área Metropolitana de Ciudad de México. Las dos funciones objetivo establecidas serán evaluadas y comparadas desde la perspectiva de su eficiencia y tiempo computacional. Para finalizar se proporcionarán algunas conclusiones.

### 7.4.1 Modelado Mes Enero

#### 7.4.1.1 Método AG1

De la *Tabla 67* se observa que la solución óptima para la función objetivo 1-Q corresponde a la granularidad (5,X,X,X,X,2) con una calidad de 0.6999.

En relación a la función objetivo  $MSE_{entr}$  la solución óptima sugerida por el **AG1** es la partición (2,2,X,X,4,2) con un  $MSE_{entr}$  de 47.94%.

Es interesante analizar los resultados obtenidos por el **AG1** para ambas funciones objetivo. En principio, se puede indicar que todas las máscaras incluyen relaciones causales con valores anteriores del O3 y la variable HD. Estas son dos entradas fundamentales según FIR. En muchas de las máscaras se encuentra que además son relevantes las variables VV y DV o HR, consiguiendo los mejores resultados en la predicción de los conjuntos de datos de prueba. Para las máscaras derivadas de la función objetivo 1-Q la variable DV es más relevante que la HR, mientras que sucede justo lo contrario para la función objetivo  $MSE_{entr}$ .

Por otra parte, las máscaras derivadas de las soluciones subóptimas sugeridas son las que presentan el mejor poder de predicción. Sin embargo, se observa que los modelos FIR obtenidos para ambas funciones objetivo son del mismo orden de magnitud ( $RMS_{prue} = 0.0179$  vs.  $RMS_{prue} = 0.0180$ ). El menor error supera en poco al obtenido por **FIR EXPERTOS** ( $RMS_{prue} = 0.0176$ ) y es inferior al error de **FIR DEFECTO** ( $RMS_{prue} = 0.0194$ ).

Con respecto al tiempo de CPU requerido por ambas funciones objetivo, se observa en la *Tabla 69* que  $MSE_{entr}$  requiere más tiempo que 1-Q para realizar las 30 ejecuciones del AG, como ya se ha justificado anteriormente.

#### **7.4.1.2 Método AG2 con granularidad suministrada por expertos (EXPERTOS+AG2)**

Se destaca de la *Tabla 62* que todas las soluciones referentes a la función objetivo 1-Q son mejores que el resultado presentado en la investigación previa [NM06]. Las calidades de las máscaras identificadas son siempre superiores a 0.80 mientras que la máscara de **FIR EXPERTOS** tiene una calidad muy inferior de 0.58. Lo mismo sucede con las máscaras derivadas de la función objetivo  $MSE_{entr}$  (*Tabla 65*).

En cuanto a la precisión de los modelos (última columna) se observa, una vez más que no siempre se obtiene el menor  $RMS_{prue}$  cuando se usan los mejores modelos obtenidos por FIR. Sin embargo, los menores  $RMS_{prue}$  conseguidos por soluciones subóptimas en ambas funciones objetivo son inferiores a los de referencia, siendo de 0.0160 para la de 1-Q y de 0.0164 para la de  $MSE_{entr}$ .

De la *Tabla 64* se puede observar que se duplica el tiempo de CPU a medida que se aumenta el número de evaluaciones.

#### **7.4.1.3 Método AG2 con granularidad suministrada por AG1 (AG1 + AG2)**

Observando la *Tabla 70* se aprecia que todas las soluciones sugeridas por el AG en las evaluaciones de la función objetivo (valores de 1-Q) son mejores con respecto a aquellas obtenidas cuando se usa por defecto el método EFP para obtener las funciones de pertenencia de cada clase, es decir, el resultado de referencia de la sección *D* de la *Tabla 75*.

En relación a la función objetivo  $MSE_{entr}$  se puede indicar (Véase *Tabla 72*), que todas las soluciones sugeridas por el **AG2** (valores de  $MSE_{entr}$ ) son mejores que la de referencia de la sección *D* (Véase *Tabla 75*). Sin embargo, las mejores predicciones se obtienen de las soluciones subóptimas.

El mejor resultado de predicción logra un valor  $RMS_{prue}$  de 0.0169, siendo inferior al mejor resultado obtenido en estudios previos para el mismo problema (Véase *Tabla 59*).

#### **7.4.1.4 Método AG3**

La *Tabla 75* presenta máscaras con alta calidad en comparación a las obtenidas usando la función objetivo  $MSE_{entr}$ , las encontradas por **FIR EXPERTOS** (*Tabla 59*) y por **FIR DEFECTO** (sección *C* de cada tabla). Sin embargo, la capacidad de predicción de las dos funciones objetivo propuestas en este caso es muy similar, no

destacando una sobre la otra.

El mejor resultado logrado por **FIR EXPERTOS** tiene un  $RMS_{prue}$  de 0.0176 superior al error de 0.0169 obtenido por el **AG3**. Para la función objetivo 1-Q el mejor modelo FIR encontrado obtuvo un  $RMS_{prue}$  de 0.0177.

#### 7.4.1.5 Conclusiones generales

Los sistemas difusos evolutivos (SDEs) propuestos en esta memoria, permiten obtener modelos FIR para el mes de Enero con igual o mejor habilidad de predicción que los obtenidos por [NM06] y **FIR DEFECTO**. La combinación **EXPERTOS + AG2** ha sido el método más eficaz para obtener los mejores resultados de predicción del conjunto de datos de prueba, aunque le siguen muy de cerca **AG1+AG2** y **AG3**. La *Tabla 80* muestra estos resultados. En general, la función objetivo  $MSE_{entr}$  tuvo mejor desempeño que la función objetivo 1-Q.

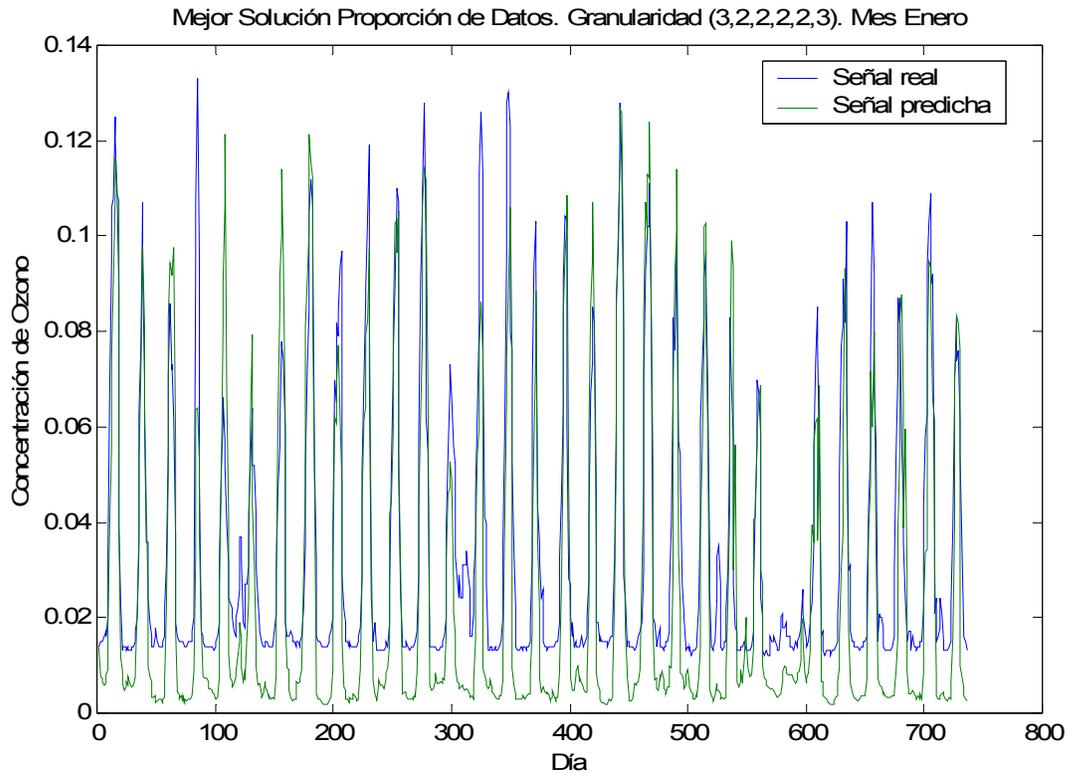
**Tabla 80.**  $ECM_{prue}$  de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para el **Mes de Enero** usando: AG1+EFP, EXPERTOS+AG2, AG1+AG2, AG3, FIR EXPERTOS.

<b>AG1+EFP</b>	<b>EXP.+AG2</b>	<b>AG1+AG2</b>	<b>AG3</b>	<b>FIR DEF.</b>	<b>FIR EXP.</b>
0.0179	0.0160	0.0169	0.0169	0.0194	0.0176

Los mejores modelos de los SDEs identifican las variables HD, VV y valores previos de  $O_3$  como las variables más relevantes para la predicción de las futuras concentraciones de ozono. Estos resultados están en concordancia con los modelos de ecuaciones diferenciales obtenidos en trabajos previos [RO96]. Sin embargo, cuando se usa la función objetivo 1-Q, muchos modelos encuentran relevante, además de las variables anteriormente mencionadas, la variable HR (humedad relativa). Contrariamente, cuando se trata de la función  $MSE_{entr}$  la variable HR no aparece prácticamente en ninguno de los modelos identificados. En este caso es la variable DV (dirección del viento) la que toma un papel relevante en la mayoría de las máscaras identificadas. Consideramos que los modelos derivados del uso de la función  $MSE_{entr}$  son más confiables en cuanto a la selección de variables relevantes puesto que tienen mejor desempeño en el proceso de predicción.

La *figura 48* muestra la señal real y de predicción obtenida usando el mejor modelo derivado de la función objetivo  $MSE_{entr}$ . Vemos que la predicción es capaz de seguir de manera bastante precisa la señal real. A pesar que hay picos que no llega a seguir adecuadamente, sí que alcanza la mayoría de ellos. Recordemos que estamos realizando predicciones a largo plazo, de más de 700 valores de golpe, en donde los

valores de ozono en instantes anteriores son datos predichos por el propio modelo FIR. Si se tiene esto en cuenta podemos considerar que el trabajo de predicción realizado por estos modelos es muy bueno.



**Figura 48.** Predicción del mejor modelo FIR para Mes Enero. Problema concentraciones de ozono en México. Función objetivo  $MSE_{entr}$ . Granularidad (3,2,2,2,2,3) con  $Q=0.6548$  y distribución de landmarks: HD:(0.36,0.43,0.21)-HR:(0.58,0.42)-TMP:(0.43,0.57)-DV:(0.46,0.54)-VV:(0.60,0.40)- $0_3$ :(0.31,0.22,0.47). Máscara:(7,8,12,17,18).  $RMS_{prue}=0.0164$ . Método **EXPERTOS+AG2**.

Una cosa más a tener en cuenta es el papel que realiza la variable hora del día. Las predicciones realizadas por modelos FIR que no incluyen la variable HD en ningún instante de tiempo son predicciones muy pobres que no siguen la frecuencia de la señal real. Esto es debido a que HD está realizando el papel de los precursores del ozono. Si tenemos en cuenta que no tenemos variables disponibles que nos registren la carga de circulación, los puestos de carne asada en la calle, etc... esta información está indirectamente capturada por la variable hora del día. Es por ello que esta variable es muy relevante para la predicción del ozono y aparece en todos los modelos.

## 7.4.2 Modelado Mes Agosto

### 7.4.2.1 Método AG1

La *Tabla 68* indica que la solución óptima para la función objetivo 1-Q corresponde a la granularidad (3,4,X,X,X,2) con una calidad de 0.5846. En este caso se sugiere siempre una granularidad de tres clases para la variable hora del día mientras que para la variable de salida se sugieren 2 o 3 clases.

Para la función objetivo  $MSE_{entr}$  la solución óptima encontrada es la partición (X,3,2,X,2,3) con un  $MSE_{entr}$  de 23.44% ya que fue la que presentó el menor  $MSE_{entr}$ . El error de prueba en este caso es bastante elevado de 0.3499.

Los modelos FIR obtenidos con la función objetivo  $MSE_{entr}$  tienen mayor eficiencia en la predicción del conjunto de prueba que los derivados de la función objetivo 1-Q, pero tiene un costo computacional más elevado (Véase *Tabla 69*). El menor  $RMS_{prue}$  fue de 0.0191 siendo éste mayor al error del modelo **FIR EXPERTOS** de la *Tabla 60* ( $RMS_{prue}=0.0180$ ) e inferior al error de **FIR DEFECTO** de 0.6336.

### 7.4.2.2 Método AG2 con granularidad suministrada por expertos (EXPERTOS+AG2)

De la *Tabla 66* se observa que el mejor resultado de predicción cuando se implementa el **AG2** tiene un error significativamente inferior al obtenido por el resultado de referencia (C), es decir, al mejor resultado obtenido cuando se utiliza **FIR EXPERTOS** (Véase *Tabla 60*).

Por otra parte, cabe señalar que en la mayoría de las máscaras (Véanse *Tablas 63 y 66*) aparece la variable humedad relativa como fundamental para obtener una buena predicción. Así mismo, la variable velocidad del viento y hora del día son importantes para obtener un buen modelo.

### 7.4.2.3 Método AG2 con granularidad suministrada por AG1 (AG1 + AG2)

De las *Tablas 71 y 73* se confirma, una vez más, que la función objetivo  $MSE_{entr}$  proporciona particiones que derivan en máscaras con más poder de predicción que las obtenidas usando la función 1-Q. También aquí, las relaciones causales de las variables HD, HR con la variable de salida  $O_3$ , son las que consiguen mayor habilidad de predicción. Cabe destacar que la temperatura también aparece muy frecuentemente en las máscaras encontradas. El menor error encontrado por este método es mayor al encontrado por el método **EXPERTOS+AG2**.

#### 7.4.2.4 Método AG3

En las *Tablas 76 y 78* se observa que los menores errores de predicción se logran con las máscaras de la función objetivo  $MSE_{entr}$  ( $RMS_{prue}=0.0174$ ).

Si comparamos los menores  $RMS_{prue}$  obtenidos por el **AG3** en ambas funciones objetivo con los resultados de **FIR DEFECTO** y **AG1** se tiene que estos últimos tienen errores superiores a aquellos. En este caso los errores más pequeños se logran con la combinación **AG1+AG2**.

#### 7.4.2.5 Conclusiones generales

Los SDEs obtuvieron de nuevo modelos FIR para el mes de Agosto con mejor habilidad de predicción que los obtenidos por [NM06] y **FIR DEFECTO**. Al igual que en el modelado del mes de Enero, la combinación **EXPERTOS + AG2** ha sido el método más eficaz para obtener los mejores resultados y la función objetivo  $MSE_{entr}$  la que tuvo mejor desempeño. La *Tabla 81* muestra esta información resumida. De este modo, se señala que los modelos FIR obtenidos usando los SDEs son buenos modelos para predecir el comportamiento real del ozono en la ciudad de México.

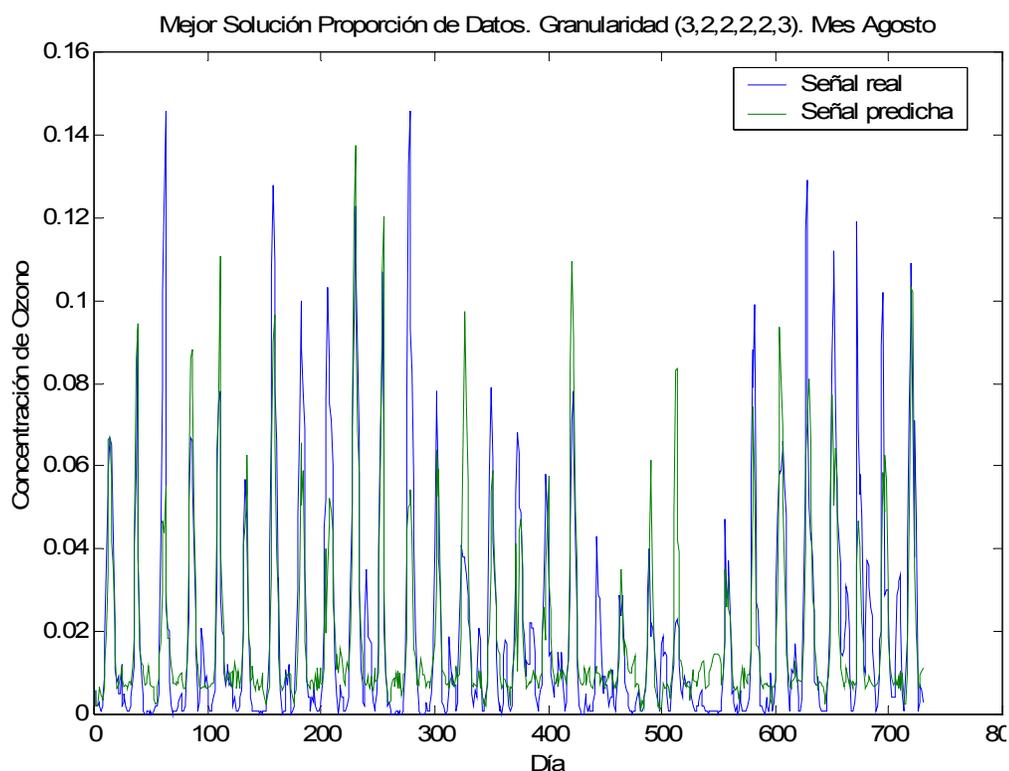
**Tabla 81.**  $ECM_{prue}$  de los mejores resultados obtenidos (desde la perspectiva del error de predicción del conjunto de datos de prueba) para el **Mes de Agosto** usando: AG1+EFP, EXPERTOS+AG2, AG1+AG2, AG3, FIR EXPERTOS.

<b>AG1+EFP</b>	<b>EXP.+AG2</b>	<b>AG1+AG2</b>	<b>AG3</b>	<b>FIR DEF.</b>	<b>FIR EXP.</b>
0.0191	0.0164	0.0174	0.0174	0.6336	0.0180

Los mejores modelos de los SDEs del mes de agosto identifican mayoritariamente las variables HD, HR, VV y valores pasados del  $O_3$  como las más relevantes para la predicción de las futuras concentraciones de ozono. Este hecho está en concordancia con los trabajos previos [NM06] en donde se determinó que la humedad relativa era una variable fundamental en los modelos del período lluvioso, como es el caso del mes de agosto.

Las *figura 49* corresponde al mejor modelo FIR, desde la perspectiva del error de predicción del conjunto de datos de prueba, para el mes de Agosto usando los SDEs desarrollados. Analizando la figura vemos que los picos altos de la señal se alcanzan en la mayoría de las ocasiones, aunque hay picos que no se llegan a predecir correctamente. Hay que recordar aquí que se están realizando predicciones a largo plazo, es decir la predicción de todo el mes (más de 700 valores) se realiza en una sola ejecución y los valores de ozono en instantes previos son en cada momento los

valores del ozono predichos por el mismo modelo FIR. Bajo esta perspectiva, los errores se pueden considerar muy bajos y los modelos encontrados están haciendo un buen trabajo. Estos modelos pueden usarse en la práctica para realizar predicciones reales de la concentración de ozono hora a hora, día a día o semana a semana con precisión suficiente para garantizar su utilidad.



**Figura 49.** Predicción del mejor modelo FIR para **Mes Agosto**. Problema concentraciones de ozono en México. Función objetivo  $MSE_{entr}$ . Granularidad  $(3,2,2,2,2,3)$  con  $Q=0.5772$  y distribución de landmarks:  $HD:(0.36,0.37,0.27)$ - $HR:(0.50,0.50)$ - $TMP:(0.57,0.43)$ - $DV:(0.60,0.40)$ - $VV:(0.52,0.48)$ - $0_3:(0.32,0.33,0.35)$ . Máscara:  $(5,12,13,14,18)$ .  $RMS_{prue} = 0.0164$ . Método **EXPERTOS+AG2**.



## **CAPITULO VIII:**

---

**Una Aportación Adicional. Método  
Multiobjetivo de Aprendizaje para  
Razonamiento Inductivo**



## Capítulo VIII: Una Aportación Adicional. Método Multiobjetivo de Aprendizaje para Razonamiento Inductivo

En el Capítulo V se presentó un método de aprendizaje automático que permite obtener a la vez los parámetros de fusificación de la Metodología FIR (número de clases y funciones de pertenencia) usando funciones monobjetivo para la evaluación de los cromosomas. En este capítulo se propone otro método para el mismo fin pero ahora utilizando AGs Multiobjetivo (AGMOs).

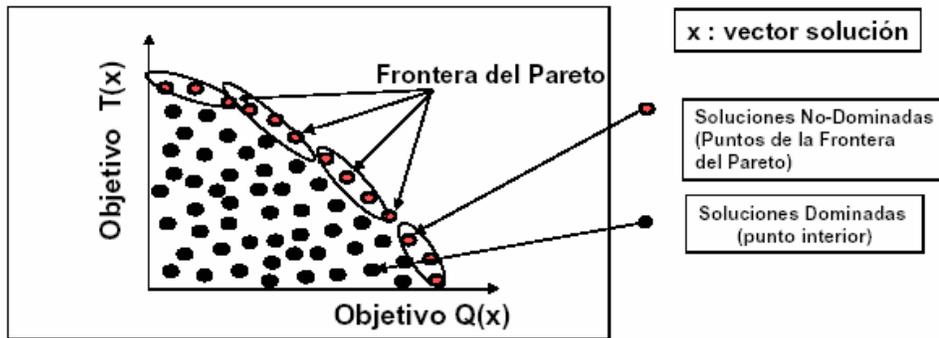
Este capítulo comprende cinco secciones. En un principio, se presenta una reseña del estado del arte de los AGMOs. Los siguientes apartados establecen las pautas del método multiobjetivo propuesto, sus componentes principales y las ventajas e inconvenientes de su utilización. Por último, se muestran los resultados preliminares obtenidos para el caso de estudio: Concentraciones de ozono en la Ciudad de México.

### 8.1 Acerca de los algoritmos genéticos multiobjetivo

La mayoría de los problemas de optimización del mundo real [BSEK06] tienen un nivel de complejidad elevado debido a que poseen varias funciones objetivo, si bien éstas tienden a simplificarse transformándolas (todas excepto una) en restricciones adicionales al problema. De este modo, un problema de optimización multiobjetivo (MOP) se puede definir como el problema de encontrar un vector de variables de decisión que satisfaga ciertas restricciones y optimice un vector de funciones cuyos elementos representan las funciones objetivo [CVL02, Deb01, VL00a]. Estas funciones forman una descripción matemática de criterios de rendimiento que están normalmente en conflicto entre ellos. Por la tanto, el término “optimización” se refiere a la búsqueda de una solución tal que contenga valores aceptables para todas las funciones objetivo.

En general, la optimización multiobjetivo no se restringe a la búsqueda de una única solución, sino de un conjunto de soluciones llamadas *soluciones no-dominadas*. Cada solución de este conjunto se dice que es un *óptimo de Pareto* y, al representarlas

en el espacio de los valores de las funciones objetivo, conforman lo que se conoce como *frente o frontera de Pareto* (Véase figura 50). Dado un problema concreto, la obtención del frente de *Pareto* es la principal finalidad de la optimización multiobjetivo.



**Figura 50.** *Frontera del Pareto para [Max Q(x), Max T(x)].*

Las diversas técnicas existentes para obtener el frente de Pareto se pueden clasificar en 3 categorías: enumerativas, determinísticas y estocásticas [CVL02]. En los últimos años, los métodos estocásticos han sido ampliamente estudiados; en particular, un gran número de autores han trabajado en el área de *Algoritmos Evolutivos (AEs)* [FF95, Coe99, VL00a], sobre todo en *Algoritmos Genéticos (AGs)*. Estos métodos no garantizan la solución óptima, pero ofrecen soluciones aceptables para un amplio rango de problemas de optimización en los que los métodos deterministas encuentran dificultades. La búsqueda enumerativa, que es un método determinista en el que no se emplea heurística alguna, constituye una estrategia conceptualmente simple de optimización, y está basada en la evaluación de cada posible solución dado un espacio de búsqueda finito. El inconveniente de esta técnica es su inherente ineficiencia, ya que puede ser computacionalmente costosa e incluso prohibitiva a medida que el espacio de búsqueda crece.

El uso de los AEs para resolver problemas de optimización multiobjetivo se plantea por primera vez a finales de los 1960s en la tesis doctoral de Rosenberg [Ros67]. Sin embargo, el primer intento real de atacar un problema multiobjetivo fue el *Vector Evaluated Genetic Algorithm (VEGA)* desarrollado por Schaffer en su tesis doctoral y presentado en la 1<sup>ra</sup> Conferencia Internacional de Algoritmos Genéticos y sus Aplicaciones celebrada en Pittsburg, 1985 [Sch84, Sch85]. Esta nueva área de investigación denominada *Optimización Multiobjetivo Evolutiva (EMOO)*, ha crecido considerablemente como indica el incremento notable (significativo en los últimos 15 años) de artículos técnicos en conferencias internacionales y revistas, sesiones

especiales y grupos de interés en Internet<sup>23</sup>. Todos los algoritmos desarrollados en esta área son llamados *Algoritmos Evolutivos Multiobjetivo (AEMOs)*.

En el Capítulo II se ha señalado que existen cuatro tipos de AEs bien definidos que han servido como base a la mayoría del trabajo desarrollado en el campo de la IA: los *Algoritmos Genéticos (AGs)*, las *Estrategias Evolutivas (EEs)*, la *Programación Evolutiva (PE)* y la *Programación Genética (PG)*. Sin embargo, en este capítulo se describe una aportación adicional que se basa en el uso de AGs para resolver problemas de optimización multiobjetivo, es decir, *Algoritmos Genéticos Multiobjetivo (AGMOs)*. A continuación se describen algunos aspectos relacionados a la nueva temática tratada que son de gran interés.

En un sentido general, los AGMOs se definen como aquellos AGs que permiten resolver problemas de optimización multiobjetivo, es decir, problemas en donde se deben satisfacer diferentes objetivos. Por ejemplo, un sistema de control de aire acondicionado debería: a) Minimizar el consumo de energía, b) Maximizar el confort y c) Maximizar la estabilidad del sistema de control. Por lo tanto, no existe una única solución sino que es necesario satisfacer varias simultáneamente (que forman la frontera de *Pareto*).

El mecanismo de trabajo de los AGMOs es parecido al de los AGs tradicionales, estos requieren de una evolución de varias poblaciones soluciones al problema y el uso de mecanismos que mantengan la diversidad en la población para conseguir un conjunto de soluciones no dominadas lo más grande posible. Asimismo, se inician con un conjunto de configuraciones aleatorias llamada población inicial. Cada individuo (cromosoma) en la población representa una solución al problema de optimización. En cada generación, los individuos son evaluados usando una función de adaptabilidad (*fitness*). Basados en ese valor algunos individuos, llamados padres, son seleccionados. La probabilidad de selección de un individuo está relacionada con su adaptabilidad. Se busca asignar una mayor probabilidad de selección a los mejores individuos. Luego un número de operadores genéticos son aplicados a los padres para producir nuevos individuos que formarán parte de la nueva población. El proceso continúa intentando obtener soluciones cada vez mejores hasta que un criterio de parada sea satisfecho.

Pueden considerarse en general, dos tipos principales de AEMOs:

---

<sup>23</sup> C. Coello ha elaborado un portal en Internet sobre Optimización Multiobjetivo Evolutiva, titulado "Evolutionary Multiobjective Optimization (EMOO) Web Page", con muchos enlaces a artículos de revistas, artículos de conferencias, libros, capítulos de libros, reportes técnicos, tesis de master, tesis doctorales, eventos actuales, software de dominio público y principales grupos de trabajo alrededor del mundo. Disponible vía <http://www.lania.mx/~ccoello/EMOO/>

1. Los algoritmos que no incorporan el concepto de óptimo de *Pareto* en el mecanismo de selección del algoritmo evolutivo sino que usan funciones agregativas lineales. Estos implementan un modelo evolutivo utilizando pesos para la agregación de los objetivos. Entre ellos están *Vector Evaluated Genetic Algorithm* (VEGA) [Sch84, Sch85], *Variable Objective Weighting GA* (VOW-GA) [HL92], *Random Weights GA* (RW-GA) [MI95, TIG98], etc.
2. Los algoritmos que jerarquizan a la población de acuerdo a si un individuo es no dominado (usando el concepto de óptimo de *Pareto*), es decir, utilizan un modelo evolutivo que genera poblaciones de soluciones no dominadas. Entre ellos se encuentran *Multiobjective Optimization Genetic Algorithm* (MOGA) [FF93], *Non-dominated Sorting GA* (NSGA) [SD93, SD94], *Niched Pareto GA* (NPGA) [HN93], NSGA II [DAPM00, DPAM02], *Multi-Sexual GA* [LE97], *Pareto Archived Evolution Strategy* (PAES) [KC99a], *micro-Genetic Algorithm* (micro-GA) [CT01a, CT01b], micro-GA2 [TC03], etc.

Históricamente se puede considerar la existencia de dos generaciones de AEMOs:

1. **Primera Generación:** Caracterizada por el uso de jerarquización de *Pareto* y nichos. Algoritmos relativamente simples. También se produjeron enfoques más rudimentarios (por ejemplo, funciones agregativas lineales). Entre los más destacados están VEGA, MOGA, NSGA, NPGA, etc.

Como se ha señalado el precursor del primer AEMO fue David Schaffer [Sch84, Sch85] quien extendió el programa GENESIS de Grefenstette [Gre84] para incluir funciones multiobjetivo. El enfoque de Schaffer utilizó una extensión del *Simple Genetic Algorithm* (SGA) que él llamó *Vector Evaluated Genetic Algorithm* (VEGA) y que difiere del primero en la manera en la cual el mecanismo de selección se lleva a cabo. La selección divide a la población, y en cada sub-población se realiza la selección basada únicamente en uno de los objetivos. Es así como en VEGA, la población total es dividida en  $k$  partes iguales (suponiendo que el problema tiene  $k$  objetivos), y en cada una de ellas opera la selección tomando en cuenta sólo una función objetivo. Una vez realizada la selección, las sub-poblaciones se mezclan para obtener una nueva población de tamaño  $N$ , en la cual se aplicarían el resto de los operadores evolutivos. Todo este proceso se repite en cada generación.

La principal ventaja de VEGA es su facilidad de implementación (sólo se necesita modificar el mecanismo de selección de un algoritmo genético simple) y su eficiencia. Sin embargo, un problema evidente de VEGA es que no favorece a los buenos compromisos, sino que prefiere las soluciones mejores en sólo uno de los objetivos. A este problema se le conoce como especiación (por su análogo en genética) y el peligro potencial está en que se podrían tener individuos que Schaffer

llamó “*middling*”<sup>24</sup>. Este problema fue identificado y atacado por Schaffer [Sch85], utilizando restricciones al cruce, al no permitir recombinación entre individuos de la misma sub-población, así como otras heurísticas adicionales durante la selección en vez del uso de la selección aleatoria de individuos del tradicional AG. También se mostró [RPLH89] que si se utiliza selección proporcional, el esquema de VEGA equivale a una combinación lineal de las funciones objetivo. Así mismo, este método no incorpora explícitamente el concepto de dominancia de *Pareto* (de hecho, se opone a él) y no utiliza ningún mecanismo para mantener diversidad en la población.

Algunos usos de VEGA son los siguientes: contaminación de acuíferos [RER94], colocación de actuadores [Rog00], diseño de circuitos combinatorios a nivel de compuertas [CHB00] y manejo de restricciones [SR97, Coe00].

A partir de 1993, diferentes algoritmos evolutivos basados en *Pareto* fueron sugeridos: a) Fonseca y Fleming con el *Multi-Objective Genetic Algorithm* (MOGA), b) Srinivas y Deb con el *Non-dominated Sorting Genetic Algorithm* (NSGA), y c) Horn y Nafpliotis con el *Niched-Pareto Genetic Algorithm* (NPGA), los cuales profundizaron más en esta técnica de optimización multiobjetivo y cuyas características básicas son: a) evaluación de los miembros de una población con base al concepto de no dominancia de Pareto, y b) la preservación de la diversidad de soluciones. La idea de usar asignación de aptitud basada en el concepto de óptimo de Pareto fue sugerida originalmente por David Goldberg [Gol89] a fin de lidiar con las limitantes de VEGA. Goldberg sugirió también el uso de nichos o alguna técnica similar para mantener diversidad. Esto es necesario debido a que, por ruido estocástico, los algoritmos evolutivos tienden a derivar hacia una solución única independientemente de la distribución inicial de la población.

El MOGA de Fonseca y Fleming [FF93] es un algoritmo que implementa una variante de la jerarquización de *Pareto* propuesta por Goldberg y representa un esquema en el cual la jerarquía de cierto individuo depende del número de individuos en la población actual que lo dominan, como se muestra en la siguiente expresión:

$$\text{jerarquía}(x_i, t) = 1 + p_i^{(t)} \quad (53)$$

donde  $p_i^{(t)}$  es el número de individuos en la generación actual que dominan a  $x_i$  en la generación  $t$ . Como se ve, todos los individuos no dominados tendrán jerarquía 1, mientras que a los dominados se les penaliza de acuerdo a la densidad de población de la región correspondiente de la superficie de soluciones compromiso.

A continuación de la jerarquización de los individuos de la población, se

---

<sup>24</sup> Por “*middling*”, Schaffer se refirió a un individuo con eficiencia aceptable, sobre el promedio, pero no excelente para cualquiera de las funciones objetivo.

asignan las aptitudes por medio de una interpolación de las jerarquías, para ello se siguen los siguientes pasos [FF93]:

- a) Ordenar la población con base a las jerarquías de los individuos.
- b) Asignar aptitudes a los individuos interpolando desde el mejor (jerarquía 1) hasta el peor (jerarquía  $n \leq M$ , donde  $M$  es el tamaño de la población) usando el esquema propuesto por Goldberg [Gol89]. La interpolación es usualmente lineal, pero puede no serlo.
- c) Promediar las aptitudes de los individuos con la misma jerarquía, de forma que todos ellos sean muestreados de la misma forma. Este procedimiento mantiene constante la aptitud global de la población a la vez que mantiene una presión de selección adecuada.

Desde el punto de vista de Goldberg y Deb [GD91], es probable que este tipo de asignación de la aptitud produzca una presión selectiva grande que podría provocar una convergencia prematura. Para evitar esto, Fonseca y Fleming usaron el método de formación de nichos para distribuir la población sobre la región del óptimo de *Pareto*, pero en este caso sugirieron compartir la aptitud en el espacio de las funciones objetivo y no en el espacio de las variables de decisión. Asimismo, MOGA utiliza restricciones al cruce.

MOGA tiene las mismas desventajas que la jerarquización de Pareto “pura”. La principal desventaja la constituye la escalabilidad/complejidad porque el algoritmo de jerarquización es  $O(kM^2)$ , donde  $k$  son las funciones objetivo y  $M$  el tamaño de la población. Su mayor ventaja es la amplia literatura al respecto. También ha mostrado el mejor desempeño de los algoritmos evolutivos multiobjetivo de primera generación. Algunos usos de esta técnica son los siguientes: Co-síntesis de sistemas embebidos de hardware-software [DJ98], diseño de controladores magnéticos [SCFG97], diagnóstico de fallas [Mar97], optimización de armaduras planas [NA99] y manejo de bosques [DdWdB01]. Entre algunos esfuerzos por extender el MOGA están su hibridación con redes neuronales [DRFF00] y la agregación de elitismo [DdWdB01].

Por otro lado, Srinivas y Deb [SD93] propusieron el NSGA que utiliza sin cambios la idea de la jerarquización de Goldberg que ahora se basa en la clasificación de individuos (mediante un criterio de no-dominancia distinto al de MOGA) en varias capas o frentes que ellos llaman ondas (*waves*), es decir, mediante fronteras de no-dominancia. La clasificación consiste en agrupar a todos los individuos no dominados en un frente, con un valor de aptitud igual para todos los individuos. Para ello se usa el siguiente procedimiento: los individuos que son no dominados con respecto a toda la población son eliminados y se les asigna una “falsa” aptitud, la cual es proporcional al tamaño de la población, para así proporcionar un potencial

reproductivo igual para todos los individuos de este frente. Posteriormente, se obtienen los individuos no dominados de la población restante (o sea, de la segunda capa) y se les asigna una aptitud “falsa” menor que la asignada a los individuos de la primera capa. Luego se les elimina de la población. El proceso continúa hasta que todos los individuos de la población hayan sido clasificados. Puesto que los individuos en el primer frente tienen el valor de aptitud mayor, consiguen siempre más copias que el resto. Esto permite buscar las regiones no dominadas, consiguiéndose una convergencia rápida de la población hacia tales regiones.

Para mantener diversidad, el NSGA usa compartición de aptitud (en el espacio de las variables de decisión), la cual se aplica sobre los valores “falsos” de aptitud asignados a cada individuo.

El NSGA tiene también las mismas desventajas que la jerarquización de *Pareto* “pura”. En la práctica, el desempeño del NSGA resultó ser el más pobre de entre los algoritmos más representativos de la primera generación (o sea, MOGA y el NPGA). Además, era también el más costoso (computacionalmente hablando). Sin embargo, con este método, al igual que con los otros, puede resolverse cualquier número de objetivos [SD94] y pueden atacar tanto problemas de maximización como de minimización. Algunos usos de esta técnica son los siguientes: control de la calidad del agua [RMG01], diseño de sistemas de control [BW01], diseño de constelaciones [MCH98] y problemas de dinámica de fluidos [MLDP99].

A diferencia de los algoritmos antes descritos que jerarquizan la población completa, Horn y Nafpliotis [HN93] desarrollaron un AGMO, denominado NPGA, cuya idea básica es el uso de la selección mediante torneo, pero se elige a los ganadores del torneo con base a no dominancia de *Pareto* (medida con respecto a una muestra de la población) y no con base a la aptitud. De este modo se evita jerarquizar toda la población.

En el NPGA, cuando se enfrentan dos individuos en un torneo, se comparan ambos contra una fracción de la población (Horn y Nafpliotis utilizan 10%) y, si uno resulta dominado y el otro no, gana el no dominado; pero si los dos son dominados o los dos son no dominados, entonces se realiza un conteo de nichos [GR87], y gana el individuo que esté en una región menos poblada. Es decir, en caso de empates se recurre a un esquema de compartición de aptitud tanto en el espacio de las funciones objetivo como en el de las variables de decisión (se usa *equivalence class sharing*). El tamaño de la población es considerablemente más grande frente a otros enfoques usados para que el ruido del método de selección pueda tolerarse por los nichos surgidos en la población [FF94].

El NPGA es un método muy eficiente, aunque requiere más parámetros que los algoritmos evolutivos multiobjetivo tradicionales. Sin embargo, en los estudios comparativos realizados entre algoritmos de primera generación [ZVL01], normalmente el NPGA resultaba mejor que el NSGA, pero no mejor que el MOGA. Algunas aplicaciones de este método son las siguientes: análisis de espectros experimentales [GLM00], selección de características [EHM00], diseño de sistemas tolerantes a fallas [Sch95b] y diseño de sistemas de tráfico [HG97].

Los AEMOs anteriormente descritos han demostrado ser eficientes en la obtención de múltiples soluciones no dominadas para varios problemas de ingeniería. No obstante, algunos investigadores han sugerido la introducción del elitismo para mejorar sus propiedades de convergencia, por lo tanto, han dado paso a la siguiente generación de AEMOs.

2. **Segunda Generación:** Se introduce el concepto de elitismo en dos formas principales: usando selección ( $\mu + \lambda$ ) y usando una población secundaria. Se busca vencer la complejidad de la jerarquización de *Pareto* ( $O(kM^2)$ ) y de las técnicas tradicionales de nichos ( $O(M^2)$ ). Desde finales de los 1990s son los más usados y entre ellos están *Strength Pareto Evolutionary Algorithm* (SPEA) [ZT98, ZT99], SPEA2 [ZLT01], NSGA-II [DAPM00, DPAM02], NPGA2 [EMH01], *Multiobjective Optimization with Messy GA* (MOMGA) [VL00b], MOMGA-II [ZLV00, ZVL01], PAES [KC99a], *Pareto Envelope-based Selection Algorithm* (PESA) [CKO00], PESA II [CJKO01], micro-GA [CT01a, CT01b], micro-GA2 [TC03], etc.

El SPEA marca el inicio de una segunda generación de AEMOs al introducir el uso de elitismo. Se ha podido demostrar matemáticamente que el elitismo es necesario para garantizar convergencia de un AEMO al verdadero frente de *Pareto*. Este algoritmo fue introducido por Zitzler y Thiele en 1998 [ZT98] como un método que integra diferentes AEMOs, buscando combinar lo mejor de cada uno de ellos. Además del elitismo, el SPEA se diferencia de otros AEMOs de primera generación (MOGA, NSGA, etc) en la manera de la asignación de la aptitud. Este método utiliza un archivo que contiene las soluciones no dominadas encontradas a lo largo del proceso evolutivo (población externa de no dominados). Este conjunto de soluciones no dominadas se suele llamar “conjunto elite” y le permite al algoritmo cubrir de un modo más adecuado el frente de Pareto. Para cada individuo en el sistema externo, se calcula un valor de “fortaleza” (*strength*) que es similar a la jerarquización de MOGA ya que es proporcional al número de soluciones a las que un individuo domina. Luego, la aptitud de cada individuo de la población depende del valor de fortaleza de todos los individuos en el archivo externo que lo dominen [Coe03] (por lo tanto todas las soluciones externas participan en la selección) y se realiza una serie de torneos binarios. Adicionalmente, se usa una técnica de clustering llamada *average linkage method* [Mor80] para mantener diversidad.

El SPEA presenta una desventaja frente a otros AEMOs en términos de eficiencia, debido al tiempo necesario para la aplicación del algoritmo de clustering en el conjunto elite. Experimentos de comparación entre NSGA y SPEA dan mejores resultados para SPEA [ZDT00, ZT99]. Sin embargo, los resultados de NSGA+Elitismo (NSGA-II) son similares a SPEA [ZDT00]. Algunas aplicaciones de este método son: exploración de compromisos en implementaciones de software para algoritmos DSP [ZTB99], planeación de tratamientos médicos [PM01], asignación en instalaciones radiológicas [LMBZ01] y rehabilitación de sistemas de agua potable [CRFC+03].

Una versión revisada de SPEA, denominada SPEA2, fue propuesta por Zitzler y sus colegas en 2001. El SPEA2 [ZLT01] tiene las siguientes diferencias principales con respecto a su precursor: a) incorpora una estrategia de grano fino para asignar aptitud, la cual toma en cuenta la cantidad de individuos que cada solución domina y la cantidad de individuos que la dominan, b) utiliza la técnica del “vecino más cercano” para la valoración de la densidad, dirigiendo la búsqueda en forma más eficiente; y c) tiene un esquema de truncamiento de archivo que garantiza la preservación de soluciones frontera. Algunos usos de esta técnica son: control del tamaño del código y reducción de *bloat* en programación genética [BBTZ01] y diseño de aeronaves [WB03].

Por su parte, PAES propuesto por Knowles y Corne [KC99a], conceptualmente hablando, es el método más simple posible y usa una estrategia evolutiva  $(1 + 1)$ , es decir, de 2 miembros (sólo padre que es mutado para generar un hijo) adicionada con un archivo externo en el que se van almacenando los individuos no dominados encontrados a lo largo del proceso evolutivo. Cuando se encuentra un individuo no dominado es comparado con los individuos del archivo externo, y en caso de que nuevamente sea no dominado, puede entrar al mismo archivo. El mecanismo de PAES para mantener la diversidad consiste en una rejilla adaptativa, que es computacionalmente más eficiente que los métodos de nichos. El uso de la rejilla adaptativa permite retener a los individuos no dominados obtenidos y distribuirlos uniformemente.

Knowles y Corne también experimentaron con estrategias evolutivas  $(1+\lambda)$  y  $(\mu+\lambda)$ , pero aseguran que no hay mejoras significativas y sin embargo sí un aumento en el esfuerzo computacional necesario [KC00]. Entre las aplicaciones de PAES están las siguientes: problemas de telecomunicaciones [KC99b, KOC00] y manejo de bases de datos distribuidas adaptativas [KCO00].

Por otro lado, Corne y otros [CKO00] propusieron un algoritmo, llamado PESA, en el que se tiene una población pequeña principal (interna), y una población secundaria más grande, que corresponde con el archivo externo mencionado antes

(como PAES). PESA usa la rejilla adaptativa de PAES para mantener diversidad. Sin embargo, en este caso, el mecanismo de selección está basado en la medida de agrupamiento (*crowding*) usada por la rejilla auto-adaptativa. Esta misma medida se usa para decidir que soluciones ingresan al archivo externo.

El funcionamiento de PESA se describe a continuación. Durante cada iteración del algoritmo, se seleccionan aleatoriamente individuos del archivo externo, y de ellos se producen los nuevos individuos que integrarán la población principal; cuando ésta población principal se llene, los individuos no dominados se incorporarán al archivo secundario. Este algoritmo ha sido aplicado a problemas de telecomunicaciones [CKO00].

Posteriormente apareció PESA-II, de Corne y otros [CJKO01], cuya diferencia principal con PESA es la selección basada en regiones. En este esquema de selección, la unidad de selección es la hiperrejilla en vez del individuo, lo que significa que la selección considera regiones y no individuos. Esto busca reducir el costo computacional. En una fase siguiente se elige al azar un individuo de cada región seleccionada. Este algoritmo también ha sido usado en problemas de telecomunicaciones [CJKO01].

Deb y sus colegas [DAPM00] propusieron una versión revisada del NSGA, llamada NSGA-II, que además de contar con el uso de elitismo es mucho más eficiente computacionalmente hablando que su predecesor y constituye un algoritmo altamente competitivo en convergencia al *Pareto*, aunque parece que sigue teniendo problemas para generar regiones aisladas del frente de *Pareto* (problemas exploratorios) sobre todo conforme se incrementa el número de funciones objetivo [ZVL01], así como para funcionar adecuadamente con representación binaria [CVL02]. El NSGA-II [DAPM00] utiliza un operador de *crowding* (medida de concentración) para mantener la diversidad que no requiere parámetros, en vez de usar nichos. Asimismo, no utiliza memoria externa como los algoritmos anteriores (SPEA y SPEA2). El mecanismo elitista consiste en elegir los mejores  $P$  individuos de la unión de las poblaciones padre e hijo. Para ello, el NSGA-II utiliza un esquema de selección más (+) en el cual la población de padres se compara con la población de hijos.

El funcionamiento del NSGA-II [DPAM02] se describe a continuación. En cada generación, se crea un conjunto mediante la unión de la actual población y la creada mediante selección, cruce y mutación. De este conjunto se extraen los diferentes frentes (agrupados según el número de soluciones que los dominan). El frente  $F_1$  coincide con el actual frente de *Pareto*. La nueva población se crea incluyendo los frentes (de mejor a peor) hasta alcanzar el tamaño máximo. Si es necesario, el último frente se trunca atendiendo al orden basado en el *crowding*. La medida de *crowding*

se utiliza para seleccionar las soluciones más dispersas entre los individuos del último frente utilizado en la nueva población. Cuanto mayor sea la distancia de *crowding* de una solución al resto de su frente mejor, ya que hay menos concentración en esa zona.

Algunos usos del NSGA-II son los siguientes: optimización de formas [DG01], diseño de sistemas de seguridad [GGW03], extrusión de polímeros [GPVC02], manejo de la calidad del agua [DR03] y terapias de radiación [LSMB03].

Van Veldhuizen y Lamont [VL00b] propusieron la versión multiobjetivo del algoritmo desordenado *messy GA* [Deb91], denominada MOMGA. En la primera fase se generan de manera exhaustiva todos los bloques constructores hasta cierto tamaño pre-especificado usando un procedimiento determinista conocido como inicialización parcialmente enumerativa. La siguiente fase realiza la selección por medio de torneo en la población y, de ser necesario, se reduce su tamaño. Por último, la fase yuxtaposicional genera la población nueva por medio de un operador de recombinación por corte y empalme. Algunas aplicaciones de este método son las siguientes: diseño de controladores [HBP02] y funciones de prueba tradicionales [VL00b].

MOMGA-II fue propuesto por Zydallis y otros [ZVL01] y es la versión multiobjetivo del algoritmo genético desordenado rápido [GDKH93]. En este caso, los autores extendieron el *fast-messy GA*, el cual consta de tres fases: 1) Fase de inicialización, 2) Filtrado de bloques constructores, y 3) Fase yuxtaposicional. La principal diferencia del MOMGA-II con respecto a su antecesor radica en las primeras dos fases. La fase de inicialización utiliza un proceso de inicialización probabilístico que crea un número controlado de clones de los bloques constructores de un cierto tamaño predeterminado. La segunda fase, llamada de filtrado de bloques constructores, consiste en reducir la población mediante un filtrado, de forma que se almacenan los mejores bloques constructores; además, en esta fase se realiza el torneo de selección. La tercera fase es la misma que la de MOMGA. Este método ha sido aplicado a funciones de prueba tradicionales [ZVL01] y ha dado buenos resultados en la mayoría de los casos en comparación con otros AEMOs tales como NSGA, MOGA, NPGA, PAES, SPEA, NSGA-II y MOMGA, pudiendo ser considerado como un algoritmo eficiente y efectivo para atacar aplicaciones reales [ZVL01].

Posteriormente Erickson y otros [EMH01], realizaron en el 2001 una segunda versión del algoritmo NPGA. El NPGA2 jerarquiza toda la población antes de realizar los torneos. Para ello, usa la jerarquización de *Pareto* pero mantiene la selección mediante torneo del NPGA original. La segunda diferencia importante es que se usa una nueva forma de compartición de aptitud en la que se calculan los conteos de nichos utilizando individuos de la siguiente generación, aunque se

encuentre incompleta, en vez de usar los de la generación actual. A esta forma de realizar conteo de nichos se le llama *continuously updated fitness sharing* y fue propuesta previamente por Oei y otros [OGC91]. El NPGA2 ha sido aplicado en la contaminación de acuíferos [EMH01].

Por su parte, Coello y Toscano [CT01a, CT01b] desarrollaron un micro-GA, que es un algoritmo que también tiene una población principal pequeña. Un micro algoritmo genético es un algoritmo genético con un tamaño de población  $\leq$  a 5 individuos. La clave del micro algoritmo genético es realizar un proceso de reinicialización una vez que se alcanza convergencia nominal.

Este algoritmo cuenta con un archivo llamado memoria de la población, y está dividido en una parte reemplazable y una parte no reemplazable, de la cual se eligen aleatoriamente los individuos que formarán la población principal. Esta población pequeña es la que utiliza el algoritmo genético con operadores normales, que cuando converge proporciona los individuos no dominados que entrarán en la memoria externa (archivo externo). Algunos de los individuos en esta memoria externa entrarán periódicamente en la parte reemplazable de la memoria de la población. En total, el micro-GA utiliza tres tipos de elitismo y la rejilla adaptativa de PAES.

El micro-GA [CT01a, CT01b] es uno de los AEMOs más rápidos de la actualidad, superando en varios órdenes de magnitud a algoritmos muy eficientes tales como el NSGA-II. No obstante, el mayor inconveniente que presenta es que requiere un elevado número de parámetros (ocho, de los cuales al menos 3 son críticos para su desempeño). Algunas aplicaciones de esta técnica son las siguientes: diseño de aviones [CCA03], diseño estructural [CVL02] y particiones de sistemas de hardware/software [FMSZ03].

Buscando eliminar la mayor desventaja del micro-GA original, Toscano y Coello propusieron recientemente el micro-GA2 [TC03], el cual usa mecanismos de adaptación en línea que hacen que el algoritmo no requiera ningún parámetro. El micro-GA2 es capaz incluso de decidir cuándo debe detenerse (no requiere un número máximo de generaciones). El único parámetro que el usuario debe proporcionar es el tamaño de la población secundaria (aunque puede usarse un valor por omisión). Hasta ahora las únicas aplicaciones del micro-GA2 han sido funciones de prueba estándar [Tos05].

Para finalizar, se puede indicar que hoy en día los algoritmos no basados en dominancia de *Pareto* siguen siendo usados en algunos dominios (por ejemplo, optimización combinatoria) con relativo éxito. Asimismo, aún no se tiene una visión clara en el área de los AEs respecto a qué tipo de algoritmos multiobjetivo son más adecuados para qué tipo de problemas.

Además, es necesario señalar que aunque se han abordado un volumen considerable de aplicaciones hasta la fecha [Coe03, FDA04], la optimización evolutiva multiobjetivo (EMOO) todavía tiene mucho qué ofrecer en términos de dominios que aún no han sido abordados tales como visión por computador, ajuste de modelos basados en elementos finitos, reconocimiento de patrones, etc. Sin embargo, la nueva fase en la que se encuentra el área amerita mayor dedicación y la generación de ideas más profundas. Posiblemente todavía queden problemas simples por resolver, pero hay que buscarlos con mayor ahínco. Adicionalmente, muchos problemas de dificultad “intermedia” todavía no han sido resueltos adecuadamente. Claramente, el diseño de algoritmos para problemas de alto grado de dificultad será una línea que permanecerá abierta durante algún tiempo. Por lo tanto, la EMOO en la actualidad presenta muchas oportunidades de investigación y su futuro luce aún prometedor.

## **8.2 Generación de particiones difusas adecuadas usando algoritmos genéticos multiobjetivo (AGMO3)**

En la experimentación realizada en los Capítulos VI y VII de esta tesis se observa que cuando se utiliza la función objetivo  $MSE_{entr}$  se obtiene, frecuentemente, granularidades elevadas, que alcanzan en algunas variables su valor máximo de 9 etiquetas. Este no es el caso en la función objetivo 1-Q, con la que se obtienen granularidades de 6 etiquetas como máximo en las variables. Por otro lado, se confirma que la función objetivo  $MSE_{entr}$  permite, generalmente, obtener modelos FIR con mayor poder de predicción (en comparación con los encontrados usando 1-Q), pero con un costo computacional mayor. Así mismo, se hace patente que la medida de calidad  $Q$  usada por la metodología FIR para determinar el poder de predicción de una máscara no está haciendo un buen trabajo, siendo fundamental investigar medidas alternativas de calidad en un futuro próximo. Es por todo ello, que el propósito principal del método multiobjetivo propuesto en este capítulo es la obtención de modelos FIR con un nivel de complejidad reducido (mínimo número de etiquetas por variable) y una alta precisión en las predicciones (menor error MSE de predicción de una parte del conjunto de datos de entrenamiento,  $MSE_{entr}$ ). Desafortunadamente, es difícil conseguir las dos cosas al mismo tiempo y se requerirá alcanzar un compromiso entre ambos objetivos.

Normalmente, los modelos FIR con muchas etiquetas por variable son modelos complejos, éstos poseen una mayor expresividad pero tienen un menor poder de predicción (poca predictividad). La disminución del poder de predicción cuando aumenta el número de etiquetas de las variables del sistema es debido al problema de la maldición de la dimensionalidad. Si discretizamos los datos del sistema en un

número elevado de clases por variable, tendremos un modelo muy expresivo, con un alto nivel de interpretabilidad. Sin embargo, su poder de predicción o predictividad será, probablemente, reducido puesto que el número de estados posibles habrá aumentado considerablemente mientras que el conjunto de datos de entrenamiento se mantiene constante. En esta situación sucederá que el siguiente estado a predecir no se encontrará en la base de reglas patrón, puesto que no se habrá observado previamente. Esto llevará a la situación en que el modelo no pueda predecir ningún estado futuro del sistema y tenga una capacidad nula de predicción. Es así como el proceso de aprendizaje de las particiones difusas para el diseño de un modelo FIR utilizará ahora un AG multiobjetivo, el cual persigue obtener modelos que sean lo más simples posible pero manteniendo el máximo poder de predicción. Para conseguir este balance/equilibrio se consideran los siguientes objetivos:

- *Obtener modelos FIR lo más simples posible:* Para alcanzar este objetivo se penalizan los modelos FIR con niveles de granularidad altos.
- *Obtener modelos FIR lo más precisos posible:* Para lograrlo se minimizará el error de predicción de una parte del conjunto de datos de entrenamiento.

Las secciones siguientes presentan el método propuesto describiendo los componentes principales del AGMO.

## 8.3 Aspectos principales del método propuesto

### 8.3.1 Codificación de las soluciones

Del mismo modo que en el Capítulo V, cada cromosoma está compuesto de dos partes que codifican el número de clases de cada variable (granularidad) y la función de pertenencia más adecuada para cada clase:

- Número de clases ( $C_1$ ): La representación genética utilizada es la misma que ha sido descrita en la Sección 3.2.1.
- Funciones de pertenencia ( $C_2$ ): Esta representación genética toma en cuenta el número de elementos observados para cada variable y ya ha sido descrita en la Sección 4.2.1.

Usaremos, por lo tanto, la misma representación de los cromosomas que se describió en el Capítulo V para el **AG3** y que recordamos aquí:

$$C = C_1 C_2 \quad (54)$$

Si denotamos por  $E_j$  el número de clases de la variable  $j$ , la representación del número de clases para un sistema de  $N$  variables ( $C_1$ ), es definida por:

$$C_1 = (E_1, E_2, \dots, E_N) \quad (55)$$

También, si denotamos por  $D_{ij}$  la proporción de datos para la clase  $i$  de la variable  $j$ , y por  $C_{2j}$  la información de la proporción de datos para todas las clases de la variable  $j$ ,

$$C_{2j} = (D_{1j}, \dots, D_{E_{jj}}) \quad (56)$$

la representación de las funciones de pertenencia ( $C_2$ ) para un sistema de  $N$  variables (incluyendo entradas y salidas) se define como:

$$C_2 = (C_{21}, C_{22}, \dots, C_{2N}) \quad (57)$$

Lógicamente, cada vez que cambie el número de clases y/o la distribución de los landmarks de una variable por la acción de los operadores genéticos, se tendrá que construir la nueva partición difusa correspondiente.

### 8.3.2 Población inicial

La población inicial está compuesta por cuatro bloques. A excepción del primero, el resto posee el mismo número de cromosomas. En ninguno de los casos se admiten cromosomas repetidos. El proceso de generación de la misma se lleva a cabo de la misma forma que se presenta en la Sección 5.2.2 de esta memoria.

### 8.3.3 Evaluación de los cromosomas

Para evaluar un cromosoma se llevan a cabo los siguientes pasos:

- 1) Se decodifica la información del cromosoma, construyendo su correspondiente partición difusa en la estructura de datos que utiliza FIR para representar dichas particiones.
- 2) Se ejecuta el proceso de identificación de modelado cualitativo de la metodología FIR con los datos de entrenamiento, usando la partición construida en el paso anterior. Por lo tanto, se obtiene la máscara asociada a esa partición difusa con la medida más alta de calidad.
- 3) Se calculan los valores de la función de evaluación multiobjetivo:
  - $\Sigma_{tiq}$ : Corresponde a la sumatoria de la granularidad de las variables de entrada y

de la variable de salida cuya relación causal o temporal se incluye en la máscara óptima. Sin embargo, cabe aclarar aquí que si una variable aparece en la máscara en diferentes instantes de tiempo, su granularidad se contabiliza una única vez.

- $MSE_{entr}$ : Indica el error de predicción de una parte del conjunto de datos de entrenamiento. El error cuadrático medio normalizado en porcentaje (MSE), se usa para este fin (Véase ecuación 35).

### 8.3.4 Operadores genéticos

Para el diseño de los operadores genéticos, se ha considerado la estructura especial de los cromosomas (con dos niveles de información distintos, pero fuertemente relacionados) para obtener una definición adecuada que haga el mejor uso posible de la representación adoptada. A continuación se describen los mismos.

#### 8.3.4.1 Selección

Se ha utilizado el mecanismo de selección multiobjetivo MOGA [FF93], que se basa en la definición *Pareto*-optimalidad. Se dice que una solución *domina* a otra cuando tiene valores mejores o iguales que la otra en todos los objetivos salvo en uno, donde tiene un valor estrictamente mejor. El conjunto óptimo de *Pareto* está compuesto por todas las soluciones no dominadas del problema. El operador consta de cinco pasos:

1. A cada individuo se le asigna un rango que es igual al número de soluciones de la población que lo dominan más uno (las soluciones no dominadas tienen rango 1).
2. La población se ordena de forma ascendente de acuerdo con su rango.
3. Se le asigna una probabilidad de selección a cada individuo que depende de su rango (menor rango, mayor probabilidad).
4. Se calcula la media de la probabilidad de selección de cada clase de equivalencia (grupo de cromosomas con el mismo rango) para que todos los individuos de igual rango tengan la misma probabilidad de reproducirse en la siguiente generación.
5. Se crea la nueva población mediante el muestreo universal estocástico de Baker [Bak87].

#### 8.3.4.2 Cruce

Se utilizan los mismos operadores de cruce descritos en la sección 5.2.4.2, éstos toman en cuenta las características de los dos cromosomas padres implicados en el cruce:

- *Cruce cuando los dos padres tienen la misma granularidad en cada una de las variables:* Si los dos cromosomas tienen los mismos valores en  $C_1$  (cada variable tienen el mismo número de etiquetas en ambos padres), se puede suponer que la búsqueda genética ha localizado una zona prometedora del espacio que conviene explotar. Para ello, se utiliza el operador de cruce aritmético no uniforme [Mic96] sobre  $C_2$ , manteniendo el valor de  $C_1$  en ambos descendientes. El funcionamiento de este operador se estableció en la Sección 2.5.3.
- *Cruce cuando los padres tienen distinto número de etiquetas en alguna de las variables:* En este segundo caso, parece interesante utilizar la información contenida en ambos padres para explorar el espacio de búsqueda y tratar de localizar nuevas zonas prometedoras. El funcionamiento del operador de cruce es sencillo, se selecciona aleatoriamente un punto de corte dentro de  $C_1$ , y se cruzan las dos partes  $C_1$  y  $C_2$  de ambos cromosomas de acuerdo al operador clásico de cruce [Mic96].

A continuación, se recuerda como funciona este cruce mediante una representación gráfica. Sean

$$\begin{aligned} C_t &= (E_1, \dots, E_p, E_{p+1}, \dots, E_N, C_{21}, \dots, C_{2p}, C_{2p+1}, \dots, C_{2N}) \\ C'_t &= (E'_1, \dots, E'_p, E'_{p+1}, \dots, E'_N, C'_{21}, \dots, C'_{2p}, C'_{2p+1}, \dots, C'_{2N}) \end{aligned} \quad (58)$$

dos individuos que se van a cruzar en el punto  $p$ , los dos descendientes resultantes son:

$$\begin{aligned} C_t &= (E_1, \dots, E_p, E'_{p+1}, \dots, E'_N, C_{21}, \dots, C_{2p}, C'_{2p+1}, \dots, C'_{2N}) \\ C'_t &= (E'_1, \dots, E'_p, E_{p+1}, \dots, E_N, C'_{21}, \dots, C'_{2p}, C_{2p+1}, \dots, C_{2N}) \end{aligned} \quad (59)$$

De esta forma, el proceso completo de recombinación permitirá al AGMO seguir una tasa adecuada de exploración y explotación del espacio de búsqueda.

### 8.3.4.3 Mutación

Los operadores de mutación seleccionados son los mismos que se han descrito en la Sección 5.2.4.3:

- *Mutación en  $C_1$ :* El operador de mutación seleccionado es el propuesto por [Thr91]. Si el gen a mutar es de  $C_1$ , se aumenta o disminuye en una unidad la granularidad asociada (la decisión es tomada de manera aleatoria). Cuando el valor que se pretende cambiar es el mínimo (2) ó el máximo (9), se realiza el único cambio posible, esto es, incrementar o decrementar en uno la granularidad, respectivamente. Por supuesto, un cambio en  $C_1$  origina una inconsistencia en  $C_2$ , donde se encuentran almacenados los conjuntos

difusos asociados al número de clases (funciones de pertenencia) que tenía antes esa variable. Por tanto, cuando se cambia la granularidad de una variable en  $C_1$  se cambia su parte correspondiente en  $C_2$ , construyendo una partición difusa uniforme (Método EFP) con el nuevo número de clases del cromosoma.

- *Mutación en  $C_2$* : Debido a la naturaleza de los valores almacenados en esta parte del cromosoma, el operador de mutación seleccionado es similar a uno propuesto por Thrift en [Thr91] que ya ha sido descrito en la Sección 4.2.4.3.

## 8.4 Comentarios

Aunque el método propuesto en este capítulo se basa en un mecanismo de selección MOGA que está enmarcado en la primera generación de AGs multiobjetivo, y por lo tanto, no incorpora una selección elitista, creemos que éste puede permitir obtener unos buenos resultados en un primer intento y se prevé que más adelante estos resultados podrían mejorarse utilizando un AG multiobjetivo más eficaz.

La ventaja principal del **AGMO3** estriba en que en una sola ejecución (con un tiempo similar al necesitado por los AGs implementados anteriormente), permite obtener un conjunto de soluciones (unas más simples y otras más precisas para poder escoger) al alcanzar el criterio de parada que se haya estipulado. De este modo, no es necesario ejecutarlo tantas veces como los métodos **AG1**, **AG2** y **AG3** contemplados previamente en esta memoria.

## 8.5 Caso de estudio: Concentraciones de ozono en la Ciudad de México

En este apartado se presentan los resultados obtenidos por el **AGMO3** para el problema de predicción de concentraciones de ozono en la Ciudad de México descrito en el Capítulo VII. Cabe destacar que el AG multiobjetivo propuesto en este capítulo se encuentra aún en fase de ajuste y pruebas.

Se han realizado cuatro ejecuciones (con distinto valor de la semilla) para cada criterio de parada. Se estudia la función multiobjetivo descrita en la sección 8.3.3. En esta aplicación medioambiental el último 25% de la señal de entrenamiento se usa

para la evaluación de la función multiobjetivo y el 75% inicial de la señal se usa para la obtención de los modelos FIR (máscaras y bases de reglas patrón).

Del mismo modo que en el capítulo anterior, se ha trabajado bajo la perspectiva del modelado del comportamiento del ozono durante un mes específico del año [NM06]. Sin embargo, en este caso solo se ha seleccionado el mes de Enero que pertenece a la época seca. Asimismo, se utiliza la máscara candidata descrita en la ecuación 52.

Los valores establecidos para el AG multiobjetivo se presentan en la *Tabla 82*. Las probabilidades de cruce y de mutación son establecidas de acuerdo a [Gre86]. Se emplearon computadoras Pentium IV (2.66 GHz) para realizar todas las simulaciones. En [ANF06c] se muestran todos los resultados obtenidos por el método propuesto en este capítulo. Hay que señalar que se probó con una población inicial de 50 individuos pero no se obtenía mucha diversidad en el conjunto de soluciones. Los mejores resultados obtenidos por el AG multiobjetivo se presentan en las *Tablas 83 a 87*.

**Tabla 82.** Parámetros del AG multiobjetivo para el problema de la estimación de concentraciones locales máximas de ozono en México DF.

Parámetro	Valor
<i>Tamaño de la población</i> (# individuos)	100
<i>Probabilidad de cruce</i>	0.6
<i>Probabilidad de mutación</i>	0.1
<i>Criterio de parada</i> (# evaluaciones de cromosomas)	{2000, 4000, 8000, 16000}

### 8.5.1 Modelos obtenidos por el *AGMO3*

En esta sección se presentan los resultados obtenidos por el *AGMO3* propuesto en este capítulo sobre el problema de modelado de concentraciones de ozono en la ciudad de México. Las *Tablas 83 a 86* muestran los resultados obtenidos por criterio de parada (número de evaluaciones de cromosomas hechas por el AG multiobjetivo). Las tablas están organizadas como se describe a continuación. La primera columna se divide en dos secciones. La sección *A* corresponde al mejor resultado obtenido por el **AGMO3** (valor mínimo de cada parte que conforma la función multiobjetivo, es decir, el menor error  $MSE_{entr}$  y la menor sumatoria de etiquetas conseguidas) en cada una de las 4 ejecuciones. La sección *B* corresponde a los mejores resultados de predicción del **AGMO3** mostrados para las 4 ejecuciones. La segunda columna

**Tabla 83.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 2000 evaluaciones. Mes Enero.

# ejec.	Gran.	Proporción de datos	Opt. Mask / $\Sigma$ etiquetas	Q	MSE <sub>entr</sub>	RMS <sub>prue</sub>
A	1	(2,2,2,3,2,2) HD:(0.73,0.27) ; HR:(0.70,0.30) ; TMP:(0.29,0.71) ; DV:(0.33,0.33,0.34) ; VV:(0.26,0.74) ; O3:(0.71,0.29)	(4,7,12,16,18) 7 etiq.	0.7706	40.7664	0.0196
		(4,2,8,4,7,2) HD:(0.41,0.10,0.10,0.39) ; HR:(0.29,0.71) ; TMP:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) ; DV:(0.3,0.17,0.05,0.48) ; VV:(0.27,0.28,0.05,0.11,0.13,0.07,0.09) ; O3:(0.37,0.63)	(12,18) 2 etiq.	0.6323	141.2987	0.0298
	2	(5,6,2,2,8,3) HD:(0.15,0.15,0.15,0.40,0.15) ; HR:(0.16,0.16,0.17,0.17,0.17,0.17) ; TMP:(0.27,0.73) ; DV:(0.33,0.67) ; VV:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) ; O3:(0.33,0.33,0.34)	(1,12,18) 8 etiq.	0.6505	43.0374	0.0215
		(3,3,3,2,3,2) HD:(0.05,0.84,0.11) ; HR:(0.52,0.11,0.37) ; TMP:(0.07,0.86,0.07) ; DV:(0.50,0.50) ; VV:(0.29,0.08,0.63) ; O3:(0.64,0.36)	(12,18) 2 etiq.	0.6731	127.629	0.0407
	3	(6,3,2,5,3,3) HD:(0.16,0.16,0.17,0.17,0.17,0.17) ; HR:(0.33,0.33,0.34) ; TMP:(0.73,0.27) ; DV:(0.20,0.20,0.20,0.20,0.20) ; VV:(0.33,0.33,0.34) ; O3:(0.33,0.33,0.34)	(1,12,18) 9 etiq.	0.6492	43.1024	0.0216
		(2,3,3,3,3,2) HD:(0.16,0.84) ; HR:(0.33,0.33,0.34) ; TMP:(0.33,0.33,0.34) ; DV:(0.33,0.33,0.34) ; VV:(0.33,0.33,0.34) ; O3:(0.66,0.34)	(12,18) 2 etiq.	0.675	125.9375	0.0404
4	(7,4,4,3,4,9) HD:(0.05,0.05,0.07,0.13,0.35,0.27,0.08) ; HR:(0.25,0.25,0.25,0.25) ; TMP:(0.07,0.44,0.44,0.05) ; DV:(0.36,0.41,0.23) ; VV:(0.19,0.29,0.32,0.20) ; O3:(0.09,0.10,0.13,0.10,0.09,0.10,0.09,0.22,0.08)	(1,12,18) 16 etiq.	0.4822	43.1197	0.0221	
	(2,2,2,2,3,2) HD:(0.32,0.68) ; HR:(0.51,0.49) ; TMP:(0.21,0.79) ; DV:(0.47,0.53) ; VV:(0.33,0.33,0.34) ; O3:(0.75,0.25)	(12,18) 2 etiq.	0.7131	116.3262	0.0386	
B	1	(2,2,2,4,2,2) HD:(0.73,0.27) ; HR:(0.70,0.30) ; TMP:(0.29,0.71) ; DV:(0.25,0.25,0.25,0.25) ; VV:(0.26,0.74) ; O3:(0.71,0.29)	(7,12,16,18) 8 etiq.	0.7721	61.9794	0.0175
	2	(2,2,2,3,2,2) HD:(0.40,0.60) ; HR:(0.24,0.76) ; TMP:(0.27,0.73) ; DV:(0.33,0.33,0.34) ; VV:(0.30,0.70) ; O3:(0.58,0.42)	(1,12,13,18) 4 etiq.	0.6907	59.2999	0.0175
	3	(3,2,6,5,2,6) HD:(0.33,0.33,0.34) ; HR:(0.22,0.78) ; TMP:(0.09,0.51,0.14,0.05,0.12,0.09) ; DV:(0.11,0.20,0.41,0.13,0.15) ; VV:(0.81,0.19) ; O3:(0.29,0.12,0.08,0.10,0.14,0.27)	(1,12,18) 9 etiq.	0.5985	50.504	0.0194
	4	(9,5,2,4,3,5) HD:(0.05,0.21,0.07,0.06,0.05,0.14,0.18,0.18,0.06) ; HR:(0.11,0.46,0.05,0.26,0.12) ; TMP:(0.48,0.52) ; DV:(0.3,0.32,0.26,0.12) ; VV:(0.37,0.50,0.13) ; O3:(0.09,0.36,0.05,0.17,0.33)	(3,12,15,18) 7 etiq.	0.6028	60.6752	0.0204

indica el número de ejecución del **AGMO3**. La tercera y cuarta columnas presentan la solución sugerida por el **AGMO3** (número de clases y distribución de landmarks) para las variables de entrada (HD, HR, TMP, DV, VV) y la variable de salida (O<sub>3</sub>) cuando se cumple el criterio de parada estipulado. El número de elementos de la proporción de datos corresponde al número de clases para esa variable. La quinta columna contiene por un lado, la máscara óptima (en notación de posición) encontrada por FIR para esa partición difusa. Por el otro lado, contiene el valor de la sumatoria de la granularidad de las variables de entrada y de salida que se incluye en esa máscara, pero sin tener en cuenta más de una relación temporal. Es decir, si una variable aparece en la máscara en diferentes instantes de tiempo, se contabiliza una

única vez. La sexta columna muestra la calidad asociada a la máscara óptima encontrada. La séptima columna corresponde al valor del error  $MSE_{\text{entr}}$  de la función multiobjetivo. La última columna muestra el error de predicción RMS para el conjunto de datos de prueba ( $RMS_{\text{prue}}$ ).

En la *Tabla 87* se presenta el tiempo consumido, número de generaciones y número de soluciones no dominadas obtenidas por ejecución y criterio de parada del AG multiobjetivo para el mes considerado.

**Tabla 84.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 4000 evaluaciones. Mes Enero.

# ejec.	Gran.	Proporción de datos	Opt. Mask / $\Sigma$ etiquetas	Q	$MSE_{\text{entr}}$	$RMS_{\text{prue}}$		
A	1	(3,2,2,2,4,2)	HD:(0.20,0.61,0.19) ; HR:(0.52,0.48) TMP:(0.76,0.24) ; DV:(0.58,0.42) ; VV:(0.25,0.25,0.25,0.25) ; O3:( 0.55,0.45)	(12,13,14,17,18) 11 etiq.	0.6517	39.8551	0.0199	
		(3,2,3,2,2,2)	HD:(0.33,0.33,0.34) ; HR:(0.67,0.33) TMP:(0.33,0.33,0.34) ; DV:(0.38,0.62) VV:(0.35,0.65) ; O3:(0.81,0.19)	(6,12,18) 2 etiq.	0.7478	106.5236	0.0393	
	2	(6,2,5,8,3,3)	HD:(0.16,0.16,0.17,0.17,0.17,0.17) ; HR:(0.50,0.50) TMP:(0.20,0.20,0.20,0.20,0.20) DV:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) VV:(0.33,0.33,0.34) ; O3:(0.33,0.33,0.34)	(1,12,18) 9 etiq.	0.6492	43.1024	0.0216	
		(3,2,2,2,2,2)	HD:(0.33,0.33,0.34) ; HR:(0.57,0.43) TMP:(0.45,0.55) ; DV:(0.39,0.61) VV:(0.32,0.68) ; O3:(0.82,0.18)	(6,12,18) 2 etiq.	0.7545	119.4284	0.0387	
	3	(2,2,4,2,4,2)	HD:(0.76,0.24) ; HR:(0.52,0.48) TMP:(0.25,0.25,0.25,0.25) ; DV:(0.21,0.79) VV:(0.25,0.25,0.25,0.25) ; O3:(0.61,0.39)	(12,13,14,17,18) 10 etiq.	0.7208	42.8004	0.0211	
		(2,8,3,8,4,2)	HD:(0.24,0.76) HR:(0.06,0.05,0.05,0.17,0.28,0.05,0.22,0.12) TMP:(0.07,0.59,0.34) DV:(0.14,0.27,0.16,0.05,0.05,0.05,0.21,0.07) VV:(0.46,0.12,0.21,0.21) ; O3:(0.77,0.23)	(12,18) 2 etiq.	0.7185	114.1194	0.0381	
	4	(5,3,7,8,4,2)	HD:(0.05,0.26,0.05,0.19,0.45) ; HR:(0.28,0.53,0.19) TMP:(0.26,0.10,0.22,0.19,0.07,0.09,0.07) ; DV:(0.07,0.12,0.16,0.21,0.05,0.18,0.12,0.09) VV:(0.15,0.19,0.33,0.33) ; O3:(0.60,0.40)	(1,12,18) 7 etiq.	0.7397	43.1676	0.0221	
		(5,3,7,7,4,2)	HD:(0.10,0.05,0.05,0.27,0.53) ; HR:(0.45,0.20,0.35) TMP:(0.23,0.06,0.18,0.14,0.05,0.29,0.05) DV:(0.14,0.14,0.14,0.14,0.14,0.15,0.15) VV:(0.15,0.19,0.33,0.33) ; O3:(0.79,0.21)	(12,18) 2 etiq.	0.7261	111.6386	0.0375	
	B	1	(3,2,2,2,2,2)	HD:(0.33,0.33,0.34) ; HR:(0.27,0.73) TMP:(0.26,0.74) ; DV:(0.77,0.23) VV:(0.28,0.72) ; O3:(0.59,0.41)	(1,12,16,18) 7 etiq.	0.7526	51.4089	0.0169
		2	(9,2,2,7,8,3)	HD:(0.05,0.11,0.18,0.05,0.15,0.05,0.13,0.12,0.16) HR:(0.36,0.64) ; TMP:(0.40,0.60) DV:(0.40,0.14,0.05,0.13,0.12,0.06,0.10) VV:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) O3:(0.29,0.33,0.38)	(7,12,18) 12 etiq.	0.609	85.7378	0.0192
3		(2,3,2,2,2,3)	HD:( 0.67,0.33) ; HR:( 0.33,0.33,0.34) TMP:( 0.38,0.62) ; DV:( 0.66,0.34) VV:( 0.41,0.59) O3:( 0.41,0.41,0.18)	(7,10,12,16,18) 7 etiq.	0.629	48.712	0.0208	
4		(5,3,7,8,4,2)	HD:(0.05,0.26,0.05,0.19,0.45) ; HR:(0.28,0.53,0.19) TMP:(0.26,0.10,0.22,0.19,0.07,0.09,0.07) ; DV:(0.07,0.12,0.16,0.21,0.05,0.18,0.12,0.09) VV:(0.15,0.19,0.33,0.33) ; O3:(0.60,0.40)	(1,12,18) 7 etiq.	0.7397	43.1676	0.0221	

**Tabla 85.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 8000 evaluaciones. Mes Enero.

# ejec.	Gran.	Proporción de datos	Opt. Mask / $\Sigma$ etiquetas	Q	MSE <sub>entr</sub>	RMS <sub>prue</sub>	
A	1	(4,6,8,2,3,3)	HD:(0.28,0.35, 0.2,0.17) HR:(0.16,0.16,0.17,0.17,0.17) TMP:(0.12,0.12,0.12,0.12,0.13,0.13,0.13,0.13) DV:(0.67,0.33) ; VV:(0.33,0.33,0.34) O3:(0.41,0.20,0.39)	(1,4,12,16,18) 9 etiq.	0.6677	43.834	0.0190
		(3,4,3,4,2,2)	HD:(0.33,0.33,0.34) ; HR:(0.25,0.25,0.25,0.25) TMP:(0.33,0.33,0.34) ; DV:(0.25,0.25,0.25,0.25) VV:(0.11,0.89) ; O3:(0.83,0.17)	(12,18) 2 etiq.	0.7522	106.9838	0.0363
	2	(2,3,2,2,2,2)	HD:(0.75,0.25) ; HR:(0.33,0.33,0.34) TMP:(0.55,0.45) ; DV:(0.78,0.22) VV:(0.20,0.80) ; O3:(0.68,0.32)	(4,12,13,16,18) 6 etiq.	0.7486	42.007	0.0191
		(3,4,2,2,2,2)	HD:(0.33,0.33,0.34) ; HR:(0.20,0.20,0.18,0.42) TMP:(0.36,0.64) ; DV:(0.63,0.37) VV:(0.52,0.48) ; O3:(0.84,0.16)	(12,18) 2 etiq.	0.756	106.2713	0.0361
	3	(4,3,2,3,3,2)	HD:(0.25,0.25,0.25,0.25) ; HR:(0.33,0.33,0.34) TMP:(0.13,0.87) ; DV:(0.46,0.48,0.06) VV:(0.33,0.33,0.34) ; O3:(0.78,0.22)	(8,13,17,18) 12 etiq.	0.7562	42.8306	0.0169
		(3,2,3,2,3,2)	HD:(0.19,0.17,0.64) ; HR:(0.76,0.24) TMP:(0.33,0.33,0.34) ; DV:(0.53,0.47) VV:(0.43,0.43,0.14) ; O3:(0.78,0.22)	(12,18) 2 etiq.	0.7254	112.7337	0.0378
A	4	(5,4,3,6,5,3)	HD:(0.15,0.15,0.15,0.15,0.40) HR:(0.14,0.13,0.13,0.60) ; TMP: (0.25,0.25,0.50) DV:(0.16,0.16,0.17,0.17,0.17,0.17) VV:(0.25,0.18,0.19,0.19,0.19) ; O3:(0.33,0.33,0.34)	(1,12,18) 8 etiq.	0.6307	43.0066	0.0215
		(3,3,3,2,2,2)	HD:(0.18,0.65,0.17) ; HR:(0.33,0.33,0.34) TMP: (0.33,0.33,0.34) ; DV:(0.63,0.37) VV:(0.94,0.06) ; O3:(0.84,0.16)	(12,18) 2 etiq.	0.756	106.2713	0.0361
	B	1	(4,4,8,2,3,3)	HD:(0.28,0.35, 0.2,0.17) ; HR:(0.22,0.22,0.23,0.33) TMP:(0.13,0.13,0.13,0.13,0.14,0.14,0.07,0.13) ; DV:(0.67,0.33) ; VV:(0.17,0.63,0.20) O3:(0.17,0.63,0.20)	(1,12,16,18) 9 etiq.	0.651	54.0317
	2	(2,3,2,2,2,2)	HD:(0.75,0.25) ; HR:(0.33,0.33,0.34) TMP:(0.55,0.45) ; DV:(0.78,0.22) VV:(0.20,0.80) ; O3:(0.68,0.32)	(4,12,13,16,18) 6 etiq.	0.7486	42.007	0.0191
	3	(4,3,2,3,3,2)	HD:(0.25,0.25,0.25,0.25) ; HR:(0.33,0.33,0.34) TMP:(0.13,0.87) ; DV:(0.46,0.48,0.06) VV:(0.33,0.33,0.34) ; O3:(0.78,0.22)	(8,13,17,18) 12 etiq.	0.7562	42.8306	0.0169
	4	(4,2,2,2,3,2)	HD:(0.25,0.25,0.25,0.25) ; HR:(0.64,0.36) TMP: (0.54,0.46) ; DV:(0.52,0.48) VV:(0.33,0.33,0.34) ; O3:(0.84,0.16)	(4,11,13,16,18) 11 etiq.	0.7625	55.0864	0.0184

## 8.5.2 Análisis de resultados y conclusiones

Este bloque comprende un análisis de los resultados obtenidos por el AG multiobjetivo propuesto en este capítulo para la resolución del problema de estimación de concentraciones de ozono en la zona centro del Área Metropolitana de la Ciudad de México. Para finalizar se proporcionan algunas conclusiones.

**Tabla 86.** Resultados de la granularidad y las funciones de pertenencia para la estimación de la concentración de ozono en México usando el AG multiobjetivo. Criterio de parada = 16000 evaluaciones. Mes Enero.

# ejec.	Gran.	Proporción de datos	Opt. Mask / $\Sigma$ etiquetas	Q	MSE <sub>entr</sub>	RMS <sub>prue</sub>	
A	1	(4,3,4,4,3,2)	HD:(0.25,0.25,0.25,0.25); HR:(0.38,0.21,0.41) TMP:(0.14, 0.6,0.13,0.13) DV:(0.25,0.25,0.25,0.25) VV:(0.49,0.14,0.37); O3:(0.84,0.16)	(8,13,17,18) 12 etiq.	0.7708	43.551	0.0172
		(3,3,2,2,2,2)	HD:(0.20,0.58,0.22); HR:(0.28,0.45,0.27) TMP:(0.64,0.36); DV:(0.50,0.50) VV:(0.35,0.65); O3:(0.84,0.16)	(12,18) 2 etiq.	0.756	106.2713	0.0361
	2	(2,4,5,3,3,3)	HD:(0.77,0.23); HR:(0.54,0.16,0.15,0.15) TMP:(0.48,0.12,0.13,0.13,0.14) DV:(0.13,0.73,0.14) VV:(0.46,0.49,0.05); O3:(0.82,0.13,0.05)	(12,13,18) 5 etiq.	0.7873	46.5797	0.0214
		(4,2,3,5,2,2)	HD:(0.20,0.22,0.22,0.36); HR:(0.36,0.64) TMP:(0.29,0.17,0.54) DV:(0.20,0.20,0.20,0.20,0.20) VV:(0.50,0.50); O3:(0.80,0.20)	(6,12,18) 2 etiq.	0.7359	111.5277	0.0373
	3	(6,6,2,4,7,2)	HD:(0.05,0.05,0.44,0.09,0.16,0.21) HR:(0.26,0.43,0.15,0.06,0.05,0.05); TMP:(0.48,0.52); DV:(0.25,0.25,0.25,0.25) VV:(0.21,0.09,0.24,0.08,0.08,0.09,0.21) O3:(0.42,0.58)	(3,12,15,16,18) 8 etiq.	0.5991	35.1356	0.0222
		(3,2,2,2,3,2)	HD:(0.33,0.33,0.34); HR:(0.50,0.50) TMP:(0.50,0.50); DV:(0.50,0.50) VV:(0.33,0.33,0.34) O3:(0.80,0.20)	(6,12,18) 2 etiq.	0.7359	111.5277	0.0373
	4	(3,4,2,2,2,2)	HD:(0.21,0.67,0.12); HR:(0.25,0.25,0.25,0.25) TMP:(0.86,0.14); DV:(0.79,0.21) VV:(0.16,0.84); O3:(0.84,0.16)	(3,4,12,15,18) 6 etiq.	0.7566	39.4247	0.0219
		(3,3,3,3,2,2)	HD:(0.21,0.67,0.12); HR:(0.39,0.30,0.31) TMP:(0.33,0.33,0.34) DV:(0.33,0.33,0.34) VV:(0.50,0.50) O3:(0.84,0.16)	(12,18) 2 etiq.	0.756	106.2713	0.0361
B	1	(4,3,4,4,3,2)	HD:(0.25,0.25,0.25,0.25); HR:(0.38,0.21,0.41) TMP:(0.14,0.60,0.13,0.13) DV:(0.25,0.25,0.25,0.25) VV:(0.49,0.14,0.37); O3:(0.84,0.16)	(8,13,17,18) 12 etiq.	0.7708	43.551	0.0172
		(4,5,5,2,3,4)	HD:(0.32,0.07,0.31,0.30); HR:(0.20,0.20,0.20,0.20,0.20) TMP:(0.20,0.20,0.20,0.20,0.20); DV:(0.66,0.34) VV:(0.32,0.31,0.37); O3:(0.25,0.25,0.25,0.25)	(12,13,17,18) 11 etiq.	0.5501	65.6164	0.0209
	3	(4,2,3,2,3,2)	HD:(0.25,0.25,0.25,0.25); HR:(0.50,0.50) TMP:(0.09,0.08,0.83); DV:(0.89,0.11) VV:(0.33,0.33,0.34); O3:(0.80,0.20)	(8,13,17,18) 11 etiq.	0.7615	40.1446	0.0174
		(4,4,2,3,3,2)	HD:(0.25,0.25,0.25,0.25) HR:(0.25,0.25,0.25,0.25); TMP:(0.25,0.75) DV:(0.40,0.30,0.30) VV:(0.33,0.33,0.34); O3:(0.84,0.16)	(13,16,17,18) 12 etiq.	0.7636	57.254	0.0164

De las *Tablas 83, 84, 85 y 86* se puede indicar que el AG multiobjetivo cumplió su tarea de encontrar un conjunto de soluciones no dominadas que satisfacen el propósito de minimizar los valores de  $\Sigma$ etiq y MSE<sub>entr</sub>. En la sección *A* de cada tabla se observan un conjunto de soluciones no dominadas por cada ejecución que tienen el menor error MSE<sub>entr</sub> y la menor sumatoria de etiquetas.

En cuanto a la precisión de los modelos (última columna de cada tabla), se observa que las soluciones que poseen la menor sumatoria de etiquetas no poseen un buen RMS<sub>prue</sub>. Las soluciones más simples sugeridas por el AG multiobjetivo no son

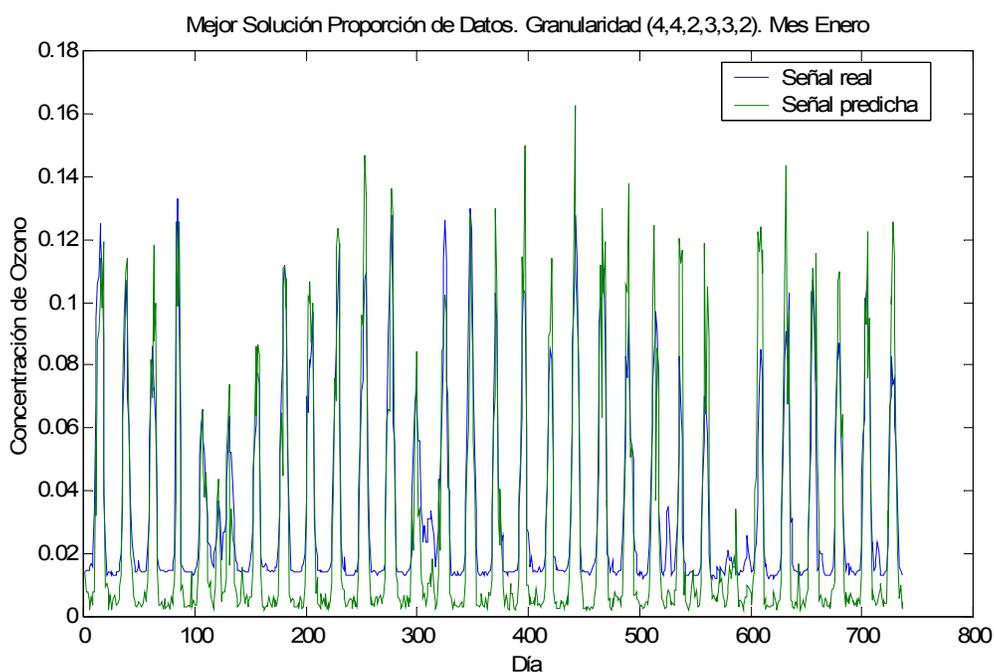
**Tabla 87.** Tiempo consumido, número de generaciones y número de soluciones no dominadas obtenidas por ejecución y criterio de parada del AG multiobjetivo. Caso Estudio: Ozono México, mes Enero.

<b>Núm. evaluaciones: 2000</b>				
<b>Tiempo Promedio</b>	<b># ejec.</b>	<b># Gen.</b>	<b>Tiempo (horas)</b>	<b># Soluciones no dominadas</b>
	1	30	11:18	24
10h:11min	2	30	12:01	25
	3	30	7:52	39
	4	30	9:31	23
<b>Núm. evaluaciones: 4000</b>				
<b>Tiempo Promedio</b>	<b># ejec.</b>	<b># Gen.</b>	<b>Tiempo (horas)</b>	<b># Soluciones no dominadas</b>
	1	61	10:40	28
18h:17min	2	62	19:16	13
	3	62	19:14	17
	4	61	23:57	16
<b>Núm. evaluaciones: 8000</b>				
<b>Tiempo Promedio</b>	<b># ejec.</b>	<b># Gen.</b>	<b>Tiempo (horas)</b>	<b># Soluciones no dominadas</b>
	1	124	25:14	18
29h:13min	2	125	28:41	17
	3	124	35:16	16
	4	124	27:39	11
<b>Núm. evaluaciones: 16000</b>				
<b>Tiempo Promedio</b>	<b># ejec.</b>	<b># Gen.</b>	<b>Tiempo (horas)</b>	<b># Soluciones no dominadas</b>
	1	250	58:35	13
50h:41min	2	250	51:52	16
	3	249	52:22	19
	4	250	39:56	22

capaces de predecir correctamente el comportamiento del ozono. Analicemos este hecho con más atención. En la *Tabla 83* vemos que las soluciones con menor  $\Sigma_{eti}$  en las 4 ejecuciones realizadas, tienen todas ellas un valor de  $\Sigma_{eti}$  de 2 y la máscara óptima que se encuentra con esta partición es (12,18), es decir que el ozono en el instante actual (posición 18 de la máscara) tiene una relación causal únicamente con la misma variable ozono una hora antes (posición 12 de la máscara). Es evidente que una máscara de complejidad 2 donde la única m-entrada corresponde a la variable de salida en el instante de tiempo anterior tiene una capacidad pésima de predicción, puesto que la señal de predicción será una constante. Recordemos que nuestro objetivo, en la aplicación que nos ocupa, es realizar una predicción a largo plazo. También aparece en las tablas, en varias ocasiones, la máscara (6,12,18), donde la posición 6 corresponde, una vez más, a un valor anterior de la variable ozono. Por lo tanto, nos encontramos en la misma situación que teníamos con la máscara (12,18). Por este motivo, estas máscaras deben ser rechazadas como posibles soluciones. Este

tipo de máscaras aparecen también en las *Tablas 84, 85 y 86*, pues lógicamente el **AGMO3** las encuentra como mejores soluciones desde la perspectiva de minimizar el objetivo  $\Sigma_{eti}$ . En este sentido, se podría pensar en incluir en el **AGMO3** una penalización a la obtención de este tipo de máscara. Esto se lograría de manera simple forzando que el mínimo valor de  $\Sigma_{eti}$  fuera mayor a la granularidad más pequeña de las variables que intervienen en el sistema.

Si analizamos los valores de  $RMS_{prue}$  en aquellas soluciones (sección *A* de las tablas) donde se está minimizando el objetivo  $MSE_{entr}$ , vemos que las soluciones con mayor error de prueba (0.0375 y 0.0222) son aquellas en las que la máscara derivada de dicha solución no incluye la variable hora del día (HD) entre sus relaciones causales con la salida. Esta es la causa fundamental del bajo rendimiento de estas soluciones en la predicción (*Tabla 84*). Las máscaras que no incluyen alguna de las posiciones 1, 7 o 13, correspondientes a diferentes instantes de tiempo de HD, son las que tienen  $RMS_{prue}$  más altos. Como ya se mencionó en el Capítulo VII, en esta aplicación la variable HD está realizando el papel de los precursores del ozono, por lo que es fundamental que esta variable forme parte del modelo.



**Figura 51.** Predicción del mejor modelo FIR para el mes Enero. Problema concentraciones de ozono en México. Granularidad (4,4,2,3,3,2) con  $Q=0.7636$  y distribución de landmarks: HD:(0.25,0.25,0.25,0.25)-HR:(0.25,0.25,0.25,0.25)-TMP:(0.25,0.75)-DV:(0.40,0.30,0.30)-VV:(0.33,0.33,0.34)- $o_3$ :(0.84,0.16). Máscara:(13,16,17,18).  $RMS_{prue} = 0.0164$ . Método **AGMO3**.

El mejor resultado logrado por **FIR EXPERTOS** (Véase *Tabla 59*) tiene un  $RMS_{prue}$  de 0.0176 ( $\Sigma etiq=10$ ) superior al error de 0.0164 ( $\Sigma etiq=12$ ) obtenido por el **AGMO3**. Asimismo, este método supera los resultados obtenidos por **FIR DEFECTO**, **AG1+EFP**, **AG1+AG2** y **AG3** de la *Tabla 80*.

El AG multiobjetivo ha permitido en una sola ejecución obtener un conjunto de soluciones (unas más simples y otras más precisas) en un tiempo razonable (*Tabla 87*) al alcanzar el criterio de parada estipulado. De este modo, no es necesario ejecutarlo tantas veces como los métodos **AG1**, **AG2** y **AG3** contemplados en esta memoria.

En resumen, el AG multiobjetivo propuesto en este capítulo ha permitido obtener modelos FIR para el mes de Enero con igual habilidad de predicción que los obtenidos por [NM06] y **FIR DEFECTO**. La combinación **EXPERTOS + AG2** sigue siendo el método más eficaz para obtener los mejores resultados de predicción del conjunto de datos de prueba, aunque le siguen muy de cerca **AGMO3**, **AG1+AG2** y **AG3**.

La *figura 51* muestra la señal real y de predicción usando el mejor modelo obtenido con el AG multiobjetivo. Analizando la figura vemos que los picos altos de la señal se alcanzan en la mayoría de las ocasiones, aunque hay picos que no se llegan a predecir correctamente.

En conclusión, se puede indicar que el AG multiobjetivo propuesto es una primera aproximación al desarrollo de este tipo de AGs para la determinación de los parámetros de fusificación de la metodología FIR. Consideramos que es necesario probarlo en un conjunto más amplio de aplicaciones para conocer bien su comportamiento. Así mismo, sería muy interesante desarrollar un AGMO de segunda generación para obtener mejores soluciones.

---

## **Conclusiones**



## CONCLUSIONES

Se dedica este último capítulo a apuntar las consideraciones que creemos más relevantes en forma de resumen. Además, se comenta algunos aspectos sobre trabajos futuros que siguen la línea de investigación aquí expuesta y sobre otras que se puedan derivar. Por último, se presentan todas las contribuciones científicas (en forma de publicaciones) originadas de nuestra investigación, así como los proyectos de investigación en donde se ha participado durante el desarrollo de esta tesis doctoral.

### Resultados obtenidos y conclusiones

Los resultados obtenidos por los SDEs desarrollados en esta tesis son muy buenos. Recordemos que en las cuatro aplicaciones estudiadas los tres AGs propuestos son los que obtienen mejores modelos con mayor eficiencia en el proceso de predicción. En general, el **AG3** y la combinación **AG1+AG2**, en ese orden, son los que han mostrado mejores resultados en todas las aplicaciones, seguidos por el **AG1+EFP**. Sin embargo el **AG3** es el que presenta mayor costo computacional. En la aplicación del sistema nervioso central humano, los tres SDEs presentados en esta tesis han conseguido resultados superiores a otras metodologías usadas en trabajos previos y al **FIR DEFECTO**, en ese orden. Así mismo, en la aplicación de la línea de media tensión en núcleos urbanos los tres AGs propuestos consiguen resultados de predicción notablemente superiores a los conseguidos por **FIR DEFECTO** y por otras metodologías en trabajos previos. También hacen un buen papel los SDEs diseñados en la aplicación de la estimación de las concentraciones de ozono en Austria. A pesar de la pobreza de los datos (tanto en cantidad como en calidad) el **AG3** y **AG1+AG2** son los que lideran la capacidad de predicción, excepto en uno de los casos del *cross validation* donde **FIR MANUAL** consigue el menor error de predicción. Sin embargo, una solución de búsqueda manual no es para nada deseable y no permite la utilización de la metodología FIR por usuarios inexpertos en lógica difusa que requieran modelar un sistema de complejidad elevada. Finalmente, en la aplicación de la estimación de las concentraciones de ozono en México, una vez más, se confirma que las diferentes combinaciones propuestas de los AGs diseñados en esta investigación superan ampliamente los resultados obtenidos por **FIR DEFECTO** y también, cosa no esperada, por **FIR EXPERTOS**, es decir, por FIR cuando los parámetros de fusificación han sido definidos por expertos en el área. Por lo tanto como conclusión general, debemos decir que los SDEs diseñados e implementados en esta tesis consiguen excelentes resultados para la tarea que les ha sido encomendada

en el entorno de la metodología FIR. Es pues el usuario quien debe decidir qué SDE resulta más conveniente para la aplicación que tiene entre manos, en función de las necesidades temporales y de precisión.

A continuación se presentan algunas conclusiones adicionales que queremos remarcar:

#### **A.- Influencia de la granularidad de las particiones difusas en el comportamiento de la metodología FIR**

A lo largo de esta memoria, se ha resaltado la importancia de la granularidad de las particiones difusas dentro del proceso de obtención del modelo FIR que represente mejor el sistema bajo estudio. Sobre este aspecto, se han extraído las siguientes conclusiones:

- En todos los sistemas difusos, el número de etiquetas (clases) por variable influye decisivamente en el comportamiento final del sistema difuso resultante. FIR no es la excepción, debido a que también está basada en lógica difusa y la determinación de los parámetros necesarios en el paso de fusificación (discretización) llega a ser relevante para la identificación de un buen modelo que capture el comportamiento del sistema de manera precisa.
- El nivel de granularidad óptimo para cada partición difusa depende del problema que se esté tratando, por lo tanto los buenos resultados obtenidos en un problema no son directamente extrapolables a otros problemas.
- Las diferencias de precisión entre los modelos FIR obtenidos con un buen nivel de granularidad en las particiones difusas y los obtenidos usando **FIR DEFECTO** (mismo número de etiquetas en todas las variables, siendo este valor por defecto de 3) son considerables. Por lo tanto, la selección de un nivel de granularidad adecuado es una tarea fundamental para el proceso de fusificación (discretización) de la metodología FIR.
- El método de búsqueda del nivel de granularidad óptimo propuesto en el Capítulo III (**AG1**) requiere poco esfuerzo computacional si se tiene en cuenta la dimensión del espacio de búsqueda manejado, y ha demostrado una buena eficiencia, especialmente en los problemas con un número elevado de variables.

#### **B.- Aprendizaje de las funciones de pertenencia para razonamiento inductivo difuso**

El AG propuesto para el aprendizaje de las funciones de pertenencia ha tenido

un buen comportamiento. El método propuesto en el Capítulo IV (**AG2**) ha logrado muy buenos resultados en la mayoría de los problemas estudiados en comparación al uso de la distribución uniforme de las funciones de pertenencia a lo largo del universo de discurso de las variables (Método EFP). El mecanismo de ajuste “local” ha permitido refinar las definiciones de las funciones de pertenencia manteniéndose inalterable la granularidad (número de clases) de las variables. La granularidad ha sido proporcionada en algunos casos por expertos en el área de interés o sugerida por el método presentado en el Capítulo III (**AG1**). Aunque se ha indicado que la adaptación no lineal de contextos en sistemas difusos resulta costosa computacionalmente, ésta nos ha permitido representar mejor los conceptos en entornos reales con alta sensibilidad en las clases extremas o en las clases intermedias, por lo tanto estas variaciones han influido directamente en el comportamiento de los modelos FIR resultantes. Es así como, una selección adecuada de las funciones de pertenencia es, también, una tarea de vital importancia para el proceso de fusificación (discretización) de la metodología FIR.

### **C.- Aprendizaje conjunto de la granularidad y las funciones de pertenencia para razonamiento inductivo difuso**

El método propuesto en el Capítulo V (**AG3**) ha demostrado ser eficaz en problemas complejos. La filosofía de ajuste “global” que subyace en dicho método ha proporcionado unos buenos resultados y ha permitido obtener sistemas difusos con buena capacidad de generalización. Sin embargo, el proceso de aprendizaje ha sido lento y dificultoso porque se trabaja con un espacio de búsqueda extenso. Asimismo, el tamaño de la población se incrementa al tener que considerar más diversidad para la población inicial.

Los valores de los parámetros escogidos para los SDEs propuestos (**AG1**, **AG2** y **AG3**) han originado un buen comportamiento general para todos los problemas tratados en esta investigación, pero estos métodos pueden ofrecer mejores resultados si se ajustan dichos valores realizando una experimentación específica.

Por último, es necesario señalar que no se puede concluir que un SDE específico o combinación de éstos es mejor que los demás, ya que ninguno tuvo mejor rendimiento (encontró soluciones más precisas y fiables) que otro en todos los casos estudiados. La excepción es quizá el **AG1**, pues ha sido el peor de los SDEs desarrollados en la mayoría de las situaciones. Sin embargo mejora substancialmente el uso de parámetros por defecto en FIR. Por todo ello, es importante tomar en consideración las ventajas e inconvenientes de cada método propuesto y las características del problema de optimización a tratar para determinar qué SDE usar en cada caso.

#### **D.- Funciones objetivo implementadas**

En esta memoria se han propuesto para la evaluación de los cromosomas en todos los SDEs desarrollados (**AG1**, **AG2** y **AG3**) dos funciones objetivos: a) la calidad de la máscara óptima y b) el error de predicción MSE de parte del conjunto de datos de entrenamiento. La calidad de una máscara  $Q$ , es un valor entre 0 y 1, donde 1 indica la más alta calidad. Por lo tanto, la primera función objetivo propuesta es  $1-Q$ , ya que la tarea era minimizar ésta. La segunda función objetivo corresponde a la predicción del error de una porción del conjunto de datos de entrenamiento. El error cuadrático medio normalizado en porcentaje (MSE), se ha usado para este fin. Es por ello, que esta función objetivo se denomina  $MSE_{entr}$ .

Se ha confirmado que el tiempo de CPU necesario por el **AG2** para efectuar el conjunto de 30 ejecuciones está relacionado directamente con el número predefinido de clases por variable. Por otro lado, en general se observa que la función objetivo  $1-Q$  necesita menos tiempo para ser evaluada pero tiene menor eficiencia en la predicción del conjunto de datos de prueba. Contrariamente, la función objetivo  $MSE_{entr}$  es más costosa desde la perspectiva del tiempo de CPU pero la eficiencia es más alta. Sin embargo, el desempeño de ambas funciones objetivo puede verse influenciado por las características, cantidad y complejidad de los datos disponibles del problema en estudio, por ser FIR una metodología de minería de datos. De este modo, se recomienda al usuario escoger la función objetivo más conveniente en función del tamaño del problema de optimización y las necesidades del propio usuario. Si se requiere la máxima precisión en el problema que se está estudiando será conveniente usar el **AG3** con la función objetivo  $MSE_{entr}$ . Contrariamente, si la precisión no es el parámetro primordial, y se requiere no excederse en el tiempo de cómputo, el **AG1** y/o el **AG2** con la función objetivo  $1-Q$  serán mejores alternativas.

#### **E.- Método multiobjetivo de aprendizaje para razonamiento inductivo difuso**

El método multiobjetivo propuesto en el Capítulo VIII (**AGMO3**) ha demostrado también ser eficaz en problemas complejos. A pesar de estar aún en proceso de ajuste y pruebas, ha permitido con los parámetros que se le han establecido obtener buenos resultados para la aplicación estudiada. La filosofía de ajuste “global” que subyace en dicho método ha permitido obtener también sistemas difusos con buena capacidad de generalización (aunque con errores  $MSE_{entr}$  más altos que los otros SDEs contemplados en esta memoria). Sin embargo, el tamaño de la población se incrementa al tener que considerar más diversidad para la población inicial, lo que puede derivar (dependiendo de la aplicación o problema que se trate) en un aumento del número de evaluaciones de cromosomas (y por ende, del número de generaciones) para que haya un proceso evolutivo más largo y se puedan explorar otras zonas prometedoras y explotar más el espacio de búsqueda. Hay que recordar

que el **AGMO3** no tiene un mecanismo elitista, por lo cual el conjunto de soluciones no dominadas en la última generación son las únicas que se tendrán a la mano para realizar la selección de aquellas que cumplan con nuestras necesidades.

El **AGMO3** ha permitido en una sola ejecución (con un tiempo similar a los que tardan los otros AGs implementados) obtener un conjunto de soluciones (unas más simples y otras más precisas para poder escoger) al alcanzar el criterio de parada que se le ha estipulado. De este modo, no fue necesario ejecutarlo tantas veces como los métodos **AG1**, **AG2** y **AG3** contemplados en esta memoria.

Se observa que las soluciones que tienen el menor valor de sumatoria de etiquetas o clases por variable (considerando las relaciones causales de la máscara óptima), es decir, los modelos más simples son los más propensos a poseer un error de predicción elevado con respecto al conjunto de datos de prueba. Se debe estudiar la posibilidad de colocar una restricción para ésta parte de la función objetivo del **AGMO3**, de manera tal que descarte soluciones que tengan sólo relaciones causales en instantes de tiempo previos en la misma variable de salida.

## **Trabajos futuros**

A continuación, se comentan algunas líneas de trabajo que quedan aún abiertas sobre la temática tratada en esta memoria:

### **Sobre la medida de calidad usada por la metodología FIR:**

En virtud de los resultados obtenidos se han comenzado a realizar esfuerzos para establecer otras medidas de calidad  $Q$  de la máscara en el proceso de modelado cualitativo de la metodología FIR. En las aplicaciones presentadas en esta memoria se ha hecho patente que esta medida no refleja con la precisión suficiente la capacidad de predicción de las máscaras.

### **Sobre otros posibles casos de estudios o aplicaciones de especial interés:**

En estos momentos, hay disponibles unos datos simulados sobre una aplicación de control en sistemas industriales que corresponden a una turbina de gas. Se requiere obtener modelos para distintas situaciones de funcionamiento de esta planta con el fin de realizar posteriormente detección de fallos. Se pretende aplicar los métodos propuestos en esta memoria para resolver la problemática planteada.

### **Sobre posibles extensiones de los métodos propuestos:**

Por un lado, se puede realizar un estudio comparativo para optimizar los parámetros de los SDEs propuestos. Mientras que para parámetros como la probabilidad de cruce o de mutación se pueden seguir recomendaciones de diversos autores duchos en el tema [Gre86], el tamaño de la población es un parámetro interesante a estudiar puesto que depende del problema que se esté tratando, de la extensión del espacio de búsqueda, de si se utilizan o no nichos, etc. De todas formas, no se debe perder de vista que los AGs deben terminar en un tiempo razonable, máxime cuando la función objetivo es costosa en tiempo de ejecución. En este sentido, el tamaño de la población es un factor que no debe tenerse en cuenta en forma aislada, sino junto a la condición de parada del AG (sea por número de generaciones o por número de evaluaciones de cromosomas). Además, puede resultar que los parámetros óptimos no sean los mismos en las dos funciones objetivo que se han implementado.

Por otro lado, se pueden estudiar como actúan los SDEs propuestos con otros operadores genéticos de cruce o mutación con codificación real [HLV98] o con nuevos mecanismos de selección frente a la resolución de los problemas estudiados en esta memoria, así como para analizar la diversidad de la población.

Por su parte, el AG multiobjetivo presentado en el Capítulo VIII debe probarse sobre un conjunto amplio de aplicaciones para estudiar cabalmente su rendimiento. Asimismo, sería interesante desarrollar otro AG multiobjetivo en donde se aprenda solamente la granularidad (como el **AG1**) y se persigan los mismos objetivos descritos.

Por último, es posible desarrollar un AG multiobjetivo elitista que sea más eficaz que el presentado en esta memoria como aportación adicional, ya que estos encuentran mejores conjuntos de soluciones (más cercanos al Pareto).

## **Contribuciones científicas**

Hasta estos momentos esta investigación ha proporcionado varias contribuciones de interés para la comunidad científica nacional e internacional. A continuación se presentan las publicaciones que se han derivado:

### **Reportes de investigación:**

- J. Acosta, A. Nebot and J. M. Fuertes, “Método Multiobjetivo de Aprendizaje para Razonamiento Inductivo Difuso,” Report de Recerca LSI-

06-35-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, Diciembre 2006.

- J. Acosta, A. Nebot and J. M. Fuertes, “Modelado de las concentraciones locales de ozono en la Zona Centro del Area Metropolitana de la Ciudad de México,” Report de Recerca LSI-06-34-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, Noviembre 2006.
- J. Acosta, A. Nebot and J. M. Fuertes, “Predicción de concentraciones locales de ozono usando una metodología híbrida de Soft Computing,” Report de Recerca ESII-RR-06-06, Dept. d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain, Marzo 2006.
- J. Acosta, A. Nebot and J. M. Fuertes, “Modelado del Sistema Nervioso Central Humano usando una metodología híbrida de Soft Computing,” Report de Recerca LSI-05-57-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, Noviembre 2005.
- J. Acosta, A. Nebot and J. M. Fuertes, “Aprendizaje Global de los parámetros de discretización del Razonamiento Inductivo Difuso,” Report de Recerca ESII-RR-05-10, Dept. d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain, Octubre 2005.
- J. Acosta and A. Nebot, “Uso del razonamiento inductivo difuso para solucionar un problema de distribución de energía eléctrica,” Report de Recerca LSI-05-22-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, Mayo 2005.
- J. Acosta, A. Nebot and J. M. Fuertes, “Un método de aprendizaje de las funciones de pertenencia para razonamiento inductivo,” Report de Recerca ESII-RR-04-03, Dept. d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain, Junio 2004.

#### **Artículos en congresos nacionales e internacionales:**

- J. Acosta, A. Nebot, P. Villar and J.M. Fuertes, “Estimating the maintenance cost in electric networks using an Hybrid Soft Computing Methodology,” in *Proceedings 2<sup>nd</sup> European Modeling and Simulation Symposium (EMSS 2006)*, Barcelona, Spain, October 4-6, 2006, pp. 349-358.
- J. Acosta, A. Nebot and J.M. Fuertes, “Hybrid Evolutionary Data Analysis Technique for Environmental Modeling,” in *Actas de las XI Jornadas de*

*Ingeniería del Software y Bases de Datos (JISBD 2006)*, Sitges, Spain, October 3-6, 2006, pp. 251-260.

- J. Acosta, A. Nebot and J.M. Fuertes, “Modelado del Sistema Nervioso Central Humano usando metodologías híbridas de Soft Computing,” in *Actas de las II Jornadas de Investigación en Automática, Visión y Robótica (AVR'06)*, Barcelona, Spain, July 4-6, 2006, pp. 54-61.
- J. Acosta, A. Nebot and P. Villar, “Dealing with Electrical Distribution Problems using the Fuzzy Inductive Reasoning Methodology,” in *Proceedings of the Ninth IASTED International Conference Artificial Intelligence and Soft Computing (ASC 2005)*, Benidorm, Spain, September 12-14, 2005, pp. 268-274.
- J. Acosta, A. Nebot and P. Villar, “Optimization of fuzzy membership functions by genetics algorithms,” in *Proceedings of 5th Internacional Conference on Control, Virtual Instrumentation, and Digital Systems (CICINDI'04)*, Ciudad de México, México, September 7-10, 2004, 10 pages, ISBN: 970-36-0189-8.
- J. Acosta, A. Nebot, P. Villar and J.M. Fuertes, “Algoritmos genéticos para el aprendizaje de particiones difusas en razonamiento inductivo,” in *Actas del Tercer Congreso Espanol de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB'04)*, Córdoba, Spain, February 4-6, 2004, pp. 110-117.
- J. Acosta, A. Nebot, P. Villar and J.M. Fuertes, “Ajuste automático de particiones difusas para razonamiento inductivo usando algoritmos genéticos,” in *Actas de las I Jornadas de Investigación en Automática, Visión y Robótica (AVR'04)*, Barcelona, Spain, February 03-04, 2004, 7 pages, ISBN: 84-7653-845-6.
- J. Acosta, A. Nebot, P. Villar and J.M. Fuertes, “Learning of Fuzzy Partitions in Human Central Nervous System Modeling using Genetic Algorithms,” in *Proceedings of 2004 Wester Simulation MultiConference (WMC'04)*, San Diego, CA, USA, January 18-22, 2004, pp. 90-95.

#### **Artículos en revistas internacionales:**

- J. Acosta, A. Nebot, P. Villar and J.M. Fuertes, “Optimization of fuzzy partitions for inductive reasoning using genetic algorithms,” *International Journal of Systems Science*, en proceso de revisión.

Se está actualmente trabajando en la elaboración de los siguientes artículos:

- Un artículo para la revista *Environmental Modelling* en donde se expongan los resultados obtenidos por los métodos propuestos en esta memoria sobre el caso de estudio de predicciones de concentraciones de ozono en México.
- Un artículo para la revista *International Journal of General Systems* en donde se presente el método de aprendizaje conjunto de la granularidad y funciones de pertenencia así como los resultados de su aplicación a varios casos de estudios considerados en esta memoria.

Se prevé que estas publicaciones serán enviadas para su consideración a las revistas mencionadas antes de la lectura de esta tesis doctoral.

## **Participación en proyectos de investigación**

Esta tesis fue desarrollada dentro del marco de los siguientes proyectos de investigación:

- Investigador Principal: Dr. Alfredo Vellido. Título: "Diseño de un sistema de apoyo a la decisión en oncología clínica basado en métodos avanzados de *soft computing* y visualización," Consejo Interministerial de Ciencia y Tecnología (CICYT TIN2006-08114). Entidad Participante: Universitat Politècnica de Catalunya, Barcelona, Spain. Duración, desde: 2006 hasta: 2009.
- Investigador Principal: Dra. Angela Nebot. Título: "Metodologías híbridas de soft computing y su aplicación a la detección de anomalías en estaciones depuradoras de aguas residuales," Consejo Interministerial de Ciencia y Tecnología (CICYT DPI2002-03235). Entidad Participante: Universitat Politècnica de Catalunya, Barcelona, Spain. Duración, desde: 2002 hasta: 2005.
- Investigador Principal: Dra. Angela Nebot. Título: "Metodologías híbridas para el análisis de fallos en sistemas dinámicos complejos," Consejo Interministerial de Ciencia y Tecnología (CICYT TAP-990747). Entidad Participante: Universitat Politècnica de Catalunya, Barcelona, Spain. Duración, desde: 1999 hasta: 2002.

## Referencias Bibliográficas

- [AA02] S.A. Abdul-Wahab and S.M. Al-Alawi, "Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks," *Environmental Modelling and Software*, vol. 17, no. 3, pp. 219-228, 2002.
- [Aar89] E. H. Aarts, *Simulated Annealing and Boltzman Machines: A stochastic approach to Combinatorial Optimization and Neural Computing*. Wiley, 1989.
- [ACC+00] R. Alcalá, J. Casillas, O. Cordón, F. Herrera, and I. Zwir, "Techniques for learning and tuning fuzzy rule-based systems for linguistic modeling and their applications," in *Knowledge Engineering Systems: Techniques and Applications*, C. T. Leondes, Ed. Academic Press, 2000, vol. 3, pp. 889-941.
- [Aco00] J. Acosta, "Aplicación de las redes neuronales para la identificación de un sistema de reacción continuo," Master's thesis, Dirección de Investigación y Postgrado, Universidad Nacional Experimental Politécnica "Antonio José de Sucre" (UNEXPO), Barquisimeto, Venezuela, 2000.
- [ACP06] G. Ascia, V. Catania, and D. Panno, "An Integrated Fuzzy-GA Approach for Buffer Management," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 4, pp. 528-541, Aug. 2006.
- [ACT99] E. Alba, C. Cotta, and J. M. Troya, "Evolutionary design of fuzzy logic controllers using strongly-typed GP," *Mathware Soft Comput.*, vol. 6, no. 1, pp. 109-124, 1999.
- [ADB02] G. Alpaydin, G. Dündar, and S. Balktr, "Evolution-based design of neural fuzzy networks using self-adapting genetic parameters," *IEEE Trans. Fuzzy Systems*, vol. 10, no. 2, pp. 211-221, Apr. 2002.
- [AJP96] G. Acuña, H. Jorquera, and R. Pérez, "Neural Network Model for Maximum Ozone Concentration Prediction," *Lectures Notes in Computer Science*, vol. 1112, no. 1, pp. 263-268, 1996.
- [AL97] E. Aarts and J.K. Lenstra, *Local Search in Combinatorial Optimization*. John Wiley & Sons, 1997.
- [AN05] J. Acosta and A. Nebot, "Uso del razonamiento inductivo difuso para solucionar un problema de distribución de energía eléctrica," Report de Recerca LSI-05-22-R, Dept. Llenguatges i Sistemes Informàtics,

- Universitat Politècnica de Catalunya, Barcelona, Spain, 2005.
- [ANF04] J. Acosta, A. Nebot and J. M. Fuertes, “Un método de aprendizaje de las funciones de pertenencia para razonamiento inductivo,” Report de Recerca ESII-RR-04-03, Dept. d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain, 2004.
- [ANF05a] J. Acosta, A. Nebot and J. M. Fuertes, “Modelado del Sistema Nervioso Central Humano usando una metodología híbrida de Soft Computing,” Report de Recerca LSI-05-57-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 2005.
- [ANF05b] J. Acosta, A. Nebot and J. M. Fuertes, “Aprendizaje Global de los parámetros de discretización del Razonamiento Inductivo Difuso,” Report de Recerca ESII-RR-05-10, Dept. d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain, 2005.
- [ANF06a] J. Acosta, A. Nebot and J. M. Fuertes, “Predicción de concentraciones locales de ozono usando una metodología híbrida de Soft Computing,” Report de Recerca ESII-RR-06-06, Dept. d’Enginyeria de Sistemes, Automàtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain, 2006.
- [ANF06b] J. Acosta, A. Nebot and J. M. Fuertes, “Modelado de las concentraciones locales de ozono en la Zona Centro del Area Metropolitana de la Ciudad de México,” Report de Recerca LSI-06-34-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 2006.
- [ANF06c] J. Acosta, A. Nebot and J. M. Fuertes, “Método Multiobjetivo de Aprendizaje para Razonamiento Inductivo Difuso,” Report de Recerca LSI-06-35-R, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 2006.
- [AW95] K.J. Anstrom and B. Wittenmark, *Adaptive Control*. Adison Wesley, 1995.
- [Bäc96] T. Bäck, *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, 1996.
- [Bak85] J. E. Baker, “Adaptative selection methods for genetic algorithms,” in *Proceedings of the First International Conference on Genetics Algorithms and Their Applications*, J. J. Grefenstette, Ed. Lawrence Erlbaum Associates, 1985, pp. 101-111.

- [Bak87] J. E. Baker, "Reducing bias and inefficiency in the selection algorithm," in *Proc. 2nd Int. Conf. Genetic Algorithms (ICGA'87)*, Hillsdale, 1987, pp. 14-21.
- [BB91] R. K. Belew and L. B. Booker, *Proceedings of the Fourth International Conference on Genetic Algorithms*. Morgan Kaufmann Publishers, 1991.
- [BBC+95] P. P. Bonissone, V. Badami, K. H. Chiang, P. S. Khedkar, K. Marcelle, and M. J. Schutten, "Industrial applications of fuzzy logic at General Electric," *Proc. IEEE*, vol. 83, pp. 450-465, Mar. 1995.
- [BBH95] K.M. Bossley, M. Brown, and C.J. Harris, "Parsimonious Neurofuzzy Modelling," *Research Journal*, Research Group on Image, Speech and Intelligent Systems, University of Southampton, 1995.
- [BBTZ01] S. Bleuler, M. Braek, L. Thiele, and E. Zitzler, "Multiobjective Genetic Programming: Reducing Bloat Using SPEA2," in *Proc. IEEE 2001 Congress on Evolutionary Computation (CEC'2001)- Vol. 1*, New Jersey, 2001, pp. 536-543.
- [BD01] H.-G. Beyer and K. Deb, "On Self-Adaptive Features in Real Parameter Evolutionary Algorithms," *IEEE Trans. Evol. Comput.*, vol. 5, no. 3, pp. 250-270, 2001.
- [BD95] A. Bardossy and L. Duckstein, *Fuzzy Rule-Based Modeling with Application to Geophysical, Biological and Engineering Systems*. CRC Press, 1995.
- [BEF84] J.C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-Means clustering algorithm", *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191-203, 1984.
- [BH91] T. Bäck and F. Hoffmeister, "Extended selection mechanisms in genetic algorithms," in *Proceedings of the Fourth International Conference on Genetic Algorithms*, R.K. Belew and L.B. Booker, Eds. Morgan Kaufmann Publishers, 1991, pp. 92-99.
- [BHS91] T. Bäck, F. Hoffmeister, and H.-P. Schwefel, "A survey of evolution strategies," in *Proceedings of the Fourth International Conference on Genetic Algorithms*, R.K. Belew and L.B. Booker, Eds. Morgan Kaufmann Publishers, 1991, pp. 2-9.
- [BKC96] P. P. Bonissone, P. S. Khedar, and Y.-T. Chen, "Genetic algorithms for automated tuning of fuzzy controllers, a transportation application," in *Proc. 5th IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE'96)*, New Orleans, LA, 1996, pp. 674-680.

- [Bon97] P.P. Bonissone, "Soft computing: the convergence of emerging reasoning technologies," *Soft Comput.*, vol. 1, no. 1, pp. 6-18, Apr. 1997.
- [Boo82] L. B. Booker, "Intelligent Behaviour as an Adaption to the Task Environment," Ph.D. thesis, University of Michigan, 1982.
- [BS93] T. Bäck and H.P. Schwefel, "An Overview of Evolutionary Algorithms for Parameter Optimization," *Evol. Comput.*, vol. 1, no. 1, pp. 1-23, 1993.
- [BS95] T. Bäck and H. P. Schwefel, "Evolution strategies I: variants and their computational implementation," in *Genetic Algorithms in Engineering and Computer Science*, J. Periaux., G. Winter, M. Galán, and P. Cuesta, Eds. John Wiley and Sons, 1995, pp. 111-126.
- [BS96] T. Bäck and H.P. Schwefel, "Evolutionary Computation: An Overview," in *Proc. 3rd IEEE Conf. Evolutionary Computation*, Piscataway, NJ, 1996, pp. 20-29.
- [BSEK06] P.P. Bonissone, R. Subbu, N. Eklund, and T.R. Kiehl, "Evolutionary Algorithms + Domain Knowledge = Real-World Evolutionary Computation", *IEEE Trans. Evol Comput.*, vol. 10, no. 3, pp. 256-280, June 2006.
- [BW01] A.L. Blumel and B.A. White, "Multiobjective Optimization of Fuzzy Logic Scheduled Controllers for Missile Autopilot Design," in *IEEE Joint 9th IFSA World Congress and 20th NAFIPS Int. Conf.- Vol. 3*, 2001, pp. 1758-1763.
- [CAN97] J. Cueva, R. Alquézar, and A. Nebot, "Experimental comparison of fuzzy and neural network techniques in learning models of the central nervous system control," in *Proc. 5th Eur. Congr. on Intelligent Techniques and Soft Computing (EUFIT'97)*, Aachen, 1997, pp. 1014-1018.
- [Cas95] J. L. Castro, "Fuzzy logic controllers are universal approximators," *IEEE Trans. Syst., Man, and Cybern.*, vol. 25, no. 4, pp. 629-635, 1995.
- [CBM95] U. Cortés, J. Béjar, and A. Moreno, *Inteligencia Artificial*. Barcelona: Ediciones UPC, 1995.
- [CCA03] H.-S. Chung, S. Choi, and J.J. Alonso, "Supersonic Business Jet Design Using Knowledge-Based Genetic Algorithm with Adaptive Unstructured Grid Methodology," presented at the 21st AIAA Applied Aerodynamics Conference, Orlando, FL, June 23-26, 2003, Paper AIAA-2003-3791.

- [CCdJH05] J. Casillas, O. Cordón, M.J. del Jesus, and F. Herrera, "Genetic Tuning of Fuzzy Rule Deep Structures Preserving Interpretability and its Interaction with Fuzzy Set Reduction," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 1, pp. 13-29, Feb. 2005.
- [CCG99] R. Corona, A.M. Chávez, and R. Gutiérrez, "Dinámica migratoria de la Ciudad de México. Serie: Tu ciudad así funciona," México: Gobierno del Distrito Federal, 1999.
- [CD96] J. L. Castro and M. Delgado, "Fuzzy systems with defuzzification are universal approximators," *IEEE Trans. Syst., Man, and Cybern.*, vol. 26, no. 1, pp. 149-152, 1996.
- [Cel91] F. Cellier, *Continuous System Modeling*. New York: Springer-Verlag, 1991.
- [CGH+04] O. Cordón, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, "Ten years of genetic fuzzy systems: current framework and new trends," *Fuzzy Sets Syst.*, vol. 141, no. 1, pp. 5-31, 2004.
- [CH95] O. Cordón and F. Herrera, "A general study on genetic fuzzy systems," in *Genetic Algorithms in Engineering and Computer Science*, J. Periaux, G. Winter, M. Galán, and P. Cuesta, Eds. John Wiley and Sons, 1995, pp. 33-57.
- [CH97] O. Cordón and F. Herrera, "A three-stage evolutionary process for learning descriptive and approximative fuzzy logic controller knowledge bases from examples," *Int. J. Approx. Reason.*, vol. 17, no. 4, pp. 369-407, 1997.
- [CHB00] C.A. Coello, A. Hernández, and B.P. Buckles, "Evolutionary Multiobjective Design of Combinational Logic Circuits," in *Proceedings of the Second NASA/DoD Workshop on Evolvable Hardware*, J. Lohn, A. Stoica, D. Keymeulen, and S. Colombano, Eds. IEEE Computer Society, 2000, pp. 161-170.
- [CHHM01] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena, *Genetic Fuzzy Systems. Evolutionary tuning and learning of fuzzy knowledge bases*. Singapore: World Scientific, 2001.
- [CHL01] S.-M. Chen, Y.-J. Horng, and C.-H. Lee, "Document retrieval using fuzzy-valued concept networks," *IEEE Trans. Syst., Man, Cybern. B*, vol. 31, no. 1, pp. 111-118, Feb. 2001.
- [CHL96] O. Cordón, F. Herrera, and M. Lozano, "On the bidirectional integration of genetic algorithms and fuzzy logic," in *Proc. Second Online Workshop on Evolutionary Computation (WEC2)*, Nagoya, 1996, pp. 13-16.

- [CHL97a] O. Cordón, F. Herrera, and M. Lozano, "A classified review on the combination fuzzy logic-genetic algorithms bibliography: 1989-1995," in *Genetic Algorithms and Fuzzy Logic Systems. Soft Computing Perspectives*, E. Sanchez, T. Shibata, and L. Zadeh, Eds. World Scientific, 1996.
- [CHL97b] O. Cordón, F. Herrera, and M. Lozano, "On the combination of fuzzy logic and evolutionary computation: a short review and bibliography," in *Fuzzy Evolutionary Computation*, W. Pedrycz, Ed. Kluwer Academic Press, 1997, pp. 55-77.
- [CHP97] O. Cordón, F. Herrera, and A. Peregrín, "Applicability of the fuzzy operators in the design of fuzzy logic controllers," *Fuzzy Sets Syst.*, vol. 86, pp. 15-41, 1997.
- [CHS98] O. Cordón, F. Herrera, and L. Sánchez, "Computing the spanish medium electrical line maintenance costs by means of evolution-based learning processes," in *Proc. 11<sup>th</sup> International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA-98-AIE)*, Castellón, 1998, vol. 1, pp. 478-482.
- [CHS99] O. Cordón, F. Herrera, and L. Sánchez, "Solving Electrical Distribution Problems using Hybrid Evolutionary Data Analysis Techiques," *Applied Intelligence*, vol. 10, pp. 5-24, 1999.
- [CHV00] O. Cordón, F. Herrera, and P. Villar, "Analysis and guidelines to obtain a good uniform fuzzy partition granularity for fuzzy rule-based systems using simulated annealing," *Int. J. Approx. Reason.*, vol. 25, pp. 187-215, 2000.
- [CJ01] J. Chenevez and Ch. Jensen, "Operational ozone forecast for the region of Copenhagen by the Danish Meteorological Institute", *Atmospheric Environment*, vol. 35, pp. 4567-4580, 2001
- [CJKO01] D.W. Corne, N.R. Jerram, J.D. Knowles and M.J. Oates, "PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, L. Spector, E. Goodman, A. Wu, W.B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M.H. Garzon and E. Burke, Eds. Morgan Kaufmann Publishers, 2001, pp. 283-290.
- [CKO00] D.W. Corne, J.D. Knowles and M.J. Oates, "The Pareto Envelope-based Selection Algorithm for Multiobjective Optimization," in *Proceedings of the Parallel Problem Solving from Nature VI Conference*, M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J.J. Merelo and H.-P. Schwefel, Eds. Springer, 2000, pp. 839-848.

- [CL05] Y.-C. Chiou and L. W. Lan, "Genetic fuzzy logic controller: an iterative evolution algorithm with new encoding method," *Fuzzy Sets Syst.*, vol. 152, no. 3, pp. 617-635, June 2005.
- [CLH02] B.-C. Chien, J.Y. Lin, and T.-P. Hong, "Learning discriminant functions with fuzzy attributes for classification using genetic programming," *Expert Syst. Appl.*, vol. 23, no. 1 pp. 31-37, 2002.
- [CLL00] I-F. Chung, C.J. Lin, and C.T. Lin, "A GA-based fuzzy adaptive learning control network," *Fuzzy Sets Syst.*, vol. 112, no. 1, pp. 65-84, May 2000.
- [CNMda92] F. Cellier, A. Nebot, F. Mugica, and A. de Albornoz, "Combined qualitative/quantitative simulation models of continuous-time processes using fuzzy inductive reasoning techniques," in *Proc. IFAC Symp. Intelligent Components and Instruments for Control Applications (SICICA '92)*, Málaga, 1992, pp. 589-593, invited paper.
- [CNMda96] F. Cellier, A. Nebot, F. Mugica, and A. de Albornoz, "Combined qualitative/quantitative simulation models of continuous-time processes using fuzzy inductive reasoning techniques," *Int. J. General Systems*, vol. 24, no. 1-2, pp. 95-116, 1996.
- [Coe00] C.A. Coello, "Constraint-handling using an evolutionary multiobjective optimization technique," *Civil Engineering and Environmental Systems*, vol. 17, pp. 319-346, 2000.
- [Coe03] C.A. Coello, "Evolutionary Multiobjective Optimization: Current and Future Challenges," in *Advances in Soft Computing: Engineering, Design and Manufacturing*, J. Benitez, O. Cordón, F. Hoffmann and R. Roy, Eds. Springer-Verlag, 2003, pp. 243-256.
- [Coe99] C.A. Coello, "A Comprehensive Survey of Evolutionary-Based Multiobjective Optimization Techniques," *Knowledge and Information Systems. An International Journal*, vol. 1, no. 3, pp. 269-308, 1999.
- [Com97] A.C. Comrie, "Comparing Neural Networks and Regression Models for Ozone Forecasting," *Air & Waste Management*, vol. 47, pp. 653-663, 1997.
- [Cor97] O. Cordón, "Una metodología para el diseño automático de Sistemas Basados en Reglas Difusas mediante Algoritmos Evolutivos," Ph.D. thesis, Dpto. Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, Granada, España, 1997.
- [CPC04] H.A. Camargo, M.G. Pires, and P.A.D. Castro, "Genetic Design of Fuzzy Knowledge Bases – a study of different approaches," in *Proc. IEEE Annual Meeting of the North American Fuzzy Information*

- Processing Society (NAFIPS'04)*, vol. 2, pp. 954–959, June 2004.
- [CRFC+03] P.B. Cheung, L.F.R. Reis, K.T.M. Formiga, F.H. Chaudhry and W. G.C. Ticona, “Multiobjective Evolutionary Algorithms Applied to the Rehabilitation of a Water Distribution System: A Comparative Study,” in *Second International Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, C.M. Fonseca, P.J. Fleming, E. Zitzler, K. Deb and L. Thiele, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 2632, 2003, pp. 662-676.
- [CT01a] C.A. Coello and G. Toscano, “Multiobjective Optimization using a Micro-Genetic Algorithm,” in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, L. Spector, E. Goodman, A. Wu, W.B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M.H. Garzon and E. Burke, Eds. Morgan Kaufmann Publishers, 2001, pp. 274-282.
- [CT01b] C.A. Coello and G. Toscano, “A Micro-Genetic Algorithm for Multiobjective Optimization,” in *First International Conference on Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C.A. Coello and D. Corne, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001, pp. 126–140.
- [CVL02] C.A. Coello, D.V. Veldhuizen, and G.B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems*. New York: Kluwer Academic/Plenum Publishers, 2002.
- [CYP96] Z Chi, H. Yan, and T. Pham, *Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition*. World Scientific, 1996.
- [DAPM00] K. Deb, S. Agrawal, A. Pratab and T. Meyarivan, “A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II,” KanGAL report 200001, Indian Institute of Technology, Kanpur, India, 2000.
- [Dar59] C. R. Darwin, *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray, 1859.
- [Das91] B.V. Dasarathy. *Nearest neighbor (NN) norms: NN pattern classification techniques*. Los Alamitos, CA: IEEE Computer Society Press, 1991.
- [Dav91a] Y. Davidor, *Genetic Algorithms and Robotics a Heuristic Strategy for Optimization*. World Scientific, 1991.
- [Dav91b] L. Davis, *Handbook of Genetic Algorithms*. New York: Van Nostrand Reinhold, 1991.

- [Daw96] R. Dawkins, *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design*. New York: W.W. Norton & Co., 1996.
- [DD01] I.G. Damousis and P. Dokopoulos, "A fuzzy expert system for the forecasting of wind speed and power generation in wind farms," in *Proc. 4<sup>th</sup> Int. Conf. Power Industry Computer Applications (PICA 2001)*, Sydney, 2001, pp. 63-69.
- [DD03] R. Drechsler and N. Drechsler, *Evolutionary Algorithms for embedded system design*. London: Kluwer Academic Publishers, 2003.
- [DdWdB01] E.I. Ducheyne, R.R. De Wulf, and B. De Baets, "Bi-objective genetic algorithm for forest management: a comparative study," in *Proc. 2001 Genetic and Evolutionary Computation Conf. – Late Breaking Papers*, California, 2001, pp. 63-66.
- [Deb01] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley and Sons, 2001.
- [Deb91] K. Deb, "Binary and Floating-Point Function Optimization using Messy Genetic Algorithms," Ph.D. thesis, Dept. Engineering Mechanics, University of Alabama, Tuscaloosa, AL 35487, 1991.
- [DeJ75] K. A. De Jong, "An Analysis of the Behavior of a Class of Genetic Adaptive Systems," Ph.D. thesis, University of Michigan, Ann Arbor, USA, 1975.
- [dFal97] I. de Falco, "An Introduction to Evolutionary Algorithms and their application to Aerofoil Design Problem – Part I: The Algorithms," invited paper at the von Karman Lecture Series 1997, Bruxelles.
- [DG01] K. Deb and T. Goel, "A Hybrid Multi-Objective Evolutionary Approach to Engineering Shape Design," in *First International Conference on Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C.A. Coello and D. Corne, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001, pp. 385-399.
- [DHR93] D. Driankov, H. Hellendoorn, and M. Reinfrank, *An Introduction to Fuzzy Control*. New York: Springer-Verlag, 1993.
- [DHS01] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. New York: John Wiley & Sons, 2001.
- [DJ98] R.P. Dick and N.K. Jha, "MOGAC: A Multiobjective Genetic Algorithm for Hardware-Software Co-synthesis of Hierarchical Heterogeneous Distributed Embedded Systems," *IEEE Trans. Computer-Aided Design*, vol. 17, no. 10, pp. 920-935, Oct. 1998.

- [dJHNS04] M.J. del Jesús, F. Hoffman, L.J. Navascués, and L. Sánchez, "Induction of Fuzzy-Rule-Based Classifiers with Evolutionary Boosting Algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 3, pp. 296-308, June 2004.
- [dMT93] M. de la Maza and B. Tidor, "An analysis of selection procedures with particular attention paid to proportional and Boltzmann selection," in *Proceedings of the Fifth International Conference on Genetic Algorithms*, S. Forrest, Ed. Morgan Kaufmann Publishers, 1993, pp. 124-131.
- [DN91] E. Davalo and P. Naim, *Neural Networks*. London: MacMillan Education LTD, 1991.
- [DO01] Y. Dote and S.J. Ovaska, "Industrial applications of soft computing: a review," *Proc. IEEE*, vol. 89, no. 9, pp. 1243-1265, Sept. 2001.
- [DP97] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. Orlando, FL: Academic Press, 1997
- [DPAM02] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, pp. 182-197, Apr. 2002.
- [DR03] J.L. Dorn and S.R. Ranjithan, "Evolutionary Multiobjective Optimization in Watershed Water Quality Management," in *Second International Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, C.M. Fonseca, P.J. Fleming, E. Zitzler, K. Deb and L. Thiele, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 2632, 2003, pp. 692-706.
- [DRFF00] N.M. Duarte, A. E. Ruano, C.M. Fonseca, and P.J. Fleming, "Accelerating Multi-Objective Control System Design Using a Neuro-Genetic Approach," in *Proc. 2000 IEEE Congress on Evolutionary Computation - Vol. 1*, New Jersey, 2000, pp. 392-397.
- [DSG93] K. A. De Jong, W. M. Spears, and D. F. Gordon, "Using genetic algorithms for concept learning," *Machine Learn.*, vol. 13, pp. 161-188, 1997.
- [ECMWF95] European Centre for Medium Range Weather Forecasts (ECMWF), *User guide to ECMWF products Version 2.1*. Reading, UK, 1995.
- [ECS89] L.J. Eshelman, A. Caruana, and J.D. Schaffer, "Biases in the Crossover Landscape," in *Proceedings of the Third International Conference on Genetics Algorithms*, J.D. Schaffer, Ed. Morgan Kaufmann Publishers, 1989, pp. 86-91.

- [EHM00] C. Emmanouilidis, A. Hunter, and J. MacIntyre, "A Multiobjective Evolutionary Setting for Feature Selection and a Commonality-Based Crossover Operator," in *Proc. IEEE 2000 Congress on Evolutionary Computation- Vol. 1*, New Jersey, 2000, pp. 309-316.
- [ELG97] S. Esquivel, A. Leiva, and R. Gallard, "Multiple crossover per couple in genetic algorithms," in *Proc. 4th IEEE Int. Conf. Evolutionary Computation (ICEC'97)*, Indianapolis, USA, 1997, pp. 103-106.
- [EMH01] M. Erickson, A. Mayer, and J. Horn, "The Niche Pareto Genetic Algorithm 2 Applied to the Design of Groundwater Remediation Systems," in *First International Conference on Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C.A. Coello and D. Corne, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001, pp. 681-695.
- [ENC04] A. Escobet, A. Nebot, and F.E. Cellier, "Visual-FIR: A new platform for modelling and prediction of dynamical systems," in *Proc. of the Summer Computer Simulation (SCSC'04)*, San José, CA, 2004, pp. 229-234.
- [Esh91] L. J. Eshelman, "The chc adaptive search algorithm: How to have safe search when engaging in non traditional genetic recombination," in *Foundations of Genetic Algorithms 1*, G.J.E. Rawlins, Ed. Morgan Kaufmann Publishers, 1991, pp. 265-283.
- [FA92] D. B. Fogel and W. Atmar, *Proceedings of the 1st Annual Conference on Evolutionary Programming*. San Diego: Evolutionary Programming Society, 1992.
- [FDA04] M. Farina, K. Deb and P. Amato, "Dynamic multiobjective optimization problems: test cases, approximations and applications," *IEEE Trans. Evol. Comput.*, vol. 8, no. 5, pp. 182-197, Oct. 2004.
- [FF93] C.M. Fonseca, and P.J. Fleming, "Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization," in *Proceedings of the Fifth International Conference on Genetic Algorithms*, S. Forrest, Ed. Morgan Kaufmann Publishers, 1993, pp. 416-423.
- [FF94] C.M. Fonseca, and P.J. Fleming, "An Overview of Evolutionary Algorithms in Multiobjective Optimization," Technical Report, Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K., 1994.
- [FF95] C.M. Fonseca, and P.J. Fleming, "An Overview of Evolutionary Algorithms in Multiobjective Optimization," *Evol. Comput.*, vol. 3, no. 1, pp. 1-16, 1995.

- [FJ96] B. Filipic and D. Juricic, "A genetic algorithm to support learning fuzzy control rules from examples," in *Genetic Algorithms and Soft Computing*, F. Herrera and J. L. Verdegay, Eds. Physica Verlag, 1996, pp. 403-418.
- [FMSZ03] W. Fornaciari, P. Micheli, F. Salice, and L. Zampella, "A First Step Towards Hw/Sw Partitioning of UML Specifications," in *IEEE/ACM Design Automation and Test in Europe (DATE'03)*, Munich, 2003, pp. 668-673.
- [Fog62] L. J. Fogel, "Toward inductive inference automata," in *Proc. Intl. Federation for Information Processing Congress*, Munich, 1962, pp. 395-399.
- [Fog63] L.J. Fogel, *Biotechnology: Concepts and Applications*. Prentice-Hall, 1963.
- [Fog64] L. J. Fogel, "On the Organization of Intellect," Ph.D. thesis, University of California, Los Angeles, CA, 1964.
- [Fog88] D. B. Fogel, "An Evolutionary Approach to the traveling salesman problem," *Biological Cybernetics*, vol. 60, pp. 139-144, 1988.
- [Fog91] D. B. Fogel, *System Identification through Simulated Evolution: A Machine Learning Approach*. Ginn Press, 1991.
- [Fog92] D. B. Fogel, "Evolving Artificial Intelligence," Ph.D. thesis, University of California, San Diego, CA, 1992.
- [Fog95] D. B. Fogel, *Evolutionary Computation. Toward a New Philosophy of Machine Intelligence*. IEEE Press, 1995.
- [Fog98] D. B. Fogel, *Evolutionary Computation: The Fossil Record*. IEEE Press, 1998.
- [For93] S. Forrest, *Proceedings of the Fifth International Conference on Genetic Algorithms*. Morgan Kaufmann Publishers, 1993.
- [FOW66] L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial Intelligence through Simulated Evolution*. New York: John Wiley & Sons, 1966.
- [FPW94] G. Franklin, J.D. Powell and M.L. Workman, *Digital Control of Dynamic Systems*. Addison Wesley, 1994.
- [FS93] J. Freeman and D. Skapura, *Redes Neuronales: Algoritmos, aplicaciones y técnicas de programación*. Addison Wesley-Ediciones Díaz de Santos, 1993.
- [FS96] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," in *Proc. 13<sup>th</sup> Int. Conf. Machine Learning*, Bari, Italy,

- 1996, pp. 148-156.
- [Ful99] R. Fuller, *Introduction to Neuro-Fuzzy Systems*. Physica-Verlag, 1999.
- [GD91] D.E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms I*, G.J.E. Rawlins, Ed. Morgan Kaufmann Publishers, 1991, pp. 69-93.
- [GDKH93] D. E. Goldberg, K. Deb, H. Kargupta, and G. Harik, "Rapid, accurate optimization of difficult problems using fast messy genetic algorithms," in *Proceedings of the Fifth International Conference on Genetic Algorithms*, S. Forrest, Ed. Morgan Kaufmann Publishers, 1993, pp. 56-64.
- [GG97] G. García and C. González, "Desarrollo de un sistema neuronal para sintonización de controladores difusos," Tesis de Pregrado, Univ. del Valle, Cali, Colombia, 1997.
- [GGP98] R. Gudwin, F. Gomide, and W. Pedrycz, "Context adaptation in fuzzy processing and genetic algorithms," *Int. J. Intell. Syst.*, vol. 13, no. 10-11, pp. 929-948, Oct./Nov. 1998.
- [GGW03] D. Greiner, B. Galván, and G. Winter, "Safety Systems Optimum Design by Multicriteria Evolutionary Algorithms," in *Second International Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, C.M. Fonseca, P.J. Fleming, E. Zitzler, K. Deb and L. Thiele, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 2632, 2003, pp. 722-736.
- [GH97] A. González and F. Herrera, "Multi-stage genetic fuzzy systems based on the iterative rule learning approach," *Mathware Soft Comput.*, vol. 4, pp. 233-249, 1997.
- [Ghi05] C. Ghiaus, "Linear fuzzy-discriminant analysis applied to forecast ozone concentration classes in sea-breeze regime", *Atmospheric Environment*, vol. 39, pp. 4691-4702, 2005.
- [GKD89] D.E. Goldberg, B. Korb, and K. Deb, "Messy genetic algorithms: motivation, analysis, and first results," *Complex Systems*, vol. 3, pp. 493-530, 1989.
- [GLM00] I.E. Golovkin, S.J. Louis, and R.C. Mancini, "Parallel Implementation of Niche Pareto Genetic Algorithm Code for X-ray Plasma Spectroscopy," in *Proc. 2000 Genetic and Evolutionary Computation Conf. - Late Breaking Papers*, Nevada, 2000, pp. 222-227.
- [Glo96] P. Y. Glorennec, "Constrained optimization of fis using an evolutionary method," in *Genetic Algorithms and Soft Computing*, F.

- Herrera and J. L. Verdegay, Eds. Physica Verlag, 1996, pp. 349-368.
- [GNMW01] P. Gómez, A. Nebot, F. Mugica, and F. Wotawa, “Fuzzy Inductive Reasoning for the Prediction of Maximum Ozone Concentration,” in *Proceedings of 13<sup>th</sup> European Simulation Symposium*, Marsella, 2001, pp. 535-542.
- [GNRA+03] P. Gómez, A. Nebot, S. Ribeiro, R. Alquézar, F. Mugica and F. Wotawa, “Local Maximum Ozone Concentration Prediction using Soft Computing Methodologies”, *Systems Analysis Modelling Simulation*, vol. 43, no. 8, pp. 1011-1031, 2003.
- [Gol02] D.E.Goldberg, *The Design of Component Genetic Algorithms: Steps Toward a Computational Theory of Innovation*. Kluwer Academic Publishers, 2002.
- [Gol89] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
- [GP01] A. González, and R. Pérez, “An experimental study about the research mechanism in SLAVE learning algorithm: hill-climbing methods versus genetic algorithm,” *Inform. Sci.*, vol. 136, no. 1-4, pp. 159-174, Aug. 2001.
- [GPVC02] A. Gaspar-Cunha, A. Poulesquen, B. Vergnes, and J. A. Covas, “Optimization of Processing Conditions for Polymer Twin-Screw Extrusion”, *Intern. Polym. Process.*, vol. 17, no. 3, p. 201-213, 2002.
- [GR87] D.E. Goldberg and J. Richardson, “Genetic algorithm with sharing for multimodal function optimization,” in *Proc. 2nd Int. Conf. Genetic Algorithms (ICGA'87)*, Hillsdale, 1987, pp. 41-49.
- [Gre84] J.J. Grefenstette, “GENESIS: A system for using genetic search procedures,” in *Proc. 1984 Conf. Intell. Systems and Machines*, Rochester, MI, 1984, pp. 161-165.
- [Gre86] J. J. Grefenstette, “Optimization of Control Parameters for Genetic Algorithms,” *IEEE Trans. Syst., Man, Cybern.*, vol. 16, no. 1, pp. 122-128, Jan./Feb. 1986.
- [Gre94] J. J. Grefenstette, *Genetic Algorithms for Machine Learning*. Kluwer Academic Press, 1994.
- [GSC00] F.A. Gers, J. Schmidhuber, and F. Cummins, “Learning to forget: continual prediction with LSTM”, *Neural Computation*, vol. 12, no. 10, pp. 2451-2471, 2000.
- [GT95] S. Goonatilake and P. Treleaven, *Intelligent Systems for Finance and Bussines*. John Willey & Sons, 1995.

- [Gür99] H.B. Gürocak, "A genetic-algorithm-based method for tuning fuzzy logic controllers," *Fuzzy Sets Syst.*, vol. 108, no. 1, pp. 39-47, Nov. 1999.
- [Han94] P.J.B. Hancock, "An empirical comparison of selection methods in evolutionary algorithms," in *Evolutionary Computing: AISB Workshop. Select Papers*, T. C. Fogarty, Ed. Springer-Verlag, 1994, pp. 80-94.
- [HBBB94] L.O. Hall, J.C. Bezdek, S. Boggavarpu, and A. Bensaid, "Genetic fuzzy clustering," in *Proc. NAFIPS'94*, San Antonio, 1994, pp. 411-415.
- [HBP02] A. Herreros, E. Baeyens, and J.R. Perán, "MRCD: a genetic algorithm for multiobjective robust control design," *Genetic Algorithm, Engineering Applications of Artificial Intelligence*, vol. 15, no. 3-4, pp. 285-301, June-Aug. 2002.
- [HD95] L. Howard and D. D'Angelo, "The GA-P: A genetic algorithm and genetic programming hybrid," *IEEE Expert*, pp. 11-15, 1995.
- [HD97a] H. Hellendoorn and D. Driankov, *Fuzzy Model Identification: Selected Approaches*. Berlin: Springer, 1997.
- [HD97b] H. Heider and T. Drabe, "A cascaded genetic algorithm for improving fuzzy-system design," *Int. J. Approx. Reason.*, vol. 17, no. 4, pp. 351-368, Nov. 1997.
- [HDB96] M. T. Hagan, H.B. Demuth, and M.H. Beale, *Neural Network Design*. Boston, MA: PWS Publishing, 1996.
- [Her90] A. Herrera, "Utilitarismo y ecología", *Estudios*, vol. 22, pp. 92-104, 1990.
- [HG97] P. Haastrup and Â. Guimarães, "Exploring the use of Multi-Objective Genetic Algorithms for Reducing Traffic Generated Urban Air and Noise Pollution," in *Proc. 5th European Cong. Intelligent and Soft Computing*, Aachen, 1997, pp. 819-825.
- [Hir93] K. Hirota, *Industrial Applications of Fuzzy Technology*. Springer-Verlag, 1993.
- [HK04] J.S. Heo and D.S. Kim, "A new method of ozone forecasting using fuzzy expert and natural network systems", *Science of the Total Environment*, vol. 325, pp. 221-237, 2004.
- [HL02] D. Hanselman and B. Littlefield, *Matlab the Language of Technical Computing. Matlab Notebook User's Guide-Versión 5.3*. The MathWorks Inc., 2002.

- [HL92] P. Hajela and C. Y. Lin, "Genetic search strategies in multicriterion optimal design," *Structural Optimization*, vol. 4, pp. 99-107, 1992.
- [HLV95] F. Herrera, M. Lozano, and J.L. Verdegay, "Tuning fuzzy controllers by genetic algorithms," *Int. J. Approx. Reason.*, vol. 12, no. 3-4, pp. 299-315, Apr./May 1995.
- [HLV96] F. Herrera, M. Lozano, and J.L. Verdegay, "Dynamic and Heuristic Fuzzy Connectives-Based Crossover Operators for Controlling the Diversity and Convergence of Real Coded Genetic Algorithms," *Int. J. Intell. Syst.*, vol. 11, no. 12, pp. 1013-1040, Dec. 1996.
- [HLV97] F. Herrera, M. Lozano and J. L. Verdegay, "Fuzzy connectives based crossover operators to model genetic algorithms population diversity," *Fuzzy Sets Syst.*, vol. 92, no. 1, pp. 21-30, 1997.
- [HLV98] F. Herrera, M. Lozano, and J. Verdegay, "Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis," *Artificial Intell. Review*, vol. 12, pp. 265-319, 1998.
- [HM97] F. Herrera and L. Magdalena, "Genetic fuzzy systems," in *Fuzzy Structures. Current Trends. Lecture Notes of the Tutorial: Genetic Fuzzy Systems. Seventh IFSA World Congress (IFSA'97)*, R. Mesiar and B. Riecan, Eds. Tatra Mountains Mathematical Publications, 1997, vol. 13, pp. 93-121.
- [HN01] F. Hoffmann and O. Nelles, "Genetic Programming for model selection of TSK-fuzzy systems," *Inform. Sci.*, vol. 136, no. 1-4, pp. 7-28, 2001.
- [HN93] J. Horn and N. Nafpliotis, "Multiobjective Optimization using the Niche Pareto Genetic Algorithms," IlliGAL Report 93005, Department of Computer Science, University of Illinois, Urbana, Champaign, 1993.
- [Hof01] F. Hoffmann, "Evolutionary Algorithms for Fuzzy Control System Design," *Proc. IEEE*, vol. 89, no. 9, pp. 1318-1333, Sept. 2001.
- [Hol62] J. Holland, "Outline for a logical theory of adaptative systems," *J. the ACM*, vol. 3, pp. 297-314, 1962.
- [Hol75] J. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor: The University of Michigan Press, 1975.
- [HP97] F. Hoffmann and G. Pfister, "Evolutionary design of a fuzzy knowledge base for a mobile robot," *Int. J. Approx. Reason.*, vol. 17, no. 4, pp. 447-469, Nov. 1997.
- [HR78] J. H. Holland and J. S. Reitman, "Cognitive systems based on

- adaptive algorithms,” in *Pattern-Directed Inference Systems*, D. A. Waterman and F. Hayes-Roth, Eds. Academic Press, 1978.
- [HS95] K. Hirota and M. Sugeno, *Industrial Applications of Fuzzy Technology in The World*, World Scientific, 1995.
- [HS97] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory”, *Neural Computation*, vol. 9, no. 9, pp. 1735-1780, 1997.
- [HT93] H. Hellendoorn and C. Thomas, “Defuzzification in fuzzy controllers,” *Int. J. Intell. Fuzzy Syst.*, vol. 1, pp. 109-123, 1993.
- [Hub89] G.L. Huber, “Physical, chemical, and biologic properties of tobacco, cigarette smoke, and other tobacco products”, *Seminars in Resp. Med.*, vol. 10, pp. 297-332, 1989.
- [HV96] F. Herrera and J. L. Verdegay, *Genetic Algorithms and Soft Computing*. Physica-Verlag, 1996.
- [IM96] H. Ishibuchi and T. Murata, “A genetic-algorithm-based fuzzy partition method for pattern classification problems,” in *Genetic Algorithms and Soft Computing*, F. Herrera and J. L. Verdegay, Eds. Physica Verlag, 1996, pp. 555-578.
- [INM99] H. Ishibuchi, T. Nakashima, and T. Murata, “Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 29, no. 5, pp. 601-618, Oct. 1999.
- [IS96] P. Ioannou and J. Sun, *Robust Adaptive Control*. Prentice Hall, 1996.
- [Jan93a] J.-S. R. Jang, “ANFIS: adaptive-network-based fuzzy inference system,” *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 3, pp. 665-684, 1993.
- [Jan93b] C. Z. Janikov, “A knowledge-intensive genetic algorithm for supervised learning,” *Machine Learn.*, vol. 13, pp.189-228, 1993.
- [Jin00] Y. Jin, “Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement,” *IEEE Trans. Fuzzy Systems*, vol. 8, pp. 212-222, 2000.
- [Jin03] Y. Jin, *Advanced Fuzzy Systems Design and Applications. Series: Studies in Fuzziness and Soft Computing. Vol. 112*, New York: Physica-Springer, 2003.
- [Jin04] Y. Jin, *Knowledge Incorporation in Evolutionary Computation. Series: Studies in Fuzziness and Soft Computing. Vol. 167*. Heidelberg: Springer-Verlag, 2004.

- [JN97] T. Jerez and A. Nebot, "Genetic algorithms vs. classical search techniques for identification of fuzzy models," in *Proc. 5th Eur. Congr. Intelligent Techniques and Soft Computing (EUFIT'97)*, Aachen, Germany, 1997, pp. 769-773.
- [Joh93] R. Johansson, *System Modeling and Identification*. Prentice Hall, 1993.
- [JSM97] J.S.R. Jang, C.T. Sun, and E. Mizutani, *Neuro-fuzzy and Soft Computing*. New York: Prentice-Hall, 1997.
- [KA05] M. Maya and R. Alhaji, "Genetic algorithm based framework for mining fuzzy association rules," *Fuzzy Sets Syst.*, vol. 152, no. 3, pp. 587-601, June 2005.
- [Kar91a] C. Karr, "Genetic algorithms for Fuzzy Controllers," *AI Expert*, vol. 6, no. 2, pp. 26-33, Feb. 1991.
- [Kar91b] C. Karr, "Applying genetics to Fuzzy Logic," *AI Expert*, vol. 6, no. 3, pp. 38-43, Mar. 1991.
- [KC00] J.D. Knowles and D.W. Corne, "Approximating the Nondominated Front Using the Pareto Archived Evolution Strategy," *Evol. Comput.*, vol. 8, no. 2, pp. 149-172, 2000.
- [KC99a] J.D. Knowles and D.W. Corne, "The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Multiobjective Optimisation," in *Proc. 1999 Cong. Evolut. Comput. (CEC'99)*, Washington D.C., USA, July 6-9, 1999, pp. 98-105.
- [KC99b] J.D. Knowles and D.W. Corne, "Assessing the Performance of the Pareto Archived Evolution Strategy," in *Proceedings of the 1999 Genetic and Evolutionary Computation Conference. Workshop Program*, Orlando, FL, 1999, pp. 123-124.
- [KCO00] J.D. Knowles, D.W. Corne, and M.J. Oates, "On the Assessment of Multiobjective Approaches to the Adaptive Distributed Database Management Problem," in *Proceedings of the Sixth International Conference on Parallel Problem Solving from Nature (PPSN VI)*, M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J.J. Merelo and H.-P. Schwefel, Eds. Springer, 2000, pp. 869-878.
- [KFAK99] J. Koza, B. Forrest, D. Andre and M. Keane, *Genetic Programming III: Darwinian Invention and Problem Solving*. Morgan Kaufmann Publishers, 1999.
- [Kli85] G. Klir, *Architecture of Systems Problem Solving*. New York: Plenum Press, 1985.
- [KOC00] J.D. Knowles, M.J. Oates, and D.W. Corne, "Multiobjective

- Evolutionary Algorithms Applied to two Problems in Telecommunications,” *BT Technology Journal*, vol. 18, no. 4, pp. 51-64, Oct. 2000.
- [Koz90] J. R. Koza, “Genetic Programming: A paradigm for genetically breeding populations of computer programs to solve problems,” Department of Computer Science, Stanford University, Stanford, CA, Tech. Rep. STAN-CS-90-1314, 1990.
- [Koz92] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. The MIT Press, 1992.
- [Koz94] J. R. Koza, *Genetic Programming II: Automatic Discovery of Reusable Programs*. The MIT Press, 1994.
- [Kun00] L. I. Kuncheva, *Fuzzy Classifier Design*. Heidelberg: Physica-Verlag, 2000.
- [KY95] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. New Jersey: Prentice-Hall, 1995.
- [LC90] D. Li and F.E. Cellier, “Fuzzy measures in inductive reasoning,” in *Proc. Winter Simulation Conference*, New Orleans, LA, 1990, pp. 527-538.
- [LE97] J. Lis and A.E. Eiben, “A Multi-Sexual Genetic Algorithm for Multiobjective Optimization,” in *Proc. 1997 IEEE Int. Conf. Evolut. Comput. (ICEC’97)*, Indianapolis, USA, April 13-16, 1997, pp. 59-64.
- [Leo99] C. T. Leondes, *Fuzzy Theory Systems: Techniques and Applications, vols. 1-4*. San Diego: Academic Press, 1999.
- [Lez00] J.L. Lezama, “Aire dividido. Crítica a la política del aire en el Valle de México. 1979-1996,” México: El Colegio de México, 2000.
- [LH04] C.H. Lin and P.J. Hsied, “A fuzzy decision support system for strategic portofolio management,” *Decision Support Systems*, vol. 38, no. 3, pp. 383-398, Dec. 2004.
- [LHP+04] A. Lengyel, K. Héberger, L. Paksy, O. Bánhidi, and R. Rajkó, “Prediction of ozone concentration in ambient air using multivariate methods”, *Chemosphere*, vol. 57, pp. 889-896, 2004.
- [Lju99] L. Ljung. *System Identification: Theory for the user*. New Jersey: Prentice-Hall, 1999.
- [LK91] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, 2nd Ed. New York: McGraw Hill, 1991.
- [LMBZ01] M. Lahanas, N. Milickovic, D. Baltas, and N. Zamboglou,

- “Application of Multiobjective Evolutionary Algorithms for Dose Optimization Problems in Brachytherapy,” in *First International Conference on Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C.A. Coello and D. Corne, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001, pp. 574-587.
- [Loi96] W. Loibl, “Trendprognose regionaler Ozonmaxima unter Einberzug verschiedener meteorologischer Daten,” Technical Report UBA-BE-058, Umweltbundesamt, Vienna, Austria, 1996.
- [LSMB03] M. Lahanas, E. Schreibmann, N. Milickovic, and D. Baltas, “Intensity Modulated Beam Radiation Therapy Dose Optimization with Multiobjective Evolutionary Algorithms,” in *Second International Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, C.M. Fonseca, P.J. Fleming, E. Zitzler, K. Deb and L. Thiele, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 2632, 2003, pp. 648-661.
- [LX06] C.-J. Lin and Y.-J. Xu, “A self-adaptative neural fuzzy network with group-based symbiotic evolution and its prediction applications,” *Fuzzy Sets Syst.*, vol. 157, no. 8, pp. 1036-1056, Apr. 2006.
- [MA75] E. H. Mamdani and S. Assilian, “An experiment in linguistic synthesis with a fuzzy logic controller,” *Int. J. Man-Mach. Stud.*, vol. 7, pp. 1-13, 1975.
- [MA98] P.M. Murphy and D.W. Aha, “UCI Repository of Machine Learning Databases,” Department of Information and Computer Science, University of California, Irvine, CA, 1998.
- [Mag97] L. Magdalena, “Adapting the gain of an FLC with genetic algorithms,” *Int. J. Approx. Reason.*, vol. 17, no. 4, pp. 327-349, Nov. 1997.
- [Mam74] E. H. Mamdani, “Applications of fuzzy algorithm for control a simple dynamic plant,” *Proc. IEE*, vol. 121, no 12, pp. 1585-1588, 1974.
- [Mar02] J.M. Marcos, “Historia y panorama actual del sistema eléctrico español,” *Física y Sociedad-Colegio Oficial de Físicos*, no. 13, pp. 10-17, 2002.
- [Mar97] T. Marcu, “A multiobjective evolutionary approach to pattern recognition for robust diagnosis of process faults,” in *Proc. IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes: SAFEPROCESS'97*, Hull, UK, 1997, pp. 1183-1188.
- [Mau94] A. Mauzy, “Estudio Global de la Calidad de Aire en la Ciudad de México,” Final Technical Report LA-12699, Instituto Mexicano del Petróleo and Los Alamos National Laboratory, México, 1994.

- [MB05] B.K. Mohanty and B. Bhasker, "Product classification in the internet bussiness- a fuzzy approach," *Decision Support Systems*, vol. 38, no. 4, pp. 611-619, Jan. 2005.
- [MC94] F. Mugica and F. Cellier, "Automated synthesis of a fuzzy controller for cargo ship steering by means of qualitative simulation," in *Proc. Eur. Simulation MultiConference (ESM'94)*, Barcelona, 1994, pp. 523-528.
- [MCH98] W. Mason, V. Coverstone-Carroll, and J. Hartmann, "Optimal Earth Orbiting Satellite Constellations via a Pareto Genetic Algorithm," in *1998 AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Massachusetts, pp. 169-177, 1998.
- [MI95] T. Murata and H. Ishibuchi, "MOGA: Multi-objective genetic algorithms," in *Proc. 2nd IEEE Int. Conf. Evol. Comput.*, Perth, Australia, November 29-December 1, 1995, pp. 289-294.
- [Mic96] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*. New York: Springer-Verlag, 1996.
- [Mit97] M. Mitchell, *An Introduction to Genetic Algorithms*. London: The MIT Press, 1997.
- [Miy89] S. Miyamoto, "Two approaches for information retrieval through fuzzy associations," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, no. 1, pp. 123-130, Jan./Feb. 1989.
- [MLDP99] N. Marco, S. Lanteri, J.-A. Desideri, and J. Périaux, "A Parallel Genetic Algorithm for Multi-Objective Optimization in Computational Fluid Dynamics," in *Evolutionary Algorithms in Engineering and Computer Science*, K. Miettinen, M.M. Mäkelä, P. Neittaanmäki, and J. Périaux, Eds. John Wiley & Sons, 1999, pp. 445-456.
- [MM02] M. Molina and L. Molina, "Air Quality in the Mexico Megacity. An Integrated Assessment," Dordrecht, Holanda: Kluwer Academic Publishers, 2002.
- [MMV03] J. Malinen, M. Maltamo, and E. Verkasalo, "Predicting the internal quality and value of Norway spruce trees by using two non-parametric nearest neighbour methods," *Forest Products Journal.*, vol. 53, no. 4, pp. 85-94, 2003.
- [MNM96] Z. Michalewicz, G. Nazhiyath, and M. Michalewicz, "A Note on Usefulness of Geometrical Crossover for Numerical Optimization problems," in *Proc. 5th Annual Conf. on Evolutionary Programming*, San Diego, CA, 1996, pp. 305-312.

- [Mor80] J. N. Morse, "Reducing the size of the nondominated set: Pruning by clustering," *Computers and Operations Research*, vol. 7, no. 1-2, pp. 55-66, 1980.
- [Mos95] E. Mosca, *Optimal Predictive and Adaptive Control*. Prentice Hall, 1995.
- [MSV93] H. Mühlenbein and D. Schlierkamp-Voosen, "Predictive models for the Breeder Genetic Algorithm I. Continuous Parameter Optimization," *Evol. Comput.*, vol. 1, no. 1, pp. 25-49, 1993.
- [MV03] F.C. Morabito and M. Versaci, "Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data", *Neural Networks*, vol. 16, pp. 493-505, 2003.
- [MV97] L. Magdalena and J. R. Velasco, "Evolutionary based learning of fuzzy controllers," in *Fuzzy Evolutionary Computation*, W. Pedrycz, Ed. Kluwer Academic Press, 1997, pp. 249-268.
- [MYS05] R. Mintz, B. R. Young, and W. Y. Svrcek, "Fuzzy logic modeling of surface ozone concentrations", *Computers & Chemical Engineering*, vol. 29, pp.2049-2059, 2005
- [NA99] S. Narayanan and S. Azarm, "On Improving Multiobjective Genetic Algorithms for Design Optimization," *Structural Optimization*, vol. 18, pp. 146-155, 1999.
- [Nau95] D. Nauck, *Beyond Neuro-Fuzzy: Perspectives and Directions*. Technical University of Braunschweig, Germany, 1995.
- [NCL96] A. Nebot, F. Cellier, and D. Linkens, "Synthesis of an anaesthetic agent administration system using fuzzy inductive reasoning," *Artificial Intell. in Medicine*, vol. 8, no. 3, pp. 147-166, 1996.
- [NCV98] A. Nebot, F. Cellier, and M. Vallerdú, "Mixed quantitative/qualitative modeling and simulation of the cardiovascular system," *Computers Methods an Programs in Biomedicine*, vol. 55, pp. 127-155, 1998.
- [Neb94] A. Nebot, "Qualitative Modeling and Simulation of Biomedical Systems Using Fuzzy Inductive Reasoning," Ph.D. thesis, Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 1994.
- [Nis93] V. Nissen, "Evolutionary algorithms in Management Science: An overview and list of references," Tech. Rep. # 9303, Papers on Economics and Evolution, Universitaet Goettingen, 1993.
- [NK93] D. Nauck and R. Kruse, "A Fuzzy Neural Network Learning Fuzzy Control Rules and Membreship Functions by fuzzy error propagation,"

- in *Proc. IEEE Int. Conf. on Neural Networks*, vol. 2, San Francisco, CA, 1993, pp. 1022-1027.
- [NKK97] D. Nauck, F. Klawoon, and R. Kruse, *Foundations of Neuro-Fuzzy Systems*. London: John Willey & Sons, 1997.
- [NM06] A. Nebot and V. Mugica, "Finding qualitative patterns in ozono behavior," in *Proceedings of 5<sup>th</sup> Mexican International Conference on Artificial Intelligence (MICAI'06)*, Apizaco, México, 2006, pp. 91-100.
- [NMG01] A. Nebot, F. Mugica, and P. Gómez, "Long term prediction of maximum ozone concentration using fuzzy inductive reasoning," in *Proc. Eur. Symp. Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems (EUNITE'01)*, Tenerife, 2001, pp. 91-101.
- [NMP97] S. Nascimiento and F. Moura-Pires, "A genetic approach for fuzzy clustering with a validity measure function," in *Proc. 2nd Int. Conf. Intelligent Data Analysis (IDA'97)*, X. Liu, P. Cohen, and M.R. Berthold, Eds. Springer-Verlag, 1997, pp. 325-335.
- [NP90] K. Narendra and K. Parthasarathy, "Identification and Control of Dynamical Systems using neuralnetworks," *IEEE Trans Neural Networks*, vol. 1, no. 1, pp. 4-27, Mar. 1990.
- [ODH04] G. Onkal-Engin, I. Demir, and H. Hiz, "Assessment of urban air quality in Istanbul using fuzzy synthetic evaluation", *Atmospheric Environment*, vol. 38, pp. 3809-3815, 2004.
- [OGC91] C. K. Oei, D. E. Goldberg, and S.-J. Chang. Tournament Selection, Niching, and the Preservation of Diversity. Technical Report 91011, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1991.
- [OHM01] D. Ortiz, C. Hervás, and J. Muñoz, "Genetic algorithm with crossover based on confidence intervals as an alternative to least squares estimation for nonlinear models," in *Proc. 4th Metaheuristics Int. Conf.*, Porto, 2001, pp. 343-347.
- [OO95] U.-M. O'Reilly and F. Oppacher, "The troubling aspects of a building block hypothesis for genetic programming," in *Foundations of Genetics Algorithms 3*, L.D. Whitley and M.D. Volsse, Eds. Morgan Kaufmann Publishers, 1995, pp. 73-88.
- [OP95] Y. Oh and D. J. Park, "Self-tuning fuzzy controller with variable universe of discourse," in *Proc. IEEE Int. Conf. on System, Man and Cybernetics*, Vancouver, 1995, pp. 2628-2632.
- [Ova04] S. Ovaska, *Computationally Intelligent Hybrid Systems: The Fusion of*

*Soft Computing and Hard Computing*. Wiley-IEEE Press, 2004.

- [Pap00] D. Papadias, "Hill climbing algorithms for content-based retrieval of similar configurations," in *Proc. 23<sup>rd</sup> Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Athens, Greece, 2000, pp. 240-247.
- [PDP+00] N. Peton, G. Dray, D. Pearson, M. Mesbah, B. Vuillot, "Modelling and analysis of ozone episodes", *Environmental Modelling and Software*, vol. 15, pp. 647-652, 2000.
- [Ped96] W. Pedrycz, *Fuzzy Modelling: Paradigms and Practice*. Kluwer Academic Press, 1996.
- [Ped97] W. Pedrycz, *Fuzzy Evolutionary Computation*. Kluwer Academic Publishers, 1997.
- [Per92] N.J. Pérez, "La contaminación del aire en la zona metropolitana de la Ciudad de México," *Revista del Instituto Nacional de Enf. Resp. Mexicano*, vol. 5, pp. 63-65, 1992.
- [PK91] D.T. Pham and D. Karaboga, "Optimum design of fuzzy logic controllers using genetic algorithms," *J. System Eng.*, vol. 1, pp. 14-118, 1991.
- [PM01] A. Petrovski and J. McCall, "Multi-objective Optimisation of Cancer Chemotherapy Using Evolutionary Algorithms," in *First International Conference on Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C.A. Coello, and D. Corne, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001 pp. 531-545.
- [PRG+04] H. Pomares, I. Rojas, J. González, M. Damas, B. Pino, and A. Prieto, "Online Global Learning in Direct Fuzzy Controllers", *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 2, pp.218-229, Apr. 2004.
- [PS82] C.H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. New York: Prentice-Hall, 1982.
- [Rec65] I. Rechenberg, "Cybernetic solution path of an experimental problem," Royal Aircraft Establishment, Library Translation No. 1122, Farnborough, Hants, UK, Aug. 1965.
- [Rec73] I. Rechenberg, *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipen der biologischen Evolution*. Frommann-Holzboorg Verlag, FRG, 1973.
- [Ree93] C.R. Reeves, *Modern Heuristic Techniques for Combinatorial Problems*. Oxford: Blackwell Scientific Publications, 1993.

- [Ree96] C.R. Reeves, "Heuristic search methods: A review," in *Operational Research: Keynote Papers 1996*, D. Johnson and F. O'Brien, Birmingham: Operational Research Society, 1996, pp. 122-149.
- [RER94] B.J. Ritzel, J.W. Eheart, and S. Ranjithan, "Using genetic algorithms to solve a multiple objective groundwater pollution containment problem," *Water Resources Research*, vol. 30, no. 5, pp. 1589-1603, May 1994.
- [RFK02] A. Rajapakse, K. Furuta, and S. Kondo, "Evolutionary learning of fuzzy logic controllers and their adaptation through perpetual evolution," *IEEE Trans. Fuzzy Systems*, vol. 10, no. 3, pp. 309-321, 2002.
- [RHBF03] R.V. Rohli, S.A. Hsu, B.W. Blanchard, and R.L. Fontenot, "Short-Range Prediction of Tropospheric Ozone Concentrations and Exceedances for Baton Rouge, Louisiana," *Weather and Forecasting*, vol. 18, pp. 371-383, 2003.
- [RMG01] Patrick M. Reed, Barbara S. Minsker and David E. Goldberg, "A multiobjective approach to cost effective long-term groundwater monitoring using an elitist nondominated sorted genetic algorithm with historical data," *Journal of Hydroinformatics*, vol. 3, no. 2, pp. 71-89, April 2001.
- [RN03] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. NJ: Prentice-Hall, 2003.
- [RO96] E. Ruiz and E. Ortiz, "Simulación matemática de la formación de ozono en la zona metropolitana de la ciudad de México," Technical Report, Instituto Mexicano de Petróleo, México, 1996.
- [Rog00] J.L. Rogers, "A Parallel Approach to Optimum Actuator Selection With A Genetic Algorithm," presented at the AIAA Guidance, Navigation, and Control Conf. *and Exhibit*, Denver, CO, Aug. 14-17, 2000, Paper AIAA-2000-4484.
- [Ros67] R S. Rosenberg, "Simulation of genetic populations with biochemical properties," Ph.D. thesis, University of Michigan, Ann Harbor, Michigan, 1967.
- [RPLH89] J.T. Richardson, M. R. Palmer, G. Liepins, and M. Hilliard, "Some Guidelines for Genetic Algorithms with Penalty Functions," in *Proceedings of the Third International Conference on Genetic Algorithms (ICGA-89)*, J. D. Schaffer, Ed. Morgan Kaufmann Publishers, 1989, pp. 191-197.
- [RRT05] K. Riveles, R. Roza, and P. Talbot, "Phenols, Quinolines, Indoles,

- Benzene, and 2-Cyclopenten-1-ones are Oviductal Toxicants in Cigarette Smoke,” *Toxicol. Sci.*, vol. 86, no. 1, pp. 141 – 151, July 2005.
- [RSP00] A. Rodgman, C.J. Smith, and T.A. Perfetti, “The composition of cigarette smoke: A retrospective, with emphasis on polycyclic components,” *Hum. Exp. Toxicol.*, vol. 19, 573–595, 2000.
- [Rus98] M. Russo, “FuGeNeSys: a Fuzzy genetic neural system for fuzzy modeling,” *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 3, pp. 373-388, Aug. 1998.
- [San97] L. Sánchez, “*Estudio de la red asturiana de baja tensión rural y urbana*,” Technical Report, Research and Development Department, Hidroeléctrica del Cantábrico (in spanish), Asturias, Spain, 1997.
- [SCC01] L. Sanchez, I. Couso, and J.A. Corrales, “Combining GP operators with SA search to evolve fuzzy rule-based classifiers,” *Inform. Sci.*, vol. 136, no. 1-4, pp. 175-191, 2001.
- [SCFG97] P. Schroder, A. J. ChipPerfield, P. J. Fleming, and N. Grum, “Multi-objective optimization of distributed active magnetic bearing controllers,” in *Proc. IEEE Int. Conf. Genetic Algorithms in Engineering Systems: Innovations and Application*, 1997, pp. 13-18.
- [Sch65] H.P. Schwefel, “Kybernetische Evolution als Strategie der experimentellen Forschung in der Strömungstechnik,” Diplomarbeit, Technische Universität Berlin, 1965.
- [Sch77] H P. Schwefel, *Numerische Optimierung von Computer-modellen mittels der Evolutionsstrategie*. Birkhäuser Verlag, 1977.
- [Sch84] J.D. Schaffer, “Some experiments in machine learning using vector evaluated genetic algorithms,” Unpublished doctoral dissertation, Vanderbilt University, 1984.
- [Sch85] J.D. Schaffer, “Multiple objective optimization with vector evaluated genetic algorithms,” in *Proc. 1st Int. Conf. Genetic Algorithms and Their Applications*, Pittsburgh, PA, July 24-26, 1985, pp. 93-100.
- [Sch94] D. Schlierkamp-Voosen, “Strategy Adaptation by Competition,” in *Proc. Second European Congress on Intelligent Techniques and Soft Computing (EUFIT'94)*, Aachen, Germany, 1994, pp. 1270-1274.
- [Sch95a] H P. Schwefel, *Evolution and Optimum Seeking. Sixth-Generation Computer Technology Series*. John Wiley & Sons, 1995.
- [Sch95b] J.R. Schott, “Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization,” Master's thesis, Department of

- Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1995.
- [Sch99] M. Schoenauer, *Evolutionary Computation: Special Issue "Scalable Evolutionary Computation"* vol. 7- no. 4. The MIT Press, 1999. ISSN 1063-6560.
- [SD91] W.M. Spears and K.A. De Jong, "On the virtues of parameterized uniform crossover," in *Proceedings of the Fourth International Conference on Genetics Algorithms*, R. K. Belew and L. B. Booker, Eds. Morgan Kaufmann Publishers, 1991, pp. 230-236.
- [SD93] N. Srinivas and K. Deb, "Multiobjective Optimization using Nondominated Sorting in Genetic Algorithms," Technical Report, Department of Mechanical Engineering, Indian Institute of Technology, Kanpur, India, 1993.
- [SD94] N. Srinivas and K. Deb, "Multiobjective Optimization using Nondominated Sorting in Genetic Algorithms," *Evol. Comput.*, vol. 2, no. 3, pp. 221-248, 1994.
- [SDI05] G. Shakhnarovich, T. Darrell, and P. Indyk. *Nearest-neighbor methods in learning and vision: Theory and Practice*. Cambridge: MIT Press, 2005.
- [Sha48] C. Shannon, "The mathematical theory of communication— 1 and 2," *The Bell Systems Tech. J.*, vol. 27, no. 1, 2, pp. 379-423, 623-656, July/Oct. 1948.
- [SISM01] Y. Suzuki, K. Itakura, S. Saga, and J. Maeda, "Signal processing and pattern recognition with soft computing," *Proc. IEEE*, vol. 89, no. 9, pp. 1297-1317, Sept. 2001.
- [SKM02] Th. Slini, K. Karatzas, and N. Moussiopoulos, "Statistical analysis of environmental data as the basis of forecasting: an air quality application", *The Science of the Total Environment*, vol. 288, pp. 227-237, 2002.
- [Smi80] S. F. Smith, "A Learning System Based on Genetic Adaptive Algorithms," Ph.D. thesis, Universidad de Pittsburgh, 1980.
- [SMPA06] S.I. Sousa, F.G. Martins, M.C. Pereira, and M.C. Alvim-Ferraz, "Prediction of ozone concentrations in Oporto city with statistical approaches," *Chemosphere*, vol. 64, no. 7, pp. 1141-1149, Aug. 2006.
- [SMSS88] K. Sagawa, L. Maughan, H. Suga and K. Sunagawa, *Cardiac Contraction and the Pressure--Volume Relationship*. New York: Oxford University Press, 1988.

- [SR97] P.D. Surry and N. J. Radcliffe, "The COMOGA Method: Constrained Optimisation by Multiobjective Genetic Algorithms," *Control and Cybernetics*, vol. 26, no. 3, pp. 391-412, 1997.
- [SSZ97] E. Sánchez, T. Shibata, and L. Zadeh, *Genetic Algorithms and Fuzzy Logic Systems. Soft Computing Perspectives*. World Scientific, 1997.
- [ST99] F. Sattar and D. B. H. Tay, "Enhancement of document images using multiresolution and fuzzy logic techniques," *IEEE Signal Processing Lett.*, vol. 6, no. 10, pp. 249-252, Oct. 1999.
- [SWKK96] A. Stohl, G. Wotawa, and H. Kromp-Kolb, "The IMPO modeling system description, sensitivity studies and applications," Technical Report, Universität für Bodenkultur, Institut für Meteorologie und Physik, Vienna, Austria, 1996.
- [SY93] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modelling," *IEEE Trans. Fuzzy Syst.*, vol. 1, no. 1, pp. 7-31, 1993.
- [Sys89] G. Syswerda, "Uniform Crossover in Genetic Algorithms," in *Proceedings of the Third International Conference on Genetics Algorithms*, J.D. Schaffer, Ed. Morgan Kaufmann Publishers, 1989, pp. 2-9.
- [SYZL05] Z. Song, J. Yi, D. Zhao, and X. Li, "A computed torque controller for uncertain robotic manipulator systems: Fuzzy approach," *Fuzzy Sets Syst.*, vol. 154, no. 2, pp. 228-226, Sept. 2005.
- [Tac94] W. A. Tackett, "Recombination, Selection, and the Genetic Construction of Computer Programs," Ph.D. thesis, Department of Computer Engineering, University of Southern California, 1994.
- [TC03] G. Toscano and C.A. Coello, "The Micro Genetic Algorithm 2: Towards Online Adaptation in Evolutionary Multiobjective Optimization," in *Second International Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, C.M. Fonseca, P.J. Fleming, E. Zitzler, K. Deb and L. Thiele, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 2632, 2003, pp. 252-266.
- [Tet95] A. G. Tettamanzi, "Evolutionary algorithms and fuzzy logic: a two-way integration," in *Proc. 2nd Joint Conf. Information Sciences*, 1995, pp. 464-467.
- [Thr91] P. Thrift, "Fuzzy logic synthesis with genetic algorithms," in *Proc. 4th Int. Conf. Genetic Algorithms (ICGA'91)*, San Diego, CA, 1991, pp. 509-513.

- [TIG98] T. Murata, H. Ishibuchi, and M. Gen, "Random Weights in Multi-Objective Genetic Algorithms," in *Proc. 2nd Int. Conf. on Engineering Design and Automation*, Maui, Hawaii, 1998. Only CD-ROM Proceedings available.
- [Tos05] G. Toscano, "On the Use of Self-Adaptation and Elitism for Multiobjective Particle Swarm Optimization," Ph.D. thesis, Computer Science Section, Department of Electrical Engineering, CINVESTAV-IPN, September 2005.
- [TS85] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. 15, pp. 116–132, Jan.–Feb. 1985.
- [VA94] H.M. Voigt and T. Anheyer, "Modal Mutations in Evolutionary Algorithms," in *Proc. 1st IEEE Conf. Evolutionary Computation (CEC'94)*, Orlando, Florida, 1994, pp. 88-92.
- [Val93] M. Vallverdú, "Modelado y simulación del sistema de control cardiovascular en pacientes con lesiones coronarias," Ph.D dissertation, Universitat Politècnica de Catalunya, Barcelona, Spain, 1993.
- [Van95] T. Van Le, "Evolutionary fuzzy clustering," in *Proc. 2nd IEEE Conf. Evolutionary Computation (ICEC'95)*, vol. 2, Perth, 1995, pp. 753-758.
- [Ven92] G. Venturini, "SIA: a supervised inductive algorithm with genetic search for learning attribute based concepts," in *Proc. Eur. Conf. Machine Learning*, Viena, 1992, pp. 280-296.
- [Vil00] P. Villar, "Aprendizaje de las Particiones Difusas para Sistemas Basados en Reglas Lingüísticas," Ph.D. thesis, Dpto. Informática, Universidade de Vigo, Ourense, España, 2000.
- [VL00a] D.A. Van Veldhuizen and G.B. Lamont, "Multiobjective evolutionary algorithms: Analyzing the state-of-the-art," *Evol. Comput.*, vol. 8, no. 2, pp. 125-147, 2000.
- [VL00b] D.A. Van Veldhuizen and G.B. Lamont, "Multiobjective Optimization with Messy Genetic Algorithms," in *Proc. 2000 ACM Symposium on Applied Computing*, Villa Olmo, Italy, 2000, pp. 470-476.
- [VLA88] P.J.M. Van Laarhoven and E.H.L. Aarts, *Simulated Annealing: Theory and Applications*. Kluwer, 1988.
- [VMC95] H.M. Voigt, H. Mühlenbein, and D. Cvetkovic, "Fuzzy Recombination for the Breeder Genetic Algorithm," in *Proceedings of the Sixth International Conference on Genetic Algorithms*, L. Eshelman, Ed.

- Morgan Kaufmann Publishers, 1995, pp. 104-111.
- [VR91] M. Valenzuela-Rendón, “The fuzzy classifier system: a classifier system for continuously varying variables,” in *Proc. 4<sup>th</sup> Int. Conf. Genetic Algorithms (ICGA'91)*, San Diego, CA, 1991, pp. 346-353.
- [Wan92] L. X. Wang, “Fuzzy systems are universal approximators,” in *Proc. 1st IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE'92)*, San Diego, CA, 1992, pp. 1163-1170.
- [Wan94] L. X. Wang, *Adaptive Fuzzy Systems and Control*. Prentice-Hall, 1994.
- [WB03] L. Willmes and T. Bäck, “Multi-criteria Airfoil Design with Evolution Strategies,” in *Second International Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, C.M. Fonseca, P.J. Fleming, E. Zitzler, K. Deb and L. Thiele, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 2632, 2003, pp. 782-795.
- [WC87] A.K.C. Wong and D.K.Y. Chiu, “Synthesizing statistical knowledge from incomplete mixed-mode data”, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 6, pp. 796-805, 1987.
- [Whi89] D. Whitley, “The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best,” in *Proceedings of the Thirth International Conference on Genetics Algorithms and Their Applications*, J. D. Schaffer, Ed. Morgan Kaufmann Publishers, 1989, pp. 116-121.
- [WL06] D. Wang and W.Z. Lu, “Interval estimation of urban ozone level and selection of influential factors by employing automatic relevance determination model”, *Chemosphere*, vol. 62, pp. 1600-1611, 2006.
- [WLWL03] W. Wang, W. Lu, X. Wang, and A.Y.T. Leung, “Prediction of maximum daily ozone level using combined neural network and statistical characteristics”, *Environmental International*, vol.29, pp. 555-562, 2003.
- [WM92] L.X. Wang and J.M. Mendel, “Generating fuzzy rules by learning from examples,” *IEEE Trans. Syst., Man, Cybern.*, vol. 22, pp. 1414–1427, 1992.
- [Wri91] A. Wright, “Genetic Algorithms for Real Parameter Optimization,” in *Foundations of Genetic Algorithms 1*, G.J.E. Rawlins, Ed. Morgan Kaufmann Publishers, 1991, pp. 205-218.
- [WS92] L. D. Whitley and J. D. Schaffer (Eds), *Proceedings of the International Workshop on Combinations of Genetic Algorithms and*

*Neural Networks*. IEEE Computer Society Press, 1992.

- [WW01] F. Wotawa and G. Wotawa, "Deriving Qualitative Rules from Neural Networks- A case study for Ozone Forecasting", *AI Communications*, vol. 14, no. 1, pp. 23-33, 2001.
- [WW99] D. Wieland and W. Wotawa, "Local Maximum Ozone Concentration Prediction Using Neural Networks," in *Proceedings of the AAI Workshop on Environmental Decision Support Systems and Artificial Intelligence (AAI'99)*, U. Corté and M. Sánchez-Marrè, Eds. Technical Report WS-99-07. The AAI Press, 1999, pp. 47-54.
- [YB06] Y. Yildirim and M. Bayramoglu, "Adaptive neuro-fuzzy based modelling for prediction of air pollution daily levels in city of Zonguldak", *Chemosphere*, vol. 63, pp. 1575-1582, 2006.
- [YKSS95] B. Yuan, G.J. Klir, and J.F. Swan-Stone, "Evolutionary fuzzy c-means clustering algorithm," in *Proc. 4th IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE'95)*, Yokohama, Japan, 1995, pp. 2221-2226.
- [Zad65] L. A. Zadeh, "Fuzzy Sets," *Inform. Control*, vol. 8, no. 3, pp. 338-353, June 1965.
- [Zad73] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. Syst., Man, and Cybern.*, vol. 3, no. 1, pp. 28-44, 1973.
- [Zad94] L. A. Zadeh, "Fuzzy Logic, Neural Neural Networks and Soft Computing ," *Communications of the ACM.*, vol. 37, no. 3, pp. 77-84, 1994.
- [ZDT00] E. Zitzler, K. Deb and L. Thiele, "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results," *Evol. Comput.*, vol. 8, no. 2, pp. 173-195, 2000.
- [Zha95] J. Zhao, "Fuzzy logic in modeling and control," Ph.D. dissertation, CESAME, Louvain la Neuve, Belgium, 1995.
- [Zim96] H. J. Zimmermann, *Fuzzy Sets Theory and its Applications*. Kluwer Academic Press, 1996
- [ZLT01] E. Zitzler, M. Laumanns and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm," Technical Report 103, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 2001.
- [ZLV00] J.B. Zydallis, G.B. Lamont and D.A. Van Veldhuizen, "Messy Genetic Algorithm Based Multi-Objective Optimization: A Comparative Statistical Analysis," in *PPSN/SAB Workshop on Multiobjective*

- Problem Solving from Nature (MPSN)*, Paris, France, 2000, pp. 1-3.  
Available: <http://neo.lcc.uma.es/emoo/EMOOconferences.html>
- [ZT98] E. Zitzler and L. Thiele, "An Evolutionary Algorithm for Multiobjective Optimization: The Strength Pareto Approach," Technical Report 43, Computer Engineering and Communication Networks Lab (TIK), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 1998.
- [ZT99] E. Zitzler and L. Thiele, "Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257-271, Nov. 1999.
- [ZTB99] E. Zitzler, J. Teich, and S.S. Bhattacharyya, "Evolutionary Algorithm Based Exploration of Software Schedules for Digital Signal Processors," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'99)-Vol. 2*, W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela, and R. E. Smith, Eds. Morgan Kaufmann, 1999, pp. 1762-1769.
- [ZVL01] J.B. Zydallis, D.A. Van Veldhuizen and G.B. Lamont, "A Statistical Comparison of Multiobjective Evolutionary Algorithms Including the MOMGA-II," in *First International Conference on Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C.A. Coello and D. Corne, Eds. Springer-Verlag, Lecture Notes in Computer Science No. 1993, 2001, pp. 226-240.

