

SISTEMA DE PREDICCIÓN DE RUIDO URBANO MEDIANTE REDES NEURONALES

Tesis Doctoral

Natalia Genaro García

**Programa Oficial de Doctorado en Tecnologías de
la Información y la Comunicación**



Universidad de Granada



Dirección:

Dr. Ignacio Requena Ramos

Dra. Montserrat Zamorano Toro

Editor: Editorial de la Universidad de Granada
Autor: Natalia Genaro García
D.L.: GR 2076-2014
ISBN: 978-84-9803-250-9

A URBAN NOISE PREDICTION SYSTEM BASED ON NEURAL NETWORKS

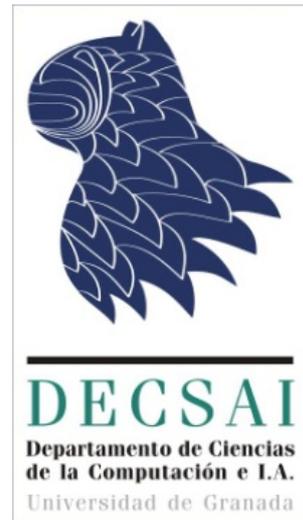
PhD Dissertation

Natalia Genaro García

Official PhD Program on Information and
Communication Technologies



Universidad de Granada



Supervisors:

Dr. Ignacio Requena Ramos

Dra. Montserrat Zamorano Toro

La doctoranda Natalia Genaro García y los directores de la tesis Ignacio Requena Ramos y Montserrat Zamorano Toro garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 27 de junio de 2014

Director/es de la Tesis



Fdo.: Ignacio Requena Ramos



Fdo.: Montserrat Zamorano Toro

Doctoranda



Fdo.: Natalia Genaro García

AGRADECIMIENTOS

Gracias Ignacio por tu apoyo, por animarme cuando lo veía todo negro, y sobre todo, por no meterme prisa justo cuando era eso lo que necesitaba.

Gracias Montse por tus consejos, no sólo los científicos.

Gracias a mis compañeros de trabajo del CSIRC, en especial a mis compañeros del Servicio de Redes y Comunicaciones. Entre unas redes y las otras, casi no daba abasto.

Gracias a mis amigos, que valoráis el esfuerzo que supone hacer una tesis y decirme “no sé cómo te has metido en ese lío teniendo la vida resuelta”. Al final lo he conseguido. Gracias Dany, M^a Mar, Sera, Viki, Virginia, Albert, Borja, Nuria, María, Nandish, M^aPaz, Pablo, Sandra, Elena, Manu y los sobris: Martín, Nacho y Nico.

Gracias especiales a mis amigos Miku y Álvaro, que ya habéis “sufrido” la tesis y me dais buenos consejos para salir victoriosa de esto. Me encanta que aunque estéis lejos forméis parte también de mi día a día.

Gracias mi familia política, Carlos, M^a Belén y Blanca, por tratarme como una más de la familia.

Gracias a mis padres, porque decirme que “mi trabajo es estudiar” me ha llevado hasta aquí. Aún sigo estudiando, y con razón me llamáis “cursillista”. A partir de ahora me dedicaré a algo mejor pagado: ser madre.

Gracias a mi hermana Noemí, a ti y a tu pequeña familia, Alma y Silvio. No sólo por el apoyo logístico, sino también el emocional, que vale más que mil tesis juntas.

Gracias Carlos, por animarme, explicarme, corregirme, pero sobre todo gracias por dármelo todo sin dudarlo y aguantar estoicamente y con cariño las llamadas de Bruno a su mamá mientras terminaba la tesis. No me habría animado a hacerla sin ti, pero tampoco la habría acabado si no siguieras a mi lado.

Gracias Bruno, porque tan pequeño como eres ya me has enseñado lo que es de verdad el amor incondicional.

ÍNDICE DE CONTENIDO

RESUMEN	xviii
ABSTRACT	xx
INTRODUCCIÓN	1
Capítulo I: Herramientas	9
1.1 El Sonido y el Ruido	12
Parámetros establecidos para la medición.....	15
Elementos involucrados en la propagación del sonido	16
1.2 Modelos matemáticos para la predicción de ruido urbano	21
Modelo lineal.....	21
Modelo lineal multivariante [13].....	21
Modelo MOPU o ISO [14].....	22
Modelo Francés [16]	24
Modelo CORTN [17]	26
Modelo RLS 90 [18].....	26
Modelo Plankwerk '96 [19].....	27
Modelo E. Gaja [20].....	28
Modelo lineal multivariante granadino [9]	29
1.3 Redes Neuronales Artificiales y otras técnicas Soft Computing.....	36
Redes Neuronales Artificiales	36
Otras Técnicas Soft Computing	38
1.4 Extracción y Selección de Características.....	39
Introducción	39
Relevancia de las características	41
Métodos de Extracción y Selección de características.....	43
Métodos de búsqueda usados en selección de características.....	49
1.5 Software	52
Weka.....	52
Matlab	53
mRMR.....	53
Capítulo II: Taxonomía de las Aplicaciones de Técnicas Soft Computing en el Ruido Ambiental	55
2.1 Soft Computing y Análisis del Ruido.....	58

2.2 Redes Neuronales Artificiales y Ruido Urbano	60
A. Predicción del nivel de ruido.....	60
B. Predicción del grado de molestia	74
C. Clasificación.....	77
D. Predicción del flujo de tráfico.....	79
2.3 Ruido y otras técnicas de Soft Computing.....	89
A. Técnicas Difusas	90
B. Algoritmos Genéticos.....	101
C. Modelos Ocultos de Markov.....	103
Capítulo III: Redes Neuronales para la Predicción del Ruido Urbano	109
3.1 Proceso de obtención de datos	112
Herramientas	112
Tiempo e intervalos.....	113
Variables que se miden.....	114
Puntos de medición.....	115
Datos.....	122
3.2 Red Neuronal Artificial para la predicción de niveles sonoros	122
Pruebas iniciales.....	124
3.2 Resultados de la red neuronal propuesta.....	127
3.4 Entrenamiento con 9 neuronas ocultas	134
3.5 Comparativa entre la red neuronal y los modelos predictivos existentes.....	134
Capítulo IV: Selección y Extracción de Características.....	139
4.1 Extracción de características: Análisis de Componentes Principales.....	142
4.2 Selección de características.....	146
A. Métodos que clasifican variables de forma individual	147
1. Relief.....	147
2. Evaluador de atributos según la ganancia de información.....	151
B. Métodos que clasifican subconjuntos de variables: tipo filtro	154
1. Selección de subconjuntos basado en la correlación	154
2. Evaluación de la consistencia de los subconjuntos	157
3. Método de mínima redundancia - máxima relevancia	161
C. Métodos que clasifican subconjuntos de variables: tipo wrapper..	166
1. Tipo Wrapper con conjunto de entrenamiento	166
2. Tipo Wrapper con validación cruzada	167

4.3 Resumen de resultados	173
Capítulo V: Validación del modelo con nuevos conjuntos de datos	183
5.1 Red Neuronal Artificial para la totalidad de los datos disponibles.....	187
5.2 Selección de características con el nuevo conjunto de datos	191
MÉTODO 1: Relief-f.....	191
MÉTODO 2: Selección de subconjuntos de características basadas en correlación.....	194
MÉTODO 3: Wrapper con Validación cruzada	196
5.3 Resumen de resultados	199
Capítulo VI: Conclusiones y Trabajo Futuro.....	205
6.1 Conclusiones.....	207
6.2 Trabajos Futuros.....	210
6.3 Publicaciones Relacionadas	211
Bibliografía	212

Índice de tablas

TABLA 1. LOS NIVELES DE RUIDO PERMITIDOS EN LA CIUDAD DE GRANADA.....	5
TABLA 2. VARIACIÓN DE LA CONSTANTE α SEGÚN EL TIPO DE PAVIMENTO	22
TABLA 3. VALOR DE E SEGÚN LA PENDIENTE	24
TABLA 4. COMPARATIVA DE LOS MODELOS PREDICTIVOS DE RUIDO URBANO	31
TABLA 5. TAXONOMÍA DE LAS PUBLICACIONES ENCONTRADAS	60
TABLA 6. RESULTADOS DE LA PRIMERA PRUEBA DE [81].....	61
TABLA 7. RESULTADOS DE LA SEGUNDA PRUEBA DE [81].....	62
TABLA 8. INDICADORES DE LOS RESULTADOS PARA TRÁFICO ININTERRUMPIDO EN [83].....	67
TABLA 9. INDICADORES DE LOS RESULTADOS PARA TRÁFICO INTERRUMPIDO EN [83].....	68
TABLA 10. RESULTADOS DE LA RED NEURONAL DE [84].....	68
TABLA 11. TEST ESTADÍSTICOS PARA COMPARAR LA RED NEURONAL Y CoRTN EN [84]	68
TABLA 12. VARIABLES DE ENTRADA DE LA RED NEURONAL DE [85]	70
TABLA 13. RESULTADOS DE LA EJECUCIÓN DE LA RED NEURONAL DE [87]	73
TABLA 14. COMPARATIVA DE LOS TRES MODELOS PREDICTIVOS USADOS EN [88]	73
TABLA 15. RESULTADOS DE LA RED NEURONAL DE PROGRAMACIÓN HACIA ATRÁS EN [88]	73
TABLA 16. RESULTADOS DE LA RED NEURONAL DE REGRESIÓN GENERAL EN [88]	74
TABLA 17. RESULTADOS DE LA PREDICCIÓN DEL MODELO DE REGRESIÓN LINEAL DE [88].....	74
TABLA 18. FACTORES MEDIDOS EN LA TOMA DE DATOS DE [93].....	76
TABLA 19. COMPORTAMIENTO DE LA RED NEURONAL DE [93] EN LA PREDICCIÓN DEL CONFORT ACÚSTICO.....	77
TABLA 20. PORCENTAJE DE SONIDOS MAL CLASIFICADOS PARA LOS 3 EXPERIMENTOS LLEVADOS A CABO EN [94]	78
TABLA 21. RESUMEN DE RESULTADOS DE [104]	86
TABLA 22. RESULTADOS DE LOS DISTINTOS MODELOS PREDICTIVOS DE [106]	88
TABLA 23. RESULTADOS COMPARATIVOS DEL MODELO DIFUSO DE [111] CON UN MODELO MATEMÁTICO	90
TABLA 24. RESULTADOS DEL SISTEMA POPFNN-TV R COMPARADOS CON FFBP DE [108]	100
TABLA 25. COMPARACIÓN DE RESULTADOS DE LOS DISTINTOS MODELOS DE [132].....	102
TABLA 26. PRECISIÓN DEL MODELO PROPUESTO EN [136] Y [137] SEGÚN EL TIPO DE SONIDO.....	105
TABLA 27. VARIABLES DE ENTRADA CONSIDERADAS POR LA RED.....	115
TABLA 28. VARIABLES DE ENTRADA CONSIDERADAS POR LA RED Y SU NORMALIZACIÓN.....	123
TABLA 29. RESULTADOS DE LAS 12 EJECUCIONES PARA ELEGIR UNA ESTRUCTURA DE RED DEFINITIVA	125
TABLA 30. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO DE DATOS 1	128
TABLA 31. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO DE DATOS 2	129
TABLA 32. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO DE DATOS 3	129
TABLA 33. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO DE DATOS 4	129
TABLA 34. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO DE DATOS 5	129
TABLA 35. RESULTADOS DE LA RED DEL CONJUNTO DE TEST 3.....	131
TABLA 36. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED NEURONAL CON 9 NEURONAS EN LA CAPA INTERMEDIA.....	134
TABLA 37. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO 1 (CON ACP)	143
TABLA 38. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO 2 (CON ACP)	143
TABLA 39. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO 3 (CON ACP)	143
TABLA 40. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO 4 (CON ACP)	144
TABLA 41. RESULTADOS DE LAS CINCO EJECUCIONES DE LA RED CON EL CONJUNTO 5 (CON ACP)	144
TABLA 42. RANKING DE VARIABLES PARA EL MÉTODO RELIEF-F	147
TABLA 43. ORDENACIÓN DE LAS VARIABLES SEGÚN EL MÉTODO RELIEF-F	148
TABLA 44. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (RELIEF-F)	149
TABLA 45. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (RELIEF-F)	149
TABLA 46. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (RELIEF-F)	149
TABLA 47. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (RELIEF-F)	150
TABLA 48. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (RELIEF-F)	150
TABLA 49. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (RELIEF-F)	150
TABLA 50. CLASES AL DISCRETIZAR L_{AEQ}	151
TABLA 51. RANKING DE VARIABLES SEGÚN EL CÁLCULO DE LA GANANCIA DE INFORMACIÓN.....	152
TABLA 52. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (INFOGAIN)	152
TABLA 53. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (INFOGAIN)	153
TABLA 54. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (INFOGAIN).....	153

TABLA 55. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (INFOGAIN)	153
TABLA 56. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (INFOGAIN)	153
TABLA 57. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (INFOGAIN)	154
TABLA 58. RESULTADOS PARA EL MÉTODO DE SELECCIÓN DE CARACTERÍSTICAS BASADO EN LA CORRELACIÓN	155
TABLA 59. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 9 VARIABLES (CORRELACIÓN)	155
TABLA 60. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 9 VARIABLES (CORRELACIÓN)	156
TABLA 61. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (CORRELACIÓN)	156
TABLA 62. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (CORRELACIÓN)	156
TABLA 63. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 11 VARIABLES (CORRELACIÓN)	156
TABLA 64. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 11 VARIABLES (CORRELACIÓN)	157
TABLA 65. RESULTADOS PARA EL MÉTODO DE SELECCIÓN DE CARACTERÍSTICAS SEGÚN LA CONSISTENCIA	158
TABLA 66. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 9 VARIABLES (CONSISTENCIA)	158
TABLA 67. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 9 VARIABLES (CONSISTENCIA)	159
TABLA 68. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (CONSISTENCIA)	159
TABLA 69. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (CONSISTENCIA)	159
TABLA 70. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (CONSISTENCIA)	159
TABLA 71. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (CONSISTENCIA)	160
TABLA 72. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 11 VARIABLES (CONSISTENCIA)	160
TABLA 73. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 11 VARIABLES (CONSISTENCIA)	160
TABLA 74. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 14 VARIABLES (CONSISTENCIA)	160
TABLA 75. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 14 VARIABLES (CONSISTENCIA)	161
TABLA 76. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (MRMR - MID)	162
TABLA 77. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (MRMR - MID)	162
TABLA 78. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (MRMR - MID)	163
TABLA 79. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (MRMR - MID)	163
TABLA 80. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (MRMR - MID)	163
TABLA 81. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (MRMR - MID)	163
TABLA 82. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (MRMR - MIQ)	165
TABLA 83. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 10 VARIABLES (MRMR - MIQ)	165
TABLA 84. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (MRMR - MIQ)	165
TABLA 85. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 12 VARIABLES (MRMR - MIQ)	165
TABLA 86. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (MRMR - MIQ)	166
TABLA 87. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 15 VARIABLES (MRMR - MIQ)	166
TABLA 88. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 17 VARIABLES (WRAPPER – TRAIN)	167
TABLA 89. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 17 VARIABLES (WRAPPER - TRAIN)	167
TABLA 90. RESULTADOS DEL MÉTODO WRAPPER CON VALIDACIÓN CRUZADA (10 ITERACIONES)	168
TABLA 91. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 7 VARIABLES (WRAPPER – CROSS10)	168
TABLA 92. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1, 7 VARIABLES (WRAPPER – CROSS10)	168
TABLA 93. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,11 VARIABLES(WRAPPER – CROSS10)	169
TABLA 94. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,11 VARIABLES(WRAPPER – CROSS10)	169
TABLA 95. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,19 VARIABLES(WRAPPER – CROSS10)	169
TABLA 96. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,19 VARIABLES(WRAPPER – CROSS10)	170
TABLA 97. RESULTADOS DEL MÉTODO WRAPPER CON VALIDACIÓN CRUZADA (25 ITERACIONES)	171
TABLA 98. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,10 VARIABLES(WRAPPER – CROSS25)	171
TABLA 99. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,10 VARIABLES(WRAPPER – CROSS25)	171
TABLA 100. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,13 VARIABLES(WRAPPER – CROSS25)	172
TABLA 101. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,13 VARIABLES(WRAPPER – CROSS25)	172
TABLA 102. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,15 VARIABLES(WRAPPER – CROSS25)	172
TABLA 103. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,15 VARIABLES(WRAPPER – CROSS25)	172
TABLA 104. RESUMEN RESULTADOS DE LOS MÉTODOS DE SELECCIÓN DE CARACTERÍSTICAS EVALUADOS	174
TABLA 105. RESUMEN RESULTADOS DE LOS MÉTODOS DE SELECCIÓN DE CARACTERÍSTICAS EVALUADOS	175
TABLA 107. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,10 VARIABLES (SELECCIÓN PROPIA)	176
TABLA 108. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 2,10 VARIABLES (SELECCIÓN PROPIA)	176
TABLA 109. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 3,10 VARIABLES (SELECCIÓN PROPIA)	177
TABLA 110. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 4,10 VARIABLES (SELECCIÓN PROPIA)	177
TABLA 111. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 5,10 VARIABLES (SELECCIÓN PROPIA)	177

TABLA 112. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,12 VARIABLES (SELECCIÓN PROPIA)	177
TABLA 113. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 2,12 VARIABLES (SELECCIÓN PROPIA)	178
TABLA 114. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 3,12 VARIABLES (SELECCIÓN PROPIA)	178
TABLA 115. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 4,12 VARIABLES (SELECCIÓN PROPIA)	178
TABLA 116. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 5,12 VARIABLES (SELECCIÓN PROPIA)	178
TABLA 117. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 1,15 VARIABLES (SELECCIÓN PROPIA)	179
TABLA 118. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 2,15 VARIABLES (SELECCIÓN PROPIA)	179
TABLA 119. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 3,15 VARIABLES (SELECCIÓN PROPIA)	179
TABLA 120. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 4,15 VARIABLES (SELECCIÓN PROPIA)	179
TABLA 121. RESULTADOS DE LAS 5 EJECUCIONES DE LA RED CON EL CONJUNTO 5,15 VARIABLES (SELECCIÓN PROPIA)	180
TABLA 122. TABLA DE VARIABLES CON VARIABLES ELEGIDAS EN NEGRITA	181
TABLA 123. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, PARA LOS CINCO CONJUNTOS, 7HN	187
TABLA 124. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, PARA LOS CINCO CONJUNTOS, 9HN	187
TABLA 125. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, PARA LOS CINCO CONJUNTOS, 12HN	188
TABLA 126. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, PARA LOS CINCO CONJUNTOS, 15HN	188
TABLA 127. RESULTADOS DE RELIEF-F EN LOS DATOS INICIALES Y NUEVOS SEGÚN EL NÚMERO DE VECINOS	191
TABLA 128. RANKING DE VARIABLES SEGÚN RELIEF-F	192
TABLA 129. COMPARATIVA DE LA APLICACIÓN DE RELIEF-F SEGÚN LOS DATOS	192
TABLA 130. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 10 VARIABLES (RELIEF)	193
TABLA 131. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 12 VARIABLES (RELIEF)	193
TABLA 132. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 15 VARIABLES (RELIEF)	193
TABLA 133. VARIABLES SELECCIONADAS SEGÚN EL MÉTODO DE BÚSQUEDA PARA LOS DATOS NUEVOS Y ANTIGUOS	195
TABLA 134. COMPARATIVA DE LA APLICACIÓN DEL MÉTODO DE CORRELACIÓN SEGÚN LOS DATOS	195
TABLA 135. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 9 VARIABLES (CORRELACIÓN)	195
TABLA 136. DATOS INICIALES Y NUEVOS	196
TABLA 137. DATOS INICIALES	196
TABLA 138. COMPARATIVA DE VARIABLES ELEGIDAS SEGÚN EL CONJUNTO DE DATOS	197
TABLA 139. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 11 VARIABLES (WRAPPER-CROSS 10)	197
TABLA 140. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 16 VARIABLES (WRAPPER-CROSS 10)	197
TABLA 141. COMPARATIVA DE VARIABLES ELEGIDAS SEGÚN EL CONJUNTO DE DATOS	198
TABLA 142. DATOS INICIALES Y NUEVOS	198
TABLA 143. DATOS INICIALES	198
TABLA 144. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 12 VARIABLES (WRAPPER-CROSS 25)	199
TABLA 145. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED NEURONAL, 16 VARIABLES (WRAPPER-CROSS 25)	199
TABLA 146. RESUMEN DE RESULTADOS DE SELECCIÓN DE CARACTERÍSTICAS CON NUEVO CONJUNTO DE DATOS	201
TABLA 147. RESUMEN DE VARIABLES ELEGIDAS TRAS LA SELECCIÓN DE CARACTERÍSTICAS	202
TABLA 148. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, 12HN. 10 VARIABLES (SELECCIÓN PROPIA)	203
TABLA 149. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, 12HN. 12 VARIABLES (SELECCIÓN PROPIA)	203
TABLA 150. RESULTADOS DE LAS SALIDAS MEDIAS DE LA RED, 12HN. 15 VARIABLES (SELECCIÓN PROPIA)	203

Índice de Figuras

FIGURA 1. RANGO DE AUDICIÓN.....	13
FIGURA 2. CURVA DE FLETCHER Y MUNSON.....	14
FIGURA 3. CURVAS DE PONDERACIÓN A, B Y C.....	15
FIGURA 4. ATENUACIÓN ATMOSFÉRICA SEGÚN DISTANCIA A LA FUENTE [11].....	17
FIGURA 5. ESTRUCTURA DE UNA RED NEURONAL FEEDFORWARD.....	37
FIGURA 6. REPRESENTACIÓN DEL ESPACIO DE CARACTERÍSTICAS SEGÚN EL TIPO.	42
FIGURA 7. ESTRUCTURA DEL MODELO PMC DE [81].....	61
FIGURA 8. RESULTADOS DE LA RED NEURONAL DE [81].....	62
FIGURA 9. ARQUITECTURA FINAL DEL MODELO [39].....	63
FIGURA 10. CLASIFICACIÓN DEL FILTRO LVQ DE [39].....	64
FIGURA 11. RESULTADOS DEL CONJUNTO DE ENTRENAMIENTO DE [82].....	65
FIGURA 12. RESULTADOS DE CONJUNTO DE TEST [82].....	65
FIGURA 13. COMPARACIÓN NIVEL SONORO MEDIDO - CALCULADO DE [82].....	65
FIGURA 14. DISTRIBUCIÓN DE LOS DATOS DE [82].....	66
FIGURA 15. COMPARACIÓN DE VALOR MEDIDO Y CALCULADO PARA LNP EN TRÁFICO ININTERRUMPIDO EN [83].....	66
FIGURA 16. COMPARACIÓN DE VALOR MEDIDO Y CALCULADO PARA LNP EN TRÁFICO INTERRUMPIDO EN [83].....	67
FIGURA 17. RED NEURONAL UTILIZADA EN [84].....	69
FIGURA 18. RESULTADOS DE PREDICCIÓN DE LEQ EN [85].....	71
FIGURA 19. EVOLUCIÓN DEL ERROR CUADRÁTICO MEDIO EN EL ENTRENAMIENTO DE LA RED NEURONAL [86].....	71
FIGURA 20. COMPARACIÓN DEL L_{Aeq} MEDIDO Y ESTIMADO POR LOS DIFERENTES MODELOS CLÁSICOS EN [87].	72
FIGURA 21. ESTRUCTURA DEL MODELO NEURO DIFUSO DE [91].....	75
FIGURA 22. ESQUEMA DEL SISTEMA AUTOMÁTICO DE RECONOCIMIENTO DE RUIDO [96].....	78
FIGURA 23. ESTRUCTURA DEL MODELO NEURO-DIFUSO DE [98].....	80
FIGURA 24. ESTRUCTURA DE LA RED DE [100].....	82
FIGURA 25. VELOCIDAD PREDICHA POR LA RED CON 20 SEG DE ANTELACIÓN [101].....	82
FIGURA 26. VELOCIDAD PREDICHA POR LA RED CON 60 SEG DE ANTELACIÓN [101].....	83
FIGURA 27. RED NEURONAL PARA PREDICCIÓN DE LA DURACIÓN DEL TRAYECTO DE [101].....	84
FIGURA 28. ESTRUCTURA DEL MODELO NEURO-DIFUSO DE [102].....	84
FIGURA 29. FRECUENCIA DE REDES ÓPTIMAS SEGÚN EL HORIZONTE DE PREDICCIÓN EN MINUTOS BASADO EN EL ERROR RELATIVO MEDIO ABSOLUTO [103].....	86
FIGURA 30. MODELO DE RED NEURONAL DE ONDÍCULA DE [105].....	87
FIGURA 31. REDES NEURONALES PARA LA PREDICCIÓN DEL FLUJO DE TRÁFICO EN [109].....	88
FIGURA 32. FLUJO DE DESARROLLO DE LA RED NEURONAL CON SUAVIZADO EXPONENCIAL EN [110].....	89
FIGURA 33. MODELO DE PREDICCIÓN DEL GRADO DE MOLESTIA [126] [121] [117].....	92
FIGURA 34. DISTRIBUCIÓN DE LOS SUJETOS [126].....	94
FIGURA 35. DISTRIBUCIÓN DE LOS SUJETOS [126].....	94
FIGURA 36. MODELO DIFUSO DE AGREGACIÓN [119] [123].....	95
FIGURA 37. FUNCIONES DE PERTENENCIA DE LOS CUATRO TÉRMINOS LINGÜÍSTICOS INDICADOS EN [125] Y [120].....	96
FIGURA 38. ARQUITECTURA DE CLASIFICACIÓN MULTINIVEL DE [128].....	97
FIGURA 39. DIAGRAMA DE BLOQUES DEL MODELO DIFUSO DE [129].....	98
FIGURA 40. PROCESO DE COMBINADO DE MODELOS PREDICTIVOS CON UN FRBS EN [131].....	99
FIGURA 41. ESTRUCTURA DEL SISTEMA POPFNN-TVR DE [107] [108].....	100
FIGURA 42. GRÁFICA COMPARATIVA PARA EL MODELO DE [113].....	101
FIGURA 43. DIAGRAMA DE FLUJO DEL PROCESO DE OPTIMIZACIÓN MEDIANTE AG EN [133].....	103
FIGURA 44. ESTRUCTURA DEL SISTEMA DE [134] Y [135].....	104
FIGURA 45. EVOLUCIÓN DE LA PRECISIÓN SEGÚN EL NÚMERO DE ESTADOS EN [136] Y [137].....	104
FIGURA 46. COMPARATIVA DEL PRIMER MÉTODO DE [97].....	106
FIGURA 47. ESTRUCTURA DEL SISTEMA MoE DE [97].....	107
FIGURA 48. COMPARATIVA DEL SEGUNDO MÉTODO DE [97].....	107
FIGURA 49. CAMINO DE RONDA.....	116
FIGURA 50. GRAN VÍA.....	116
FIGURA 51. AVENIDA DE MURCIA.....	117
FIGURA 52. MESONES.....	117
FIGURA 53. MÉNDEZ NÚÑEZ.....	118

FIGURA 54. REYES CATÓLICOS	118
FIGURA 55. CAMINO DE LAS VACAS	119
FIGURA 56. DOCTOR OLÓRIZ.....	119
FIGURA 57. REAL DE CARTUJA	120
FIGURA 58. GRAN CAPITÁN	120
FIGURA 59. GONZALO GALLAS.....	121
FIGURA 60. NUEVA DEL SANTÍSIMO	121
FIGURA 61. GRÁFICA COMPARATIVA DEL CONJUNTO ENTRENAMIENTO, RED 4, EJECUCIÓN 2	126
FIGURA 62. GRÁFICA COMPARATIVA DEL CONJUNTO TEST, RED 4, EJECUCIÓN 2	127
FIGURA 63. ESTRUCTURA DE LA RED NEURONAL PROPUESTA.....	127
FIGURA 64. GRÁFICA COMPARATIVA DEL CONJUNTO ENTRENAMIENTO 3, EJECUCIÓN 4.....	130
FIGURA 65. GRÁFICA COMPARATIVA DEL CONJUNTO TEST 3, EJECUCIÓN 4.....	131
FIGURA 66. GRÁFICA COMPARATIVA DE LOS RESULTADOS DE LOS MODELOS CLÁSICOS Y LA RED NEURONAL CON EL VALOR L_{AEQ} MEDIDO PARA EL CONJUNTO DE DATOS 1	136
FIGURA 67. GRÁFICA COMPARATIVA DE LOS RESULTADOS DE LOS MODELOS CLÁSICOS Y LA RED NEURONAL CON EL VALOR L_{AEQ} MEDIDO PARA EL CONJUNTO DE DATOS 1	137
FIGURA 68. GRÁFICA COMPARATIVA RESULTADOS CONJUNTO ENTRENAMIENTO 5, EJECUCIÓN 1 (ACP).....	145
FIGURA 69. GRÁFICA COMPARATIVA RESULTADOS CONJUNTO TEST 5, EJECUCIÓN 1 (ACP)	146
FIGURA 70. GRÁFICA COMPARATIVA DE LOS RESULTADOS DE LOS MODELOS CLÁSICOS Y LA RED NEURONAL CON EL VALOR L_{AEQ} MEDIDO PARA EL CONJUNTO DE DATOS 1 (A)	189
FIGURA 71. GRÁFICA COMPARATIVA DE LOS RESULTADOS DE LOS MODELOS CLÁSICOS Y LA RED NEURONAL CON EL VALOR L_{AEQ} MEDIDO PARA EL CONJUNTO DE DATOS 1 (B)	190

RESUMEN

El ruido es un problema global. En 1972 la Organización Mundial de la Salud (OMS) clasificó el ruido como un contaminante. Desde entonces, la mayoría de los países industrializados han promulgado leyes y reglamentos locales para prevenir y reducir la contaminación ambiental acústica. Otro objetivo es alertar a la gente de los peligros de este tipo de contaminación. En este contexto, los planificadores urbanos necesitan herramientas que les permitan evaluar el grado de contaminación acústica. Los científicos de muchos países han modelado el ruido urbano, utilizando una amplia gama de enfoques, pero sus resultados no han sido tan buenos como se esperaba.

Actualmente existen una serie de modelos predictivos de ruido urbano basados en métodos matemáticos tradicionales, que predicen el nivel de ruido a partir de un reducido número de variables. Tras el estudio del estado del arte de literatura científica que relaciona métodos de Soft Computing y predicción de ruido urbano, se ha evidenciado que la mayoría de estos estudios no obtienen unos resultados mucho mejores que los obtenidos con modelos matemáticos tradicionales.

Esta tesis desarrolla un modelo para la predicción de ruido ambiental urbano mediante Redes Neuronales Artificiales (RNA). El modelo se basa en el análisis de las variables consideradas como influyentes por expertos en la materia y se aplica a los datos recogidos en diferentes tipos de calles. Para ello, en colaboración con expertos en ruido de la UGR, se han obtenido una serie de medidas del ruido urbano en la ciudad de Granada, que permiten el entrenamiento de la RNA que se propone.

Los dos grandes objetivos de la tesis son:

- **Estudio del estado del arte de literatura científica relacionando métodos de Soft Computing y predicción de ruido urbano.**

El estudio del estado del arte ha hecho patente que la mayoría de los estudios no obtienen unos resultados mucho mejores que los obtenidos con modelos matemáticos. Por ello, la aplicación de métodos de Soft Computing que permitan predecir con más exactitud el nivel de ruido ambiental puede ser una solución. En esta tesis se concluye, además, que no hay muchos artículos prediciendo el ruido urbano con redes neuronales, y la mayoría de ellos usan modelos de redes neuronales similares a las de este trabajo. No obstante, todos ellos utilizan un conjunto de datos de entrada menos completo, en número y naturaleza de las variables, que el de este trabajo y todos los que obtienen resultados más parecidos lo hacen en fechas posteriores y citando, en la mayoría de las ocasiones, las publicaciones de los resultados explicitados en esta memoria.

- **Desarrollo de un modelo de predicción de ruido urbano basado en redes neuronales preciso para modelar los niveles de presión sonora.**

Este segundo objetivo tiene a su vez varios subobjetivos que se desarrollan en siguientes capítulos:

En primer lugar, construir un modelo de predicción válido en la ciudad de Granada, y que se pueda extender fácilmente a otras ciudades. Se ha desarrollado una RNA, con 25 entradas o variables a ser tenidas en cuenta, y el nivel equivalente de ruido como salida, entrenada con 289 registros obtenidos en Granada. La única capa oculta tiene 7 neuronas (también con 9 se han obtenido resultados similares). Se obtiene así un modelo que predice los niveles de ruido urbano de Granada, con excepcional eficacia, ya que casi ningún ejemplo supera el 5% del error. El entrenamiento es rápido, lo que permite usar la red en tiempo casi real.

En segundo lugar, se ha pretendido reducir el número de entradas, seleccionando las variables más relevantes. Un análisis de Componentes Principales ha permitido saber el número mínimo de variables a considerar, y varios métodos de selección de características han permitido obtener varios conjuntos incrementales de variables con resultados razonablemente eficaces. Los mejores resultados se han obtenido con los métodos RELIEF y el de evaluación de subconjuntos de variables por consistencia. Finalmente se seleccionan tres conjuntos (incrementales) con 10, 12 y 15 variables y se entrena de nuevo la RNA con ellos. Los mejores resultados se obtienen con 15 variables, estando los resultados casi al mismo nivel que los obtenidos con las 25 variables, y superando ampliamente a los métodos clásicos con los 3 conjuntos de variables seleccionados.

Finalmente, y para descartar que el éxito se pueda haber obtenido por un posible exceso de homogeneidad en los datos de partida, también se han utilizado datos, también de Granada, pero que no fueron obtenidos para el presente trabajo, llegando así, tras un adecuado preproceso, a un total de 640 registros. Esto ha permitido validar el sistema desarrollado, construyendo una nueva RNA (25 entradas, 1 salida y 12 neuronas ocultas) entrenada con estos nuevos datos. Los resultados siguen siendo muy buenos, aunque no tan espectaculares como los obtenidos previamente, pero amplia y claramente mejores que los obtenidos por los modelos matemáticos. También se lleva a cabo un proceso de Selección de Características, seleccionando finalmente un conjunto con 15 variables, obteniendo unos resultados mejores que los modelos matemáticos “clásicos” y suficientemente buenos para la predicción del ruido urbano.

ABSTRACT

Noise is a global problem. In 1972 the World Health Organization (WHO) classified noise as a pollutant. Since then, most industrialized countries have enacted laws or local regulations that regulate noise levels to reduce noise pollution and to alert people on the dangers of this kind of pollution. Urban planners have a need for tools that allow them assess noise levels. Scientists in many countries have modeled urban noise, using a wide range of approaches. However, their results have not been as good as expected.

Currently there are a few models to predict urban noise and try to mathematically calculate the noise level based on a given set of characteristics. This study of the previous works relating Soft Computing and urban noise prediction made clear that most of them don't get better results than those obtained with classical mathematical models.

This thesis develops an artificial neural network (ANN) to create a urban noise prediction system based on the analysis of influential characteristics. To that end, and in collaboration with urban noise experts in the University of Granada, different measurements of urban noise in the city of Granada were collected. The ANN was trained with these data.

The two main goals of this thesis are:

- **Review of the literature**

A review of the literature of the previous works describing urban noise prediction models with Soft Computing methods showed that most of them don't obtain better results than those obtained with classical mathematical models. Therefore the application of Soft Computing methods to predict more accurately environmental noise could lead to a solution. This thesis concludes that there are not many papers relating urban noise prediction and ANN, and most of them use similar models to the one developed in this work. However, all of them use a narrower range of data and obtain similar results, always after this work was published. Most of them also cite this work.

- **A neural network for urban noise prediction**

This goal has at the same time three subgoals:

The first subgoal was the development of a prediction model that could be used not only in the city of Granada, but also easily applicable to other cities. 25 characteristics constituted the input of the ANN, whereas the equivalent noise level was the output. It was trained and tested with 289 records, all of them collected in Granada. Its one and only hidden layer had 7 neurons (similar results were obtained with 9 neurons too). This model predicts exceptionally accurate noise levels in Granada, as almost no instance exceed 5% of error. Training is fast, so this could be real-time used.

The second goal was the reduction in the number of inputs, selecting the most relevant ones. A Principal Component Analysis allowed knowing how many variables could be considered. After applying different feature selection techniques, reasonably well performance was obtained. Best results were obtained with RELIEF and Consistency-based Subset evaluation. Finally three sets of 10, 12 and 15 variables were selected, and the ANN gave similar results to those given with 25 variables, and much better than those obtained by classical mathematical models.

Finally, to discard that success could have been obtained by an excess of homogeneity in the input data, new data was analyzed. These data were also collected in Granada, but were not obtained specifically for the aim of this study. After an appropriate preprocess a total of 656 records were obtained. This has allowed validating the developed system, building a new RNA (25 inputs, 1 output and 12 hidden neurons) trained with the new data. The results are still very good, though not as stunning as those obtained previously, but wide and clearly better than those obtained by mathematical models. A feature selection process was also carried out, finally selecting a set of 15 variables, obtaining better results than the mathematical models and good enough for the prediction of urban noise.

INTRODUCCIÓN

El hombre ha convivido con el ruido desde el principio de los tiempos. Ya en la Suiza del siglo XV se decretó la prohibición de circulación de carretas en mal estado, con el objetivo de evitar ruidos molestos a viandantes y vecinos. Hasta hace poco tiempo, el ruido era considerado solo un accidente producto de la actividad humana, y con el que el hombre debía convivir, pero desde hace algo más de treinta años, las autoridades han empezado a reaccionar ante este serio problema.

La primera declaración internacional que contempló las consecuencias del ruido se remonta a 1972, cuando la Organización Mundial de la Salud (OMS) decidió catalogarlo genéricamente como un tipo más de contaminación. Siete años después, la Conferencia de Estocolmo clasificaba al ruido como un contaminante específico. Aquellas primeras disposiciones oficiales fueron ratificadas posteriormente por la entonces CEE, que requirió a los países miembros un esfuerzo para regular legalmente la contaminación acústica. Más tarde, un informe publicado en 1990 presentaba a España como el segundo país con mayor índice de ruidos del mundo después de Japón, y estimaba que el 74% de la población estaba sometida a niveles superiores a los tolerables.

En las últimas décadas se percibe un cambio en las condiciones acústicas de las ciudades (incremento del parque móvil, mecanización de las actividades, cambio de usos de la vía pública, etc.), produciéndose un incremento en el nivel de ruido ambiental. Ante la evidencia de un problema mundial, la mayoría de los países del mundo, han puesto o están poniendo en marcha medidas para proteger a sus ciudadanos. Aún así, el problema persiste incluso en los lugares en los que ya se están dedicando recursos, leyes y

normativas para regular, evaluar y amortiguar fuentes de ruido o para la construcción de barreras de ruido.

“El Libro Verde de la lucha contra el Ruido” (1996) [1] elaborado por la Unión Europea, abogó por una serie de medidas entre las que se encuentran la reducción de los valores de emisión autorizados para los vehículos de carretera, el desarrollo de instrumentos económicos para la compra de vehículos silenciosos, la instauración de un control de ruido de vehículos durante la inspección técnica, fomentar el uso de revestimientos de calzada y neumáticos más silenciosos e incluso prohibir la circulación de vehículos ruidosos en determinados períodos que sean más sensibles a la población. El Libro Verde estimaba que, en términos del número de personas afectadas por el ruido, el 20% de la población (unos 80 millones de personas) sufría niveles de ruido inaceptables que causan alteraciones en el sueño, molestias y efectos adversos sobre la salud. Otros 170 millones de ciudadanos vivían, en Europa, en áreas donde los niveles de ruido causan una seria molestia durante el día. En términos financieros, el ruido ambiental cuesta a la sociedad entre un 0.2% y un 2% del Producto Nacional Bruto. Incluso la menor de estas cifras representa un coste inmenso[2].

Éste y otros estudios de la Unión Europea han conducido a la adopción de la Directiva 2002/49/CE del Parlamento Europeo y del Consejo, de 25 de junio de 2002, sobre evaluación y gestión del ruido ambiental[3], que impone a los Estados miembros la obligación de designar las autoridades y entidades competentes para elaborar los mapas de ruido y planes de acción, así como para recopilar la información que se genere, la cual, a su vez, deberá ser transmitida por los Estados miembros a la Comisión y puesta a disposición de la población.

En España, como consecuencia del Libro Verde y de la Directiva de la UE, surge la ley 37/2003 del ruido del 17 de noviembre [4], que sentó las bases contra el ruido ambiental, ya que hasta entonces se carecía de norma reguladora de ámbito estatal, y que tiene como objetivo, prevenir, vigilar y reducir la contaminación acústica, para evitar y reducir los daños que de ésta pueden derivarse para la salud humana, los bienes o el medio ambiente. Entre otras medidas, la ley del ruido no permite conceder nuevas licencias de construcción de edificaciones destinadas a viviendas, usos hospitalarios, educativos o culturales, si los índices de emisión medidos o predichos incumplen los objetivos de calidad acústica que sean de aplicación a las correspondientes áreas acústicas. Dado que esta es una ley muy reciente, está derivando en necesidades inmediatas de sistemas de análisis, predicción y evaluación del ruido ambiental y sus molestias.

Se aprueba en el año 2012 el Reglamento de Protección contra la Contaminación Acústica en Andalucía[5], distinguiendo dos tipos de objetivos de calidad acústica: niveles permitidos para áreas urbanizadas ya existentes y niveles permitidos para áreas de nueva urbanización. Este Decreto, aunque indica expresamente los niveles máximos de ruido por área y período del día, insta a los Ayuntamientos a asignar un tipo de área acústica a cada parte del territorio. También incluye un título específico con los objetivos de calidad acústica en el interior de las edificaciones, estando obligadas las actividades sensibles de producir contaminación acústica o vibraciones a cumplir los valores límite de transmisión

al interior de edificaciones e inmisión al exterior, aunque éstas ya estén constituidas o iniciadas, marcando multas de entre 600€ y 300.000€.

Desde febrero de 2001 se cuenta en la ciudad de Granada con una ordenanza municipal de protección del ambiente acústico[6]. Según dicha ordenanza, el Excmo. Ayuntamiento de Granada califica al ruido como “la mezcla compleja de sonidos con frecuencias fundamentales diferentes. En un sentido amplio puede considerarse ruido cualquier sonido que interfiere en alguna actividad humana”. En la Tabla 1 se relacionan los niveles de ruido permitidos por esta ordenanza en la ciudad de Granada, según zona y periodo del día.

Tabla 1. Los niveles de ruido permitidos en la ciudad de Granada

ZONA	Día (7h-23h)	Noche (23h-7h)
Zona de equipamiento sanitario	60 dbA	50 dbA
Zona con residencia, servicios terciarios, no comerciales o equipamientos no sanitarios	65 dbA	55 dbA
Zona con actividades comerciales	70 dbA	60 dbA
Zona con actividad industrial o servicio urbano excepto servicios de administración	75 dbA	70 dbA

En la Universidad de Granada se lleva algún tiempo realizando un estudio del ruido urbano en la ciudad de Granada, cuya motivación es ser la más ruidosa de las ocho capitales andaluzas, sobretodo, por culpa del tráfico. En concreto, el alto número de ciclomotores que circulan por las calles hace de Granada la segunda ciudad con más decibelios de todos los municipios de más de 50.000 habitantes de Andalucía. Tanto de día como de noche Granada es la más ruidosa, si bien la diferencia con Málaga, la menos ruidosa, es apenas de dos decibelios, tal y como revela un estudio de la Consejería de Medio Ambiente realizado entre 1995 y 1998.

En el análisis del ruido ambiental urbano hay al menos tres aspectos que habitualmente se consideran, el nivel de ruido medido en decibelios, la molestia para los ciudadanos y la clasificación del ruido para identificar la fuente. La construcción de mapas de ruido es un tema importante en el análisis del ruido, y para ello la predicción del nivel de ruido es esencial. En esta línea, la presente memoria se va a centrar en la predicción del nivel de ruido ambiental urbano.

Muchos países y/o ciudades tienen establecidos en su normativa, una forma de predecir el nivel de ruido, generalmente mediante una función matemática, que siempre usa el flujo de tráfico como principal componente, y cuyo resultado se ajusta en función de los valores que algunas variables concretas puedan tener. La eficiencia de estos modelos

matemáticos “clásicos” no es alta, por lo que la búsqueda de nuevas formas de predecir el ruido ambiental está claramente motivada.

Por las características de los datos que los investigadores manejan, y de los resultados que se pretenden obtener, es claro que la aplicación de Técnicas de Soft Computing (fundamentalmente Redes Neuronales y Técnicas Difusas), no solo es conveniente, sino interesante. La capacidad de resolver problemas complejos no tratables con métodos clásicos es una de las razones que invitan a utilizar Soft Computing para el modelado y análisis del ruido urbano lo que además da un abanico de nuevas soluciones y posibilidades de análisis. Otro aspecto importante a considerar, es la posibilidad de obtener una aplicación que permita desarrollar la investigación en este aspecto, hasta llegar, por qué no, a un método válido para muchos países y/o ciudades.

Un primer vistazo a las publicaciones que relacionan técnicas de Soft Computing con el análisis del ruido urbano, pone de manifiesto que el campo está poco estudiado y lo publicado es susceptible de futuras mejoras, convirtiendo estos trabajos en interesantes novedades a nivel internacional. En este contexto se aborda el presente trabajo, y parece interesante hacer una revisión del estado del arte no solo relativo a las redes neuronales, sino también a otras técnicas de Soft Computing que se hayan utilizado para el análisis del ruido urbano. La revisión de la literatura existente hasta el momento en los temas indicados permitirá conocer las carencias que hay en la actualidad, obtener sugerencias e ideas para el trabajo a desarrollar y al mismo tiempo dará la posibilidad de continuar con trabajos inacabados o que no hayan obtenido los resultados esperados.

Por otro lado, los datos obtenidos gracias a la toma de muestras del nivel de ruido en las distintas calles del núcleo urbano invitan a predecir este nivel en situaciones nuevas (p.e., en calles de nueva construcción, sobre todo en áreas que vayan a estar o ya estén pobladas). Disponiendo de datos de las variables relevantes al problema (como pueden ser el potencial flujo de automóviles o el tipo de pavimento de la calle), el fin será obtener el nivel de ruido a soportar por los vecinos de la nueva zona. Para ello, y motivado porque los métodos matemáticos que se utilizan actualmente no son suficientemente eficaces, se usarán, como metodología más adecuada, las Redes Neuronales Artificiales (RNA), que serán entrenadas con los datos disponibles. Se ha optado por esta vía porque actualmente las RNA muestran ser una muy buena solución a problemas de este tipo, dada su capacidad de aprendizaje a partir de ejemplos, la flexibilidad y la tolerancia a fallos (debido a su capacidad de almacenamiento distribuido).

Inicialmente, se pretende considerar un conjunto amplio de variables que puedan tener influencia sobre el nivel de ruido, que serán las entradas a la RNA. Una vez seleccionado el conjunto de variables de entrada, se tomarán datos de ruido en la ciudad de Granada, para tener un adecuado conjunto de ejemplos que permita entrenar a la RNA, y ver si la metodología es suficientemente eficaz. Para ello se contará con la colaboración de expertos en el tema de ruido ambiental.

Posteriormente, si el proceso de aprendizaje con la RNA es suficientemente exitoso, se intentará reducir el número de variables a considerar, eliminando aquellas que,

mediante un proceso de selección de características, puedan considerarse no relevantes (o al menos poco relevantes). Para ello, se aplicarán diferentes métodos de Extracción y Selección de Características lo que permitirá no sólo conocer qué variables son más determinantes e influyentes en el nivel de ruido ambiental, sino que además dará la posibilidad de construir un sistema capaz de predecir dicho nivel con un menor número de variables.

Finalmente, se tratará de validar el sistema desarrollado, con nuevos conjuntos de datos que no se hayan obtenido directamente para este trabajo. Esto permitirá definir la naturaleza de los datos obtenidos en un primer momento, así como dar a la RNA utilizada la capacidad de predecir el nivel de ruido en un mayor número de casos y situaciones.

En consecuencia, los objetivos globales de la presente memoria son dos. Por un lado, la revisión del estado del arte sobre la utilización de técnicas de Soft Computing en el análisis del ruido ambiental urbano, y por otro, la utilización de RNA como metodología para aprender un sistema de predicción del nivel del ruido ambiental, que sea más eficaz que los utilizados actualmente. El segundo objetivo se puede dividir en 3 subobjetivos más específicos, con lo que los objetivos concretos de la memoria son los cuatro siguientes:

1. Revisión del estado del arte en publicaciones que relacionen el ruido ambiental con técnicas de Soft Computing.
2. Desarrollo de un sistema, basado en Redes Neuronales, para predecir el nivel de ruido urbano en Granada, de acuerdo con un conjunto inicial de variables de entrada.
3. Reducción del número de variables de entrada relevantes, mediante un proceso de extracción y selección de características.
4. Validación del sistema desarrollado con datos no obtenidos directamente para este trabajo.

El contenido del resto de la memoria se estructura de la siguiente forma:

- En el capítulo I, de herramientas, se van a describir todos los aspectos desarrollados por otros autores que se consideran necesarios para facilitar el seguimiento del resto de la memoria. Por un lado, se recordarán los conceptos y modelos necesarios para el estudio del ruido y su predicción. También se comentan los conceptos básicos sobre técnicas de Soft Computing, y de Extracción/Selección de Características, necesarios para entender el resto de la memoria.
- En el capítulo II se aborda el objetivo 1, la revisión de la literatura actual que relaciona el tratamiento del ruido urbano con técnicas Soft Computing. Primero se establece una clasificación por áreas del análisis del ruido y por técnicas de Soft Computing utilizadas, y después una descripción suficiente de las publicaciones encontradas.
- En el capítulo III se desarrolla el objetivo 2, describiendo la RNA desarrollada, que aprende la predicción de ruido, con datos disponibles de la ciudad de Granada.
- En el capítulo IV se aplican varios métodos de Extracción y Selección de Características para reducir el número de variables relevantes, para obtener un conjunto final que permita entrenar la RNA con suficiente eficacia, cumpliendo así el objetivo 3.

- El capítulo V, aborda el objetivo 4, validando el proceso desarrollado en los capítulos anteriores, con datos que no se han obtenido específicamente para desarrollar este sistema de RNA para predecir el ruido de la ciudad de Granada.
- Posteriormente, en el capítulo VI, se incluyen las conclusiones y trabajos de futuro.
- Finalmente se incluyen las referencias utilizadas.

Como anexo, se incluye un CD con los resultados obtenidos en todos los experimentos realizados, tanto para entrenar la RNA como en la aplicación de los métodos de Selección de Características y que no se han incluido en la memoria.

CAPÍTULO I: HERRAMIENTAS

Cuando se realiza una actividad investigadora, a la hora de explicitar los resultados, es importante separar las aportaciones nuevas realizadas, de las aportaciones de otros autores, que de una forma directa o indirecta, se utilizan en la investigación desarrollada. Lógicamente, las “herramientas” aportadas con anterioridad por otros autores, y que se han utilizado en el presente trabajo, se han de referenciar y describir con suficiente detalle. En este sentido, en la presente memoria se incluye este capítulo de Herramientas.

Se presenta una descripción de las contribuciones de otros autores, que preceden al trabajo desarrollado en esta memoria y que han sido utilizadas en mayor o menor grado, o incluso de forma indirecta, para el desarrollo del trabajo presentado aquí. Las descripciones son más detalladas, cuando se ha considerado que las contribuciones descritas son menos conocidas en el ámbito científico. También se recuerdan los conceptos básicos que se han considerado necesarios para un mejor entendimiento de la investigación realizada.

El trabajo que se presenta, es una aplicación de técnicas de Soft Computing, en el proceso de análisis, fundamentalmente predicción, del ruido ambiental urbano. Por ello es importante describir lo básico del análisis del ruido y lo básico de las técnicas de Soft Computing consideradas, para lograr un texto autocompletado.

En primer lugar, se describen una serie de conceptos y definiciones sobre el ruido ambiental y su medida, empezando por el estudio del sonido desde el punto de vista de la Física. Estas definiciones ayudarán a comprender, a los no expertos, el proceso de medición de ruido, básico en el trabajo que se presenta, así como conocer qué tipo de variables se van a encontrar en el desarrollo del sistema de predicción.

Adicionalmente, es importante seleccionar y describir, con suficiente detalle, los modelos matemáticos más importantes que hasta la fecha se han propuesto para predecir el nivel de presión sonora, y conocer las características del paisaje sonoro (variables) que cada modelo utiliza. Casi todos los modelos descritos se utilizan “oficialmente”, en algunos países o ciudades, para la predicción del ruido urbano. Además, estos modelos servirán para comparar los resultados obtenidos con el modelo presentado en esta memoria.

Por otro lado, dado que la naturaleza del trabajo desarrollado está basada en las redes neuronales y también se hace referencia a otras metodologías de Soft Computing para la revisión de la literatura sobre el tema, se incluye una descripción somera de los conceptos básicos, detallando lo relativo a las Redes Neuronales Artificiales (RNA), porque la propuesta que se hace en la memoria está basada en las RNA.

También se incluye, en este capítulo, una descripción de diferentes métodos de Extracción y de Selección de Características, que se han utilizado para analizar las variables que se consideran en el modelo propuesto, y tratar de descubrir las que son realmente relevantes y las que no, ya que la eliminación en el modelo de las variables no relevantes, simplifica el proceso y reduce tiempo y costes. Se describe el Análisis de Componentes Principales y varios métodos que se han considerado métodos de referencia.

1.1 EL SONIDO Y EL RUIDO

El sonido es una vibración que se transmite por el aire. En consecuencia, el número de veces que esta vibración se produce por segundo se llama frecuencia del sonido y se mide en hercios (Hz). La percepción auditiva normal de una persona joven saludable varía aproximadamente desde 20 hasta 20000 Hz (20 kHz).

La presión sonora o acústica es producto de la propia propagación del sonido. La energía provocada por las ondas sonoras genera un movimiento ondulatorio de las partículas del aire, provocando la variación alterna en la presión estática del aire (pequeñas variaciones en la presión atmosférica). La razón de estas variaciones de presión atmosférica es que se producen áreas donde se concentran estas partículas (zonas de concentración) y otras áreas quedan menos saturadas (zonas de rarefacción¹). Las zonas con mayor concentración de moléculas tienen mayor densidad y las zonas de menor concentración tienen menor densidad. Cuando estas ondas se encuentran en su camino con el oído, la presión que ejercen sobre el mismo no es igual para toda la longitud de onda.

Así pues, la presión sonora se define como la diferencia de presión instantánea (cuando la onda sonora alcanza al oído) y la presión atmosférica estática. Comparado con la presión estática del aire, 105 Pa (Pascuales), las variaciones de presión sonora audible son muy pequeñas, en un margen que puede ir desde los 20 μ Pa hasta 100 Pa. 20 μ Pa corresponde al umbral auditivo medio de una persona. Una presión sonora de

¹ Disminución de la densidad en un cuerpo gaseoso.

aproximadamente 100 Pa es tan alta que causa dolor y por lo tanto es llamada umbral del dolor. La relación entre estos dos extremos es mayor que de un millón a uno. Aplicar de forma directa las escalas lineales (en Pa) a la medida de la presión sonora lleva a cifras enormes e inmanejables. Ya que el oído responde a los estímulos de forma logarítmica, más que lineal, es más práctico expresar los parámetros acústicos como una relación logarítmica entre el valor medido respecto a un valor de referencia. Esta relación logarítmica es llamada decibelio o dB. Aquí, la escala lineal con sus grandes cifras se convierte en una escala manejable, desde 0 dB en el umbral auditivo (20 μ Pa), hasta 130 dB, en el umbral del dolor (\sim 100 Pa). Estos rangos pueden observarse en la Figura 1.

Físicamente, no existe ninguna distinción entre sonido y ruido. Nuestro oído reconoce información en los sonidos que escuchamos. La información que no necesitamos o que no queremos pasa a ser ruido, por lo que el ruido se define como un sonido no deseado. Las características del ruido que nos hacen atender y prestar atención son tonos o cambios en el nivel sonoro. Cuanto más destacable sea el tono o más abrupto el cambio de nivel sonoro, más perceptible es el ruido. En casa, y en el trabajo, a menudo oímos ruidos procedentes de sistemas de ventilación o de calefacción, a los cuales difícilmente prestamos atención ya que no tienen características destacables. Esos ruidos nunca paran y no tienen tono, pero si de repente el ventilador se parara o empezara a zumbiar, el cambio podría llamarnos la atención o incluso molestarnos.

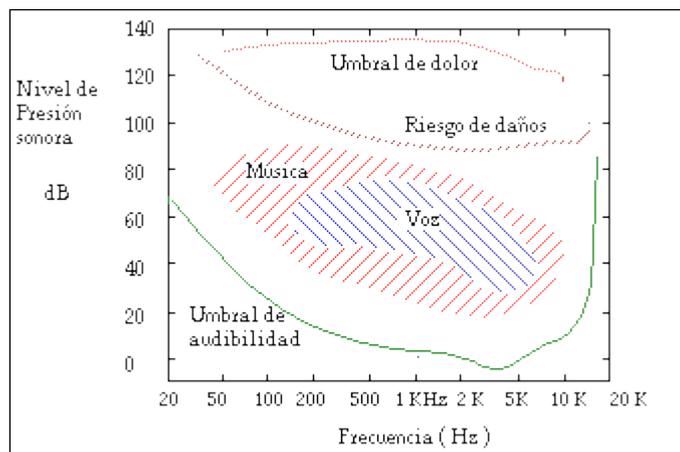


Figura 1. Rango de audición

Inicialmente, puede considerarse que un ruido es más molesto cuanto mayor nivel sonoro tenga, pero como acabamos de ver, esto no es así. Un concierto al aire libre puede ser de sumo agrado para los asistentes al mismo, pero de la misma forma puede convertirse en insoportable para los vecinos, que se habrán convertido en oyentes involuntarios.

El nivel de presión sonora tiene la ventaja de ser una medida objetiva y bastante cómoda de la intensidad del sonido, pero tiene la desventaja de que está lejos de representar con precisión lo que realmente se percibe. Esto se debe a que la sensibilidad

del oído depende fuertemente de la frecuencia. En efecto, mientras que un sonido de 1 kHz y 0 dB ya es audible, es necesario llegar a los 37 dB para poder escuchar un tono de 100 Hz, y lo mismo es válido para sonidos de más de 16 kHz. Cuando esta dependencia de la frecuencia de la sensación de sonoridad fue descubierta y medida (por Fletcher y Munson, en 1933, Figura 2), se pensaba que utilizando una red de filtrado (o ponderación de frecuencia) adecuada sería posible medir esa sensación en forma objetiva. Esta red de filtrado tendría que atenuar las bajas y las muy altas frecuencias, dejando las medias casi inalteradas.

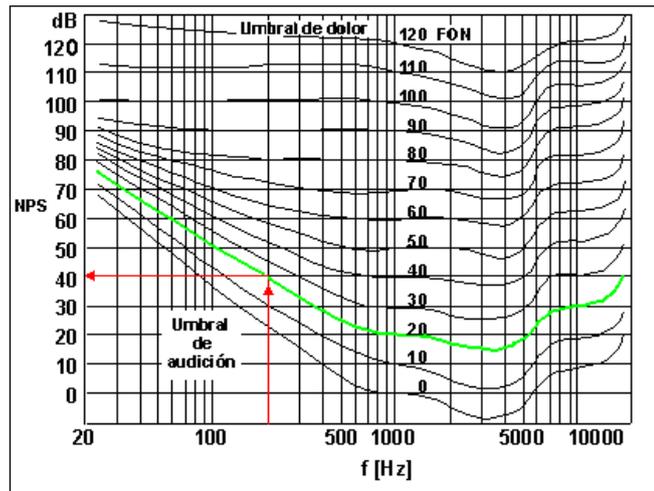


Figura 2. Curva de Fletcher y Munson

Había sin embargo algunas dificultades para implementar tal instrumento o sistema de medición. El más obvio era que el oído se comporta de diferente manera con respecto a la dependencia de la frecuencia para diferentes niveles físicos del sonido. Por ejemplo, a muy bajos niveles, sólo los sonidos de frecuencias medias son audibles, mientras que a altos niveles, todas las frecuencias se escuchan más o menos con la misma sonoridad. Por lo tanto parecía razonable diseñar tres redes o curvas de ponderación de frecuencia correspondientes a niveles de alrededor de 40 dB, 70 dB y 100 dB, llamadas A, B y C respectivamente. Son denominadas redes ya que “filtran” el nivel de ruido, bajándolo o subiéndolo para tratar de adaptarse a la respuesta en frecuencia del oído humano a diferentes niveles. La red de ponderación A se aplicaría a los sonidos de bajo nivel, la red B a los de nivel medio y la C a los de nivel elevado (Figura 3). El resultado de una medición efectuada con la red de ponderación A se expresa en decibelios A, abreviados dBA o algunas veces dB(A). Similares conceptos se aplican para las ponderaciones B y C.

En la gráfica representada en la Figura 3, el eje de abscisas representa la frecuencia y el eje de ordenadas representa el valor que se restará al nivel de presión sonora medido. Observando la curva de ponderación A, se puede ver cómo para un sonido con una frecuencia de aproximadamente 100 Hz, se reduce 20 decibelios su valor real, mientras que un sonido de 1000 Hz no es reducido. Diferentes medidas se aplican para las curvas B y C. La ponderación B se usa para niveles intermedios y es similar a la curva A, a excepción

de que la reducción de bajos es mucho menor, aunque todavía significativa (-10 dB a 60 Hz). Estudios recientes demuestran que es la mejor ponderación para usar en la medida de niveles de escucha musical. La ponderación C es muy similar a la B en agudos, y apenas aporta atenuación de las frecuencias graves. Se planteó para la evaluación de ruidos de alto nivel. Existen otras ponderaciones para aplicaciones especiales, como la D, para uso de muy alta presión como ruido aeronáutico.

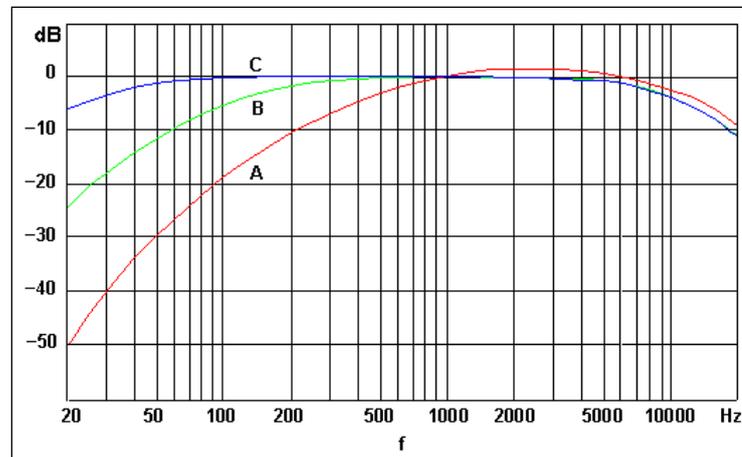


Figura 3. Curvas de ponderación A, B y C

PARÁMETROS ESTABLECIDOS PARA LA MEDICIÓN

Teniendo en cuenta todos los factores que rodean a la manifestación acústica, se establecen las valoraciones del ruido mediante índices energéticos durante un periodo de tiempo determinado y mediante análisis estadísticos de los datos que se han medido en los tiempos en consideración. A algunos de estos parámetros se les aplica la ponderación A, B o C. Aquí se incluyen las definiciones para la ponderación A, pero igualmente se hacen con las otras (B y C). Algunos de los más utilizados son:

- **Nivel L_{eq} :** nivel sonoro continuo equivalente. Se conoce universalmente como el parámetro promedio esencial. El L_{eq} es el nivel que, de haber sido constante durante el período de medición, representaría la misma cantidad de energía presente en el nivel de presión sonora medido y fluctuante. El L_{eq} se mide directamente con un sonómetro integrador y es una medida de la energía promedio en un nivel sonoro variante, es decir, es un valor acumulado. Aunque investigaciones más amplias han mostrado que el L_{eq} se correlaciona bien con la molestia, no es una medida directa de la molestia.
- **Nivel L_{Aeq} :** nivel L_{eq} ponderado según la red A (comentado en la sección anterior). Es una práctica recomendada indicar siempre la ponderación frecuencial aplicada.
- **Niveles percentiles $L_{AN,T}$:** El nivel de ruido ponderado A que se supera durante el N% del tiempo de medición T. Es decir, colocados los datos en orden creciente según el nivel de ruido equivalente, el nivel $L_{AN,T}$ es el percentil N de esa ordenación. En algunos países, el $L_{A90,T}$ (nivel de ruido excedido durante el 90% del tiempo de medición) o el

nivel $L_{A95,T}$ se utilizan como medida del nivel de ruido de fondo. El estudio de los niveles percentiles ha dado lugar a numerosos trabajos, entre ellos cabe destacar[7][8][9][10].

- **Nivel L_{AeqMax} :** El nivel L_{AeqMax} es el máximo valor obtenido para el nivel continuo equivalente ponderado A (L_{Aeq}) durante el total del periodo de medida, por lo tanto se trata de un valor puntual del L_{Aeq} , el cual es el más alto en toda la medida.
- **Nivel L_{AeqMin} :** Tiene el mismo significado que el parámetro anterior, pero para el valor mínimo obtenido, siendo el más bajo en toda la medida.
- **Nivel L_{pico} :** valor instantáneo, integrado a un periodo de tiempo menor que L_{Aeq} , que muestra el nivel de ruido más alto obtenido durante el periodo total de medida. No es un valor L_{eq} , sino que es el valor de ruido ambiental más alto. Generalmente este valor esta en ponderación C.

El ruido con tonos destacados, como el de ventiladores, compresores o sierras eléctricas, es mucho más molesto que otros ruidos. Este factor de molestia no se tiene en cuenta en las mediciones de banda ancha, en las que únicamente se evalúa la potencia sonora. Para evaluar la molestia se necesita un análisis espectral, en el que se obtiene un nivel de potencia sonora por cada una de las frecuencias.

ELEMENTOS INVOLUCRADOS EN LA PROPAGACIÓN DEL SONIDO

Una vez descritos estos parámetros, se han de considerar los elementos que van a influir en mayor o menor medida en la propagación del sonido[11]:

- **Tipo de Fuente:** Se distingue entre fuente puntual, cuando las dimensiones de la misma son pequeñas en comparación con la distancia al oyente; y fuente lineal, si la fuente es larga en una dirección y estrecha en la otra comparada con la distancia al oyente. Un ejemplo de fuente puntual puede ser un aparato de radio, mientras que el ruido del tráfico rodado en una calle se considera una fuente lineal. Ya que en este trabajo se está considerando el tráfico rodado, es necesario conocer las siguientes definiciones según establece la Recomendación de la Comisión Europea [12]:
 - *Caudal de vehículos ligeros:* número de vehículos de hasta 3500 Kg. de carga útil (automóviles, todo terrenos, furgonetas de reparto, etc.) por unidad de tiempo. Esta cantidad se expresa habitualmente como Q_{vl} .
 - *Caudal de vehículos pesados:* número de vehículos de más de 3500 Kg. de carga útil (camiones, autobuses, vehículos de limpieza, vehículos de recogida de basura, etc.) por unidad de tiempo. Normalmente expresado como Q_{vp} .
 - *Caudal de motocicletas:* número de motocicletas y ciclomotores por unidad de tiempo. Se expresa como Q_{mot} .
- **Velocidad:** La influencia de la velocidad en el nivel de presión sonora es prácticamente nulo en el caso de vías urbanas, mientras que en vías interurbanas sí se

considera. Esto se debe a que en las vías urbanas la variación de la velocidad media suele ser muy pequeña (entre 20 y 30 Km/h), y así todos los modelos y las mediciones realizadas suelen ser en las mismas condiciones, mientras que en las vías interurbanas puede variar de 60 a 140 Km/h, implicando ello variaciones apreciables en los resultados de los modelos y mediciones.

- **Absorción del aire:** El aire atenúa la propagación del sonido debido a las pérdidas energéticas de la onda al propagarse por el mismo. El grado de atenuación depende de los siguientes factores: distancia desde la fuente, contenido frecuencial del ruido (nivel sonoro por cada frecuencia), temperatura ambiental, humedad relativa y presión ambiental. Entre estos factores, los dos primeros son los considerados más influyentes, pudiendo apreciarse en la Figura 4 la relación entre ellos.

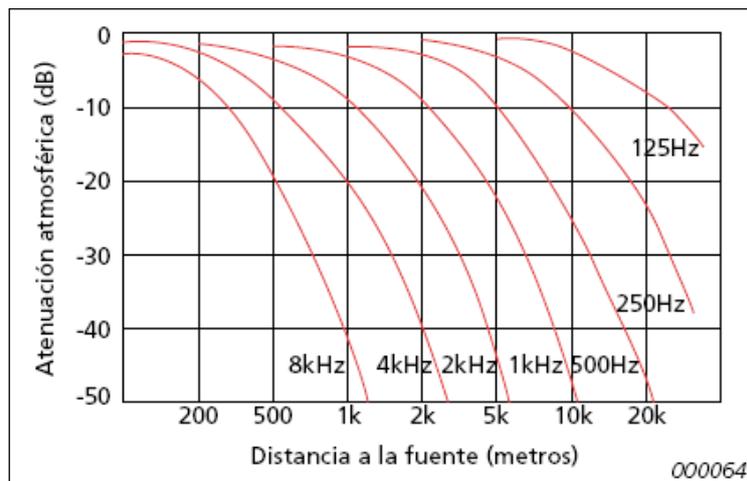


Figura 4. Atenuación atmosférica según distancia a la fuente[11]

Al considerar la distancia desde la fuente, hay estudios que incluyen también el ángulo de visión. El ángulo de visión es el ángulo que forman la línea por la que se desplaza la fuente de ruido (por ejemplo, la carretera) y la línea imaginaria trazada entre la fuente del ruido y el oyente.

El estudio del efecto del aire en la propagación del sonido es algo complejo, resumiéndose en que el efecto del aire es mayor a altas frecuencias, no atenuando bien las bajas. Para este trabajo, dado que se manejarán datos de situaciones urbanas, no se va a tener en cuenta, ya que las distancias son bastante pequeñas.

- **Vías:** El tipo de vía es un factor que influye de manera decisiva en la propagación del sonido. Las calles se clasifican de la siguiente manera dependiendo de la presencia de edificios:
 - *Calles en U:* poseen edificios a ambos lados de la misma, produciéndose el llamado “efecto cañón”, debido a la reflexión del sonido en las fachadas de los edificios.

- *Calles en L*: son aquellas que poseen edificios a un lado de la misma únicamente, produciéndose el “efecto fachada”, debido a la reflexión del sonido en las fachadas de los edificios.
- *Calles en J*: aquellas vías que poseen edificios con una diferencia de altura considerable entre ambos lados de la misma.
- *Calles en “campo libre”*: no poseen edificios a ninguno de los lados de la misma.

Otros parámetros tenidos en cuenta en este estudio respecto a las vías son:

- *Anchura de la calle*: Determina la mayor o menor distancia recorrida por las ondas sonoras antes de encontrarse con los edificios y sufrir el proceso de reflexión.
 - *Altura de la calle*: Se considera como la altura media de los edificios de la misma.
 - *Pendiente de la vía*: Inclinación de una calle respecto a la horizontal. Se diferenciará entre pendiente positiva (ascendente) y pendiente negativa (descendente). Se considera, ya que un vehículo desplazándose por una vía con alta pendiente positiva hará más uso del motor y por lo tanto emitirá un ruido mayor del que emitiría al desplazarse por una vía con pendiente nula o pendiente negativa.
- **Flujo del tráfico**: según la anteriormente mencionada recomendación de la Comisión Europea [12], es un parámetro complementario al de la velocidad, que tiene en cuenta la aceleración, desaceleración, carga del motor (por pendientes o paradas) y flujo continuo o en pulsos. Se distinguirán los siguientes tipos de flujo:
 - *Flujo continuo fluido*: Los vehículos se desplazan a velocidad casi constante por el segmento de vía considerado. Se habla de “fluido” cuando el flujo es estable tanto en el espacio como en el tiempo durante períodos de al menos diez minutos. Se pueden producir variaciones en el curso de un día, pero éstas no han de ser bruscas ni rítmicas. Además, el flujo no es acelerado ni decelerado, sino que registra una velocidad constante. Este tipo de flujo corresponde al tráfico de autopistas, autovías y carreteras interurbanas, y al de las vías rápidas urbanas (excepto en las horas punta), y grandes vías de entornos urbanos.
 - *Flujo continuo en pulsos*: Flujos con una proporción significativa de vehículos en transición (acelerando o decelerando), inestables en el tiempo (es decir, se producen variaciones bruscas del flujo en períodos de tiempo cortos) y el espacio (es decir, en cualquier momento se producen concentraciones irregulares de vehículos en el tramo de la vía considerado). Sin embargo, sigue siendo posible definir una velocidad media para este tipo de flujos, que es estable y repetitivo durante un período de tiempo suficientemente largo. Este tipo de flujo corresponde a las calles de los centros urbanos, vías importantes que se encuentran próximas a la saturación, vías de conexión o distribución con numerosas intersecciones, estacionamientos, pasos de peatones y accesos a vivienda.

- *Flujo acelerado en pulsos*: En este caso se trata de un flujo en pulsos y, por lo tanto, es turbulento. Sin embargo, una proporción significativa de los vehículos está acelerando, lo que implica que la noción de velocidad sólo tiene sentido en puntos discretos, pues no es estable durante el desplazamiento. Es el caso típico del tráfico que se observa en las vías rápidas después de una intersección, en los accesos a las autopistas, en los peajes, tras un semáforo, tras un paso de peatones, etc.
 - *Flujo decelerado en pulsos*: Es el flujo contrario al anterior, pues una proporción importante de vehículos está decelerando. Este tipo de tráfico se observa en general en las grandes intersecciones urbanas, en las salidas de autopistas y vías rápidas, en la aproximación a peajes, en la aproximación a semáforos y pasos de peatones, etc.
 - *Flujo congestionado*: En este caso el flujo de vehículos tiene velocidad casi nula ya que el número de vehículos excede de la capacidad de la ruta, provocando una alta congestión de tráfico.
 - *Flujo en pulsos o intermitente*: En este caso el caudal de vehículos es inferior a 5 vehículos por minuto.
- **Pavimento**: Por encima de una determinada velocidad, el ruido total emitido por un vehículo está dominado por el contacto entre el neumático y la carretera. Dicho ruido depende de la velocidad a que circula el vehículo, el pavimento de la vía (en particular, las superficies porosas e insonorizantes) y el tipo de vehículo. Existen varios tipos de pavimento:
 - *Asfalto liso* (hormigón o mástique asfáltico): Se trata de una superficie densa y de textura regular, en hormigón asfáltico o mástique con un tamaño máximo del árido de 11-16 mm.
 - *Pavimento poroso*: pavimento con al menos un 20 % de volumen vacío. La superficie ha de tener menos de cinco años de antigüedad. Esta restricción de edad se debe a la tendencia de las superficies porosas a perder poder absorbente con el tiempo, a medida que el vacío se llena. Si se realiza un mantenimiento especial puede levantarse esta restricción de edad. Sin embargo, una vez transcurridos los primeros cinco años, deben realizarse mediciones para determinar las propiedades acústicas del pavimento. El efecto insonorizante de este pavimento está en función de la velocidad del vehículo.
 - *Cemento hormigón y asfalto rugoso*: incluye tanto el hormigón como el asfalto de textura áspera.
 - *Adoquinado o empedrado de textura lisa*: adoquinado con una distancia entre bloques inferior a 5 mm. Es el tipo de pavimento que menos absorbe el ruido, y por tanto el que mejor propaga el mismo, obteniendo un entorno más ruidoso.
 - *Adoquinado o empedrado de textura áspera*: adoquinado con una distancia entre bloques igual o superior a 5 mm.

Al igual que el tipo de pavimento en sí, se considera también el estado del mismo.

- **Presencia de pavimentos húmedos:** Se deben de tener en cuenta situaciones en las que aparezcan pavimentos con cierta humedad, ya que en estas condiciones se produce una emisión sonora mayor que en los casos de ausencia de humedad y, si bien el clima existente en la zona no presenta demasiados períodos de lluvias, existen multitud de vías en las que aparece este tipo de problema. Esto ocurre en vías céntricas de las principales ciudades donde gracias a los servicios de limpieza municipales que utilizan gran cantidad de agua para limpiar, aparecen situaciones de pavimentos húmedos que generan una mayor emisión sonora, con el agravante que este tipo de servicios municipales son realizados en el periodo nocturno, causando un mayor grado de molestia.
- **Barreras o pantallas:** Cuando entre la fuente del ruido y el receptor hay una barrera, ésta atenúa de manera considerable el sonido. Esta reducción depende principalmente de dos factores:
 - La diferencia de la trayectoria de la onda sonora al viajar por encima de la barrera comparada con la transmisión directa al receptor.
 - El contenido frecuencial del ruido.
- **Vegetación:** la presencia de vegetación también influye, atenuando el sonido en menor medida que las barreras. En este estudio del ruido urbano no existe vegetación de suficiente espesura como para ser considerada influyente en la atenuación del sonido.
- **Viento:** La velocidad del viento aumenta con la altitud, la cual desviará la trayectoria del sonido para “hacerla converger” en el lado situado a favor del viento y crear una “sombra” en el lado de la fuente que se encuentra en contra del viento. En distancias cortas, hasta 50 m, el viento tiene una influencia pequeña en el nivel de sonido medido. Para mayores distancias, el efecto del viento se hace apreciablemente mayor.
- **Temperatura:** Los gradientes de temperatura crean efectos similares a los de los gradientes de viento, excepto en que los primeros son uniformes en todas direcciones a partir de la fuente. En un día soleado y sin viento, la temperatura disminuye con la altura, creando un efecto “sombra” del sonido. En una noche clara, la temperatura puede aumentar con la altura, “haciendo converger” el sonido en la superficie del suelo.

En general, habitualmente se tienen en cuenta todas o parte de las siguientes variables:

- Caudal de vehículos, distinguiendo vehículos ligeros, pesados y en ocasiones, motocicletas y buses.
- Velocidad de los vehículos implicados en el tráfico rodado, diferenciando en algunos modelos entre las velocidades de cada uno de los tipos de vehículos.
- Distancia desde la fuente de ruido al oyente.
- Tipo de vía.
- Anchura de la vía.
- Altura de la vía.

- Pendiente de la vía.
- Tipo de flujo del tráfico.
- Tipo de pavimento.

1.2 MODELOS MATEMÁTICOS PARA LA PREDICCIÓN DE RUIDO URBANO

Los modelos de predicción de ruido urbano tratan de, a través de cálculos matemáticos, obtener el nivel sonoro dadas unas características. Cada uno de los modelos va a tener en cuenta, en primer lugar, el caudal de vehículos, obteniendo el nivel de ruido L_{eq} básico. Posteriormente a esta fórmula básica se le añaden una serie de correcciones (en algunos de los modelos) que darán la ecuación final que predecirá el nivel de ruido equivalente. Estas correcciones suelen ser debido a la distancia, al tipo de vía, al ángulo de visión, etc.

A continuación se describen someramente los modelos, que por su bondad de resultados o interés por su utilización en la vida real, es conveniente destacar:

MODELO LINEAL

Se trata de un modelo muy básico, que se atiene simplemente a la relación lineal que existe entre el nivel continuo equivalente y el caudal de vehículos en la vía. Ello supone que a priori el grado de estimación sea inferior al resto de modelos aunque sí puede ser útil a la hora establecer de una forma bastante admisible la influencia del caudal de vehículos ($\log Q$) en los niveles de ruido ambiental medidos, para cada día tipo y en cada una de las vías seleccionadas.

Se trata básicamente de una recta de regresión entre el $\log Q$ y el L_{eq} experimental medido para cada día tipo y, además, para cada una de las vías seleccionadas.

La expresión básica para el cálculo del L_{eq} , es la siguiente:

$$L_{eq} = a + b \log Q$$

donde: a y b son constantes de regresión y Q es la densidad de tráfico (n° vehículos / hora).

MODELO LINEAL MULTIVARIANTE [13]

Deriva del modelo lineal, fundamentándose de la misma forma en encontrar una relación lineal entre el caudal de vehículos y el nivel de presión sonora continuo equivalente, pero le añade complejidad, teniendo en cuenta otras características, como:

- Velocidad media (v), en km/h.

- Distancia media desde la calzada al lugar de medición (d), en metros.
- Proporción de vehículos pesados en tanto por uno (p).

La expresión para el nivel de presión sonora equivalente es:

$$L_{eq} = 55,7 + 11,2 \log Q + 0,4 p - 0,05v - 12,7 \log d$$

MODELO MOPU O ISO[14]

Se trata del modelo desarrollado por el antiguo Ministerio de Obras Públicas, por lo que está algo desfasado. Supuso un gran avance en los años 90 en los estudios acústicos en España, incluyéndose variables totalmente innovadoras como tipos de pavimento, ángulo de visión de la calle... En definitiva, un conjunto de nuevos valores que deben a posteriori mostrar una mayor afinidad con los resultados medidos experimentalmente.

En este caso el L_{eq} se predice considerando dos términos:

$$L_{eq} = L_{eq\ 15m} + \sum \Delta_i$$

- L_{eq} básico: Es el L_{eq} que el modelo predice para una distancia de 15 metros.

$$L_{eq\ 15m} = 40 + 10 \log Q + 10 \log (1 + 19 p)$$

donde, Q es la densidad de tráfico (nº de vehículos/h); p es el porcentaje de vehículos pesados en tanto por uno.

- Correcciones: El L_{eq} básico anterior es modificado por una serie de correcciones que serán aplicables o no según el caso.

- Correcciones por distancias (Δ_1): Corrige la distancia concreta en cada punto de medida en función de la distancia de referencia (15 metros).

$$\Delta_1 = 10 \log (15 / d) 1 + \alpha$$

Donde: d es la distancia desde el centro de la calzada al punto de medida (m) ; α es una constante en función del tipo de pavimento, que varía según la Tabla 2.

Tabla 2. Variación de la constante α según el tipo de pavimento

Pavimento	α
Absorbente	0
Poco reflectante	0,25
Reflectante	0,5

- Corrección por tipo de fachada (Δ_2): Es muy importante, ya que considera el efecto fachada o efecto cañón según se trate de calles en “L” o calles en “U”, respectivamente. En los casos de campo libre esta corrección es nula.

- Efecto Fachada: Se considera una corrección constante y siempre positiva (para calles en “L”).

$$\Delta_2 = 2,5 \text{ dB}$$

- Efecto Cañón: La corrección se establece en función de un parámetro J , que considera la relación altura / anchura. (para calles en “U”).

$$\Delta_2 = - 0,208 + 3.81J - 0,585 J^2$$

donde $J = h / b$; h es la altura media los edificios (m) y b es la anchura de la calle (m)

- Corrección por ángulo de visión (Δ_3): Mediante esta corrección se tiene en cuenta que el cilindro de propagación de las ondas sonoras producido por una fuente longitudinal como es el tráfico, no es infinito. De modo que esta corrección será de importancia en aquellas zonas donde existan curvas significantes en la calzada.

$$\Delta_3 = 10 \log \theta / 180$$

Donde: θ es el ángulo de visión expresado en grados.

- Corrección por pendiente (Δ_4): Corrige el L_{eq} en función de la mayor o menor pendiente de la calzada.

$$\Delta_4 = 0,3 \delta$$

Donde: δ es el ángulo de pendiente expresado en grados.

- Corrección por tipo de calzada (Δ_5): Depende de si la calzada está asfaltada o empedrada:

$$\Delta_5 = 10 * \text{Log} (90 * TD + 30) - 20 \text{ dBA para superficies de cemento}$$

$$\Delta_5 = 10 * \text{Log} (20 * TD + 60) - 20 \text{ dBA para superficies bituminosas.}$$

donde:

TD es la Texture Depth (profundidad de la textura).

- Corrección por barreras (Δ_6): Esta corrección viene tabulada para cada tipo de barrera acústica[15].
- Corrección por cruces: Corrige el L_{eq} en el punto de medida (vía 1), para eliminar el aporte de ruido debido al tráfico de la vía perpendicular a la estudiada con la que se forma el cruce (vía 2). Se calcula L_{Aeq} de la calle 1 según modelo MOPU y L_{Aeq} de la calle 2 según el modelo lineal.

MODELO FRANCÉS[16]

Es el modelo de predicción del nivel de presión sonora más extendido, ya que obtiene los mejores resultados. Por ello, es además el modelo recomendado por la Unión Europea en la directiva 2002/49/CE[3].

Distingue dos expresiones para el cálculo del L_{eq} , según el tipo de calzada y la relación altura media / anchura (h / b):

- Vías en “U” con $h / b > 0,2$

$$L_{eq} = 55 + 10 \log (Q_{VL} + E Q_{VP}) - 10 \log L + K_h + K_v + K_c$$

donde: Q_{VL} es la densidad de vehículos ligeros (n° vehículos ligeros / h) ;
 Q_{VP} es la densidad de vehículos pesados (n° vehículos pesados / h)

E es un factor de equivalencia acústica entre vehículos ligeros y pesados que tiene en cuenta el mayor ruido que causan las marchas cortas de los vehículos pesados cuanto mayor es la pendiente. Depende pues del grado de elevación de la vía (r). Su valor viene dado en la Tabla 3.

Tabla 3. Valor de E según la pendiente

Pendiente (r, %)	E
$\leq 2\%$	10
$\leq 3\%$	13
$\leq 4\%$	16
$\leq 5\%$	18
$\leq 6\%$	20

L es la anchura de la calzada sin contar las aceras, de modo que es diferente a la anchura utilizada anteriormente ($L < b$).

K_h es la corrección por altura:

- Si $h \leq 4$ m, la corrección es nula: $K_h = 0$
- Si $h > 4$ m, la corrección, expresada en dB, se calcula como:
 $K_h = - ((h - 4) / L)$

K_v es la corrección por velocidad media:

- Si $v < 60$ Km. / h, la corrección es nula: $k_v = 0$. Éste será el caso para la mayoría de las vías urbanas.
- Si $v > 60$ Km. / h, la corrección se calcula sumando 1 dB por cada 10 Km. / h que se supere la velocidad de referencia (60 Km. / h).

K_c es la corrección por cruce, que se obtiene de igual forma a la comentada en el apartado anterior para el Modelo MOPU o Modelo ISO.

- Vías en “L” y Vías en “U” con $h / b < 0,2$

$$L_{eq} = 20 + 10 \log (Q_{VL} E Q_{VP}) + K_h + K_v + K_a$$

donde: Q_{VL} , Q_{VP} , E siguen la misma definición que el caso anterior.

La corrección por altura, K_h , es ahora:

$$K_h = 12 \log (d + L / 3)$$

donde: d es la distancia desde el centro de la calzada al punto de medida (m) ; L es la anchura de la calzada sin contar las aceras (m)

La corrección por velocidad media, K_v , es aquí:

$$K_v = 20 \log v$$

donde v es la velocidad media y se considera que:

- Si $v > 60$ Km. / h se introduce el valor máximo: $v = 60$.
- Si $v < 60$ Km. / h se introduce el valor de v .

K_a es la corrección por ángulo de visión:

$$K_a = 10 \log \theta / 180$$

donde: θ es el ángulo de visión expresado en grados.

Existe una versión más definida de este modelo, denominada “Modelo Francés Detallado” [16], en la que se aborda de forma mucho más completa que en el método simplificado las diferentes leyes de emisión y propagación del ruido debido al tráfico rodado. Dado que es un modelo muy complejo, a modo de resumen, se comentarán los pasos principales [12]:

El nivel de emisión de ruido de un vehículo se caracteriza por el nivel sonoro máximo de paso, L_{Amax} (en dB), medido a una distancia de 7,5 metros del eje de la trayectoria del vehículo. Este nivel sonoro se determina por separado para los distintos tipos de vehículos, velocidades y flujos de tráfico. Aunque se tiene en cuenta la pendiente de la vía, no sucede lo mismo con el pavimento. Para preservar la compatibilidad con las condiciones de medida originales, se deben medir las características acústicas de los vehículos que circulen sobre los revestimientos siguientes: cemento hormigón, hormigón bituminoso de muy escaso espesor 0/14, hormigón bituminoso semigranulado 0/14, sello superficial 6/10, y sello superficial 10/14. A continuación, se introduce una corrección de pavimento.

Las medidas se pueden realizar sobre vehículos aislados o sobre circuitos específicos en condiciones controladas. La velocidad del vehículo debe medirse con un radar Doppler (que posee una precisión de aproximadamente 5% a bajas velocidades). El flujo de tráfico se determinará, bien por observación subjetiva (acelerado, decelerado o fluido) o por medición. El micrófono se coloca a 1,2 metros de altura sobre el suelo y a 7,5 metros de distancia perpendicularmente al eje de desplazamiento del vehículo.

Para su uso con el modelo francés y conforme a las especificaciones de la “Guide du Bruit 1980”, el nivel de potencia sonora L_w y la emisión sonora E se calculan a partir del nivel de presión sonora L_p y la velocidad del vehículo V mediante las fórmulas siguientes:

$$L_w = L_p + 25,5$$

$$E = (L_w - 10 \text{ Log } V - 50)$$

$$L_p = L_w + 10 \text{ Log } (Q / 1000 V) - 10 \text{ Log } l + 3 - A$$

donde: L_w es el nivel de potencia sonora media de un vehículo ; Q es el caudal de vehículos ; V es la velocidad de los vehículos ; l es la longitud de la vía en metros ; A es la pérdida por absorción.

MODELO CORTN [17]

Modelo inglés, cuyo índice descriptor del nivel sonoro es el percentil L10:

$$L_{10} (1 \text{ h}) = 42.2 + 10 \times \log (q) \text{ (dBA)}$$

$$L_{10} (18 \text{ h}) = 29.1 + 10 \times \log (Q) \text{ (dBA)}$$

donde: q es el número de vehículos por hora ; Q es el número de vehículos en 18 horas.

- La corrección para otra velocidad y tráfico compuesto viene dada por:

$$K_{v,p} = 33 \times \log(V+40+500/V) + 10 \times \log(1+5P/V) - 68.8 \text{ (dBA)}$$
 donde: V es la velocidad del flujo de tráfico en Km/h.; P es la proporción de vehículos pesados en %.
- La corrección por pendiente es:

$$L = 0,3 * g$$
 donde g es la pendiente de la carretera en %.
- Por tipo de asfalto también se hace cierta corrección que va desde -3.5 dBA para superficies porosas permeables hasta sus valores máximos dados para cementos, en función de la profundidad de textura y velocidad.

MODELO RLS 90[18]

Modelo desarrollado en Alemania, que introduce un término de corrección por velocidad, siempre que sea distinta de la asumida por defecto (100 km/h para vehículos ligeros y 80 km/h para pesados).

- El nivel base se evalúa mediante la expresión:

$$L_{Aeq} = 37.3 + 10 \log\{ (Q(1+0.082 \cdot P)) \} \text{ (dBA)}$$

donde: Q es la densidad de tráfico, en vehículos/h; P es el porcentaje de vehículos pesados (PMA>2.8 Ton), en %

- El término de corrección para una velocidad distinta de la asumida por defecto (100Km/h para ligeros y 80 Km/h para pesados), es el siguiente:

$$\Delta v = L_{\text{ligeros}} - 37.3 + 10 \log \left[\frac{100 + (10^{0.1 \cdot C} - 1) \cdot P}{100 + 8.23 \cdot P} \right]$$

Donde: $L_{\text{ligeros}} = 27.7 + 10 \log [1 + (0.02 V_{\text{ligeros}})^3]$;

$L_{\text{pesados}} = 23.1 + 12.5 \log V_{\text{pesados}}$; $C = L_{\text{pesados}} - L_{\text{ligeros}}$

$30 \text{ km/h} \leq V_{\text{ligeros}} \leq 130 \text{ km/h}$; $30 \text{ km/h} \leq V_{\text{pesados}} \leq 80 \text{ km/h}$

- La corrección por pendiente viene a ser:

$$L = 0,6 \cdot g - 3 \text{ si } g > 5$$

donde g es la pendiente de la carretera en %.

- Por tipo de asfalto también se hace cierta corrección que va desde -5 dBA para superficies porosas (con más del 15% de poros) hasta +6 dBA para una superficie de adoquinado rugosa.

MODELO PLANKWERK '96[19]

Modelo implantado en los países nórdicos, tiene en cuenta de manera aislada el nivel de ruido producido por vehículos ligeros y el producido por vehículos pesados.

- El descriptor es el nivel sonoro continuo equivalente, L_{Aeq} , evaluado para un periodo de 24 horas, aunque puede ser evaluado para cualquier intervalo de tiempo.

$$L_{Aeq} = 10 \cdot \log \left[10^{\frac{L_{Aeq,10m}(\text{ligeros})}{10}} + 10^{\frac{L_{Aeq,10m}(\text{pesados})}{10}} \right]$$

Donde: $L_{Aeq,10m}(\text{ligeros}) = L_{AE,10m}(\text{ligeros}) + 10 \log \left(\frac{N_l}{T} \right)$

$L_{Aeq,10m}(\text{pesados}) = L_{AE,10m}(\text{pesados}) + 10 \log \left(\frac{N_p}{T} \right)$

$L_{AE,10m}(\text{ligeros}) = 73.5 + 25 \cdot \log \left(\frac{V}{50} \right)$ para $V \geq 40 \text{ km/h}$

$L_{AE,10m}(\text{ligeros}) = 71$ para $V < 40 \text{ km/h}$

$L_{AE,10m}(\text{pesados}) = 81 + 30 \cdot \log \left(\frac{V}{50} \right)$ para $50 \text{ km/h} \leq V \leq 40 \text{ km/h}$

$$L_{AE,10m}(\text{pesados})=81$$

para $V < 50$ km/h

Siendo N_l y N_p el número de vehículos ligeros y pesados, respectivamente, durante el intervalo de tiempo T , en segundos.

- o La corrección por pendiente es:

$$L = (2 * g / 10) + ((3 * g / 10) * \text{Log}(1 + p))$$

donde g es la pendiente de la carretera en % y p es la proporción de vehículos pesados en tanto por 1.

- o Al igual que en los dos modelos anteriores, por tipo de asfalto también se hace una corrección, que va desde -5 dBA para superficies porosas hasta +5 dBA para una superficie de adoquinado.

MODELO E. GAJA[20]

Modelo predictivo desarrollado para la ciudad uruguaya de Montevideo. Para este modelo Leq se define como:

$$Leq=49.4+10 \log (A + 2.33 \cdot M + 9.01 \cdot O + 6.84 \cdot C) + 23.266 \cdot Q - 0.3811 \cdot 10 \log d$$

donde A , M , O , C son los caudales de automóviles, motocicletas, autobuses (ómnibuses) y camiones respectivamente. En este modelo no se distingue entre vehículos pesados y ligeros, sino entre estas cuatro categorías, representando los automóviles y motocicletas a los vehículos ligeros, y los autobuses y camiones, a los vehículos pesados.

La validez de esta ecuación en un rango de ± 3 dBA está dada para los siguientes valores de los parámetros en juego:

- Q entre 200 y 4000 vehículos/hora, incluyendo vehículos livianos, pesados y motos

- d entre 3 y 15 m

- Calles en L

- Para velocidad superior a los 60 km/h, debe aplicarse la corrección lineal: $\Delta v = 0,15 \cdot v - 8,67$

Los coeficientes que se acaban de mencionar se obtuvieron dado que en la ciudad de Montevideo un valor medio de ruido de fondo es 49,4 dBA; cada autobús equivale acústicamente a 9 automóviles; cada camión a casi 7 automóviles (precisamente 6,84) y cada motocicleta a más de 2 automóviles (2,33).

MODELO LINEAL MULTIVARIANTE GRANADINO [9]

Modelo desarrollado en Granada por los departamentos de Ingeniería Civil y Física Aplicada de la Universidad de Granada que trata de adaptarse a la realidad de esta ciudad, incluyendo una característica tan importante como es el caudal de motocicletas. Este caudal es incluido en el resto de modelos (excepto en el modelo de E. Gaja, que lo menciona explícitamente) entre el caudal de vehículos ligeros. Fue precedido por otro modelo similar, no publicado, desarrollado unos años antes y que fue pionero en la inclusión del caudal de motocicletas en la ecuación de cálculo del nivel de ruido. Se describe en primer lugar esta primera versión del modelo lineal multivariante granadino.

- L_{eq} básico: Es la suma de un nivel estándar y la ponderación del caudal de tráfico.

$$L_{eq \text{ básico}} = 42 + 10 \log (Q_{VL} + E Q_{VP} + F Q_{mot})$$

donde: Q_{VL} es la densidad de vehículos ligeros; Q_{VP} es la densidad de vehículos pesados; Q_{mot} es la densidad de ciclomotores y motocicletas; E es el factor de equivalencia acústica entre vehículos ligeros y pesados y se calcula del mismo modo que en modelo francés, según la Tabla 3; F es un factor de equivalencia acústica para los ciclomotores y motocicletas.

Es experimental, de modo que se debe analizar en relación a las demás variables conocidas hasta obtener los resultados que mejor se adapten a cada caso.

- Corrección por distancia:

$$L_d = - 10 \log d^\alpha$$

donde: d es la distancia desde el centro de la calzada al punto de medida (m) y α es una constante en función del tipo de pavimento, según la Tabla 2.

- Corrección por tipo de vía:

- Calles en "U" (efecto cañón):

$$L_{cañón} = - 0,208 + 3.81J - 0,585 J^2$$

donde: $J = h / b$; h es la altura media los edificios (m) y b es la anchura de la calle (m).

- Calles en "L" (efecto fachada):

$$L_{fachada} = + 2,5 \text{ dB}$$

La segunda versión de este modelo lineal multivariante granadino fue obtenida a partir de los datos recogidos mediante la aplicación de una fórmula empírica. A partir de la obtención de un nivel de fondo de ruido ambiental característico de la ciudad, y

aplicándole posteriormente un factor de corrección por eventos anómalos, se estima con precisión el valor del L_{Aeq} . La ecuación obtenida fue la siguiente:

$$L_{Aeq} = NB + 10 \text{ Log } [(a \cdot Q_{vl}) + (b \cdot Q_{vp}) + (c \cdot Q_{mc})]$$

donde: NB es el Nivel Base ó nivel de fondo obtenido; a , b y c son los coeficientes para caudal de vehículos ligeros, pesados y motocicletas-ciclomotores, respectivamente; Q_{vl} es el caudal de vehículos ligeros; Q_{vp} es el caudal de vehículos pesados; Q_{mc} es el caudal de motocicletas y ciclomotores

Lo interesante de esta propuesta es que considera obtener un nivel de ruido de fondo (NB) para caracterizar, de una forma precisa, el ambiente acústico de la ciudad con un flujo constante de vehículos circulante y sin la aparición de eventos anómalos. Estos eventos aparecen con relativa frecuencia en las aglomeraciones urbanas, pueden ser sirenas, ladridos de perros, gritos, etc., y tienen una influencia muy notable con respecto a los niveles de ruido ambiental. Para ello, se ha considerado la utilización del percentil L_{90} como nivel de ruido de fondo.

Se ha utilizado para la elaboración de un mapa de ruido en Granada, de acuerdo a la directiva 2002/49/CE[3], que persigue evaluar y analizar la situación acústica de la ciudad, de forma que se puedan sentar las bases que orientarán el Plan de Acción que deberá ser llevado a cabo para solucionar y mejorar dicha situación acústica.

El estudio de los niveles percentiles ha dado lugar a publicaciones como [7][9]. En el primero de ellos se persigue, en primer lugar, estudiar la relación entre los niveles percentiles y L_{eq} y las fluctuaciones sonoras mediante el valor de la diferencia $L_{10}-L_{90}$ y $L_{10}-L_{eq}$ y, posteriormente, hallar una relación entre los descriptores del ruido y el caudal de vehículos. De la misma forma y como se acaba de comentar en los párrafos anteriores, en [9] hacen una propuesta de modelo de predicción del L_{eq} y finalmente sugieren L_{50} como descriptor de ruido generado por el tráfico, al advertir un elevado nivel de correlación del mismo con el caudal de tráfico [7].

Por otro lado, en[21] se analiza qué descriptores de ruido ambiental representan de mejor manera el ruido causado por el tráfico rodado, inclinándose finalmente por L_{50} , al igual que en [7]. También en [21] se desarrolla en detalle la obtención del factor de corrección por la aparición de eventos anómalos (eventos sonoros cortos y de gran intensidad), mediante el cálculo previo del ruido de fondo (L_{90}). En torno al cálculo de este factor de corrección se desarrollan las comunicaciones [8][10][22].

Durante el año 2006 se han realizado, por parte del grupo de trabajo de estos investigadores, nuevas tomas de los niveles de ruido de la ciudad de Granada, datos que servirán para alimentar al sistema que aquí se propone

En la Tabla 4 se incluye un cuadro resumen de los diferentes modelos que se acaban de presentar, que pueden servir para una comparación rápida, de una sola mirada.

Tabla 4. Comparativa de los modelos predictivos de ruido urbano

		LINEAL	LINEAL MULTIVARIANTE	MOPU
Índice descriptor ruido		L_{eq}	L_{eq}	L_{eq}
Ecuación		$L_{eq} = a + b \log Q$	$L_{eq} = 55,7 + 11,2 \log Q + 0,4 p - 0,05v - 12,7 \log d$	$L_{eq} = L_{eq\ 15m} + \text{suma}(\text{correcciones } i)$
VARIABLES ENTRADA				$L_{eq\ 15m} = 40 + 10 \log Q + 10 \log (1 + 19 p)$
Caudal (vehículos / hora)	Vehículos ligeros	Q	Q	Q
	Vehículos pesados	Q	Q, p	Q, p
Distancia			d	$\Delta_1 = 10 \log (15 / d)^{1 + \alpha}$
Tipo de vía	Calles en L			$\Delta_2 = 2,5 \text{ dB}$
	Calles en U			$\Delta_2 = -0,208 + 3,81J - 0,585 J^2$
Ángulo de visión				$\Delta_3 = 10 \log \theta / 180$ (θ : ángulo de visión expresado en grados)
Pendiente				$\Delta_4 = 0,3 \delta$ (δ : ángulo de pendiente expresado en grados)
Tipo de calzada				$\Delta_5 = 10 * \text{Log} (90 * TD + 30) - 20 \text{ dBA}$ para superficies de cemento $\Delta_5 = 10 * \text{Log} (20 * TD + 60) - 20 \text{ dBA}$ para superficies bituminosas. TD: profundidad de la textura.
Barreras				Δ_6 según[15]
Cruces				Δ_7 (L_{Aeq} de calle 1 según modelo MOPU y L_{Aeq} de calle 2 según el modelo lineal)
Anchura de la calzada				Absorbente, $\alpha=0$ / Poco reflectante, $\alpha=0.25$ / Reflectante $\alpha=0.5$
Altura				$J = h / b \rightarrow h$: altura media los edificios (metros) / b : anchura de la calle (metros)
Velocidad			v	

		FRANCES																	
		Vías en U con $h/b > 0.2$			Vías en L y vías en u con $h/b < 0.2$														
Índice descriptor ruido		Leq			Leq														
Ecuación		$Leq = 55 + 10 \log (Q_{VL} + E Q_{VP}) - 10 \log L + K_h + K_v + K_c$			$Leq = 20 + 10 \log (Q_{VL} E Q_{VP}) + K_h + K_v + K_a$														
VARIABLES ENTRADA																			
Caudal de vehículos por hora	Veh.ligeros	Qvl																	
	Veh.pesados	Qvp																	
Ángulo de visión					$K_a = 10 \log \theta / 180$														
Pendiente		<table border="1"> <thead> <tr> <th>Pendiente (r, %)</th> <th>$\leq 2\%$</th> <th>$\leq 3\%$</th> <th>$\leq 4\%$</th> <th>$\leq 5\%$</th> <th>$\leq 6\%$</th> </tr> </thead> <tbody> <tr> <td><i>E</i></td> <td>10</td> <td>13</td> <td>16</td> <td>18</td> <td>20</td> </tr> </tbody> </table>						Pendiente (r, %)	$\leq 2\%$	$\leq 3\%$	$\leq 4\%$	$\leq 5\%$	$\leq 6\%$	<i>E</i>	10	13	16	18	20
		Pendiente (r, %)	$\leq 2\%$	$\leq 3\%$	$\leq 4\%$	$\leq 5\%$	$\leq 6\%$												
<i>E</i>	10	13	16	18	20														
Cruces		K_c se calcula igual que MOPU																	
Anchura de la calzada		L																	
Altura		$K_h = 0$ si $h \leq 4$ m $K_h = -(h-4/L)$ si $h > 4$ m			$K_h = 12 \log (d + L / 3)$														
Velocidad		$K_v = 0$ si $v < 60$ km/h $K_v = +10$ db por cada 10 km/h que supera 60 km/h si $v > 60$ km/h			$K_v = 20 \log v$														
Observaciones					- <i>d</i> es la distancia desde el centro de la calzada al punto de medida (metros) - <i>L</i> es la anchura de la calzada sin contar las aceras (metros)														

		CORTN	RLS90
Índice descriptor ruido		L ₁₀	L _{eq}
Ecuación		$L_{10} = 42.2 + 10 \times \log(q)$	$L_{Aeq} = 37.3 + 10 \log\{Q(1 + 0.082 \cdot P)\}$
VARIABLES ENTRADA			
Caudal de vehículos por hora	Vehículos ligeros	q	Q
	Vehículos pesados	p, P	Q, P
Pendiente		$L = 0,3 \cdot g$	$L = 0,6 \cdot g - 3$ donde $g > 5$
Tipo de calzada		Desde -3.5 dBA*	Desde -5 dBA*
Velocidad		$K_{v,p} = 33 \times \log(V + 40 + 500/V) + 10 \times \log(1 + 5P/V) - 68.8$	$\Delta v = L_{\text{ligeros}} - 37.3 + 10 \log \left[\frac{100 + (10^{0.1 \cdot C} - 1) \cdot P}{100 + 8.23 \cdot P} \right]$
Observaciones		<p>- V es la velocidad del flujo de tráfico en Km/h.</p> <p>- P es la proporción de vehículos pesados en %.</p> <p>- g es la pendiente de la calzada en %</p> <p>*Para superficies porosas permeables, hasta sus valores máximos en función de la profundidad de textura y velocidad.</p>	<p>$L_{\text{ligeros}} = 27.7 + 10 \log [1 + (0.02 V_{\text{ligeros}} \text{ ligeros}^3)]$</p> <p>$L_{\text{pesados}} = 23.1 + 12.5 \log V_{\text{pesados}}$</p> <p>$C = L_{\text{pesados}} - L_{\text{ligeros}}$</p> <p>$30 \text{ km/h} \leq V_{\text{ligeros}} \leq 130 \text{ km/h}$</p> <p>$30 \text{ km/h} \leq V_{\text{pesados}} \leq 80 \text{ km/h}$</p> <p>- g es la pendiente de la calzada en %</p> <p>- P es el porcentaje de vehículos pesados.</p> <p>*Para superficies porosas (con más del 15% de poros) hasta +6 dBA para una superficie de adoquinado rugosa.</p>

		PLANKWERK '96	E. GAJA
Indice descriptor ruido		L_{eq}	L_{eq}
Ecuación		$L_{Aeq} = 10 \cdot \log \left[10^{\frac{L_{Aeq,10m}(ligeros)}{10}} + 10^{\frac{L_{Aeq,10m}(pesados)}{10}} \right]$	$L_{eq} = 49.4 + 10 \log (A + 2.33 \cdot M + 9.01 \cdot O + 6.84 \cdot C) + 23.266 \cdot Q^{-0.3811} - 10 \log d$
VARIABLES ENTRADA			
Caudal de vehículos por hora	Vehículos ligeros	$L_{Aeq,10m}(ligeros) = L_{AE,10m}(ligeros) + 10 \log \left(\frac{N_l}{T} \right)$ $L_{AE,10m}(ligeros) = 73.5 + 25 \cdot \log \left(\frac{V}{50} \right) \quad \text{para } V \geq 40 \text{ km/h}$ $L_{AE,10m}(ligeros) = 71 \quad \text{para } V < 40 \text{ km/h}$	A (automóviles) M (motocicletas)
	Vehículos pesados	$L_{Aeq,10m}(pesados) = L_{AE,10m}(pesados) + 10 \log \left(\frac{N_p}{T} \right)$ $L_{AE,10m}(pesados) = 81 + 30 \cdot \log \left(\frac{V}{50} \right) \quad \text{para } 50 \text{ km/h} \leq V \leq 40 \text{ km/h}$ $L_{AE,10m}(pesados) = 81 \quad \text{para } V < 50 \text{ km/h}$	C (camiones) O (buses)
Distancia			d
Pendiente		$L = (2 * g / 10) + ((3 * g / 10) * \text{Log} (1 + p))$	
Tipo de calzada		Desde -5dBA para asfaltos porosos hasta + 5dBA para adoquinados	
Velocidad			$\Delta v = 0,15 \cdot v - 8,67$ si $v > 60 \text{ km/h}$

		LINEAL MULTIVARIANTE GRANADINO I	LINEAL MULTIVARIANTE GRANADINO II
Índice descriptor ruido		L_{eq}	L_{eq}
Ecuación		$L_{eq} \text{ básico} = 42 + 10 \log (Q_{VL} + E Q_{VP} + F Q_{mot})$	$L_{Aeq} = NB + 10 \text{ Log } [(a \cdot Q_{vl}) + (b \cdot Q_{vp}) + (c \cdot Q_{mc})]$
VARIABLES ENTRADA			
Caudal de vehículos por hora	Veh.ligeros	Q_{vl}	Q_{vl}
	Veh.pesados	Q_{vp}	Q_{vp}
	Motocicletas	Q_{mot}	Q_{mc}
	Buses	Q_{vp}	Q_{vp}
Distancia de la fuente al receptor		$L_d = - 10 \log d^\alpha$	
Tipo de vía	Calles en L	$L_{fachada} = + 2,5 \text{ dB}$	
	Calles en U	$L_{cañón} = - 0,208 + 3.81J - 0,585 J^2$	
Pendiente		E se calcula igual que modelo francés	
Observaciones		<p>Absorbente, $\alpha=0$; Poco reflectante, $\alpha=0.25$; Reflectante $\alpha=0.5$</p> <p>- F es un factor de equivalencia acústica para los ciclomotores y motocicletas.</p> <p>- $J = h / b$</p> <p>- h es la altura media los edificios(metros).</p> <p>- b es la anchura de la calle (metros).</p>	<p>- NB es el Nivel Base ó nivel de fondo obtenido.</p> <p>- a, b y c son los coeficientes para caudal de vehículos ligeros, pesados y motocicletas-ciclomotores, respectivamente.</p> <p>En [21] se propuso: $NB=55.28$; $a=0.235$; $b=0.238$; $c=0.518$</p>

1.3 REDES NEURONALES ARTIFICIALES Y OTRAS TÉCNICAS SOFT COMPUTING

El término Soft Computing se refiere a un número de técnicas computacionales usadas para estudiar, modelar y analizar sistemas complejos para los que las técnicas convencionales no proporcionan resultados satisfactorios en términos de tiempo o precisión. Las redes neuronales artificiales (RNA) [23], los sistemas difusos[24], la computación evolutiva[25] y los Modelos Ocultos de Markov (HMM)[26] son algunas de las técnicas de Soft Computing utilizadas en el análisis del ruido ambiental.

De todas estas técnicas, en esta memoria se utilizan sólo las RNA, que por otro lado son ya bastante conocidas en la comunidad científica, por lo que sólo se hace una muy básica descripción, indicando alguna referencia básica importante.

De las otras técnicas sólo se indica una breve idea del concepto básico y una referencia genérica.

REDES NEURONALES ARTIFICIALES

Las redes neuronales artificiales son un paradigma de aprendizaje y procesado automático cuyo fin último es emular el cerebro humano, o al menos algunas de sus funciones, como el aprendizaje. Las RNA están compuestas de un grupo interconectado de neuronas artificiales o procesadores elementales (PE), donde las conexiones tienen peso cuya modificación representa el aprendizaje de la red. Éstas proporcionan una respuesta única o salida a partir de un conjunto de entradas desde el exterior o desde otras neuronas.

Las RNA aprenden a resolver problemas desde la observación de una serie de instancias que representa el tipo de problema. Para ello, las RNA deben ser entrenadas modificando los pesos de las conexiones, para dar una respuesta correcta a cada ejemplo del problema considerado.

La arquitectura o topología de una red consiste en la organización y disposición de las neuronas y sus conexiones en la red. Dependiendo de la topología o estructura de la red, las RNA pueden tener diferentes características tales como tener (o no) diferentes capas (de entrada, ocultas y de salida), tener propagación hacia adelante o hacia atrás, o usar diferentes funciones para evaluar el rendimiento de la red.

Las redes neuronales pueden ser clasificadas en tres tipos, de acuerdo a la forma en la que sus elementos principales están conectados:

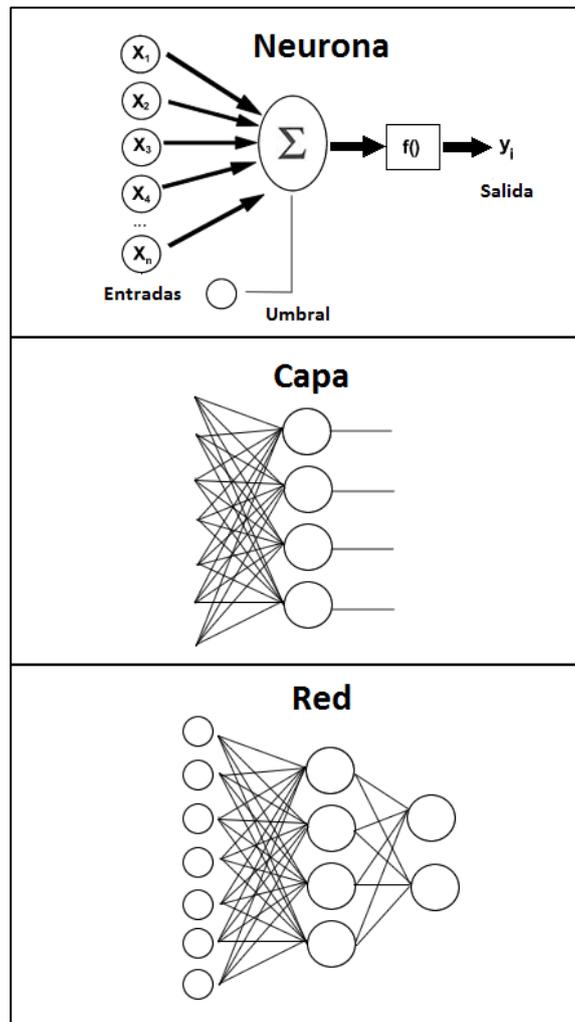


Figura 5. Estructura de una red neuronal feedforward

- Redes neuronales feedforward: solo contienen conexiones hacia adelante o feedforward entre capas. Esto significa que una capa no puede ser conectada a otra que recibe la señal antes que ella en la dinámica de la computación. Ejemplos de tales redes son el Perceptrón[27][28], Adaline y Madaline [29] y Perceptrón Multicapa (PCM) [30][31]. Una red neuronal feedforward multicapa se muestra en la Figura 5, donde las neuronas están organizadas en capas, y múltiples capas forman una red neuronal. Estas redes tienen un interés particular para nosotros en este trabajo.

Las redes feedforward usan aprendizaje supervisado, lo que quiere decir que para cada ejemplo se sabe la salida correcta (o esperada). Por lo tanto, el proceso de entrenamiento de la RNA modifica los pesos para que la salida de la red coincida con la salida esperada. El rendimiento de la red es evaluado en función de los errores (por ejemplo, el error cuadrático medio) entre la salida calculada y la esperada sobre todos los ejemplos de entrenamiento. Si la salida

correcta no se sabe (o no se usa), el entrenamiento es no supervisado, y el comportamiento de la red se evalúa con una función de distancia.

Las redes PCM, también llamadas redes backpropagation, son el tipo más conocido y usado de RNA. Estas redes usan el algoritmo backpropagation, un método de aprendizaje supervisado para enseñar a la RNA cómo hacer una determinada tarea. El algoritmo minimiza el error (comúnmente cuadrático) usando métodos basados en el gradiente descendente. Consecuentemente, una parte esencial del algoritmo es el cálculo de derivadas parciales del error de acuerdo a los parámetros (pesos) de la red neuronal. Este método fue descrito por Werbos en 1974 [30], pero alcanzó la fama cuando fue usado por Rumelhart, Hinton y Williams en 1986 [31].

- Redes neuronales recurrentes o feedback: estas redes se diferencian de las anteriores en que tienen conexiones hacia atrás, así que no se distinguen capas, y por lo tanto, la información puede volver a anteriores procesadores elementales en la dinámica de la red. Ejemplos de estas redes son ART[32], Memoria bidireccional asociativa [33], Cognitrón [34], redes de Hopfield [35], máquina de Boltzmann [36].
- Redes autoorganizadas: tienen una capa de neuronas de salida totalmente interconectadas (recurrente) conectadas a un conjunto de neuronas de entrada, con las que no hay feedback. Un ejemplo de este tipo de redes son las redes de Kohonen[37].

OTRAS TÉCNICAS SOFT COMPUTING

Sistemas Difusos[24]

Las técnicas difusas son aquellas que usan la lógica difusa como principal motor en sus razonamientos. La lógica difusa fue propuesta por primera vez por Zadeh en 1965[38]. Ésta es capaz de manejar información que es aproximada, con valores de verdad que se encuentran entre 0 y 1, es decir, entre completamente verdadero y completamente falso.

Algoritmos Genéticos[25]

Los algoritmos genéticos son heurísticas que imitan el proceso de evolución natural. Se encuentran enmarcados dentro de la computación evolutiva, ya que se basan en conceptos como selección, mutación, herencia y cruce. Generalmente se usan para obtener resultados útiles a problemas de optimización y búsqueda [39].

Modelos Ocultos de Markov (HMM)[26]

Los Modelos Ocultos de Markov son una herramienta estadística para modelar secuencias (o procesos de Markov) que puede estar caracterizadas por un proceso subyacente de parámetros desconocidos, generando una secuencia

observable. Fueron descritos por primera vez por Markov en 1913 [40], pero fue Baum quien desarrolló la teoría de los HMM en los años 60 [41].

1.4 EXTRACCIÓN Y SELECCIÓN DE CARACTERÍSTICAS

INTRODUCCIÓN

Aunque a priori la extracción y la selección de características puedan parecer definir un mismo proceso, son diferentes maneras de conseguir una reducción de características.

También llamada selección de variables o selección de subconjuntos de variables, la selección de características pretende, sobre un conjunto de variables, obtener el mejor subconjunto de variables que represente de forma óptima a la variable objetivo y que idealmente sea necesario y suficiente. Esto es, dado un conjunto de datos de entrada de N ejemplos y M características $X = \{x_i, i=1, \dots, M\}$ y la variable objetivo C , el problema de la selección de características es encontrar, en el espacio de observación de M dimensiones, R^M , un subespacio R^p que aproxime de forma óptima a C [42].

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_p \end{bmatrix}$$

La extracción de características pretende, por otro lado, encontrar una transformación de las variables en un espacio de menos dimensiones. Esta transformación puede implicar una combinación lineal o no de las variables originales y puede ser supervisada o no. Uno de los métodos más comunes usados en extracción de características es el Análisis de Componentes Principales, que se comentará más adelante.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_p \end{bmatrix} = f \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

También se puede considerar como un método de evaluación de características la ponderación de características[43], que otorga un peso determinado a cada característica. Otro caso similar es el de la discretización [44], que permite discretizar más o menos las características de forma que éstas serán tenidas en cuenta en mayor o menor medida. Ambas técnicas pueden considerarse una forma de selección o extracción de características, según que valores se les otorguen a los pesos o según el intervalo de la discretización.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \rightarrow \begin{bmatrix} w_1 \cdot x_1 \\ w_2 \cdot x_2 \\ w_3 \cdot x_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ w_m \cdot x_m \end{bmatrix}$$

Mientras que la extracción es más potente a la hora de concentrar conocimiento en un número menor de características, los beneficios potenciales de la selección de características son muchos más:

- a) Facilitar la visualización y comprensión de los datos. Disponer de un menor número de variables o características permitirá asimilar mejor los datos y obtener una representación gráfica de ellos más sencilla.
- b) Reducción de las mediciones y requisitos de almacenamiento. La obtención y toma de datos se simplificará, al ser menor el número de características necesarias para obtener un resultado. Esto tendrá un beneficio directo sobre las necesidades de almacenamiento de dichos datos.
- c) Reducción de los tiempos de utilización. Cuando un sistema ha de trabajar en tiempo real, éste se ve reforzado en cuanto a velocidad de respuesta si la cantidad de datos a manejar es menor.

- d) Obviar características. El disponer de un número de características que han sido eliminadas, dará información extra sobre el problema, indicando qué grado de relevancia tiene cada una de ellas.

RELEVANCIA DE LAS CARACTERÍSTICAS

Para entender mejor la selección de características, a continuación se describen algunos conceptos como la relevancia de características. Estas definiciones se tendrán en cuenta en los diferentes algoritmos que se comentarán más adelante.

En esta sección se va a tener en cuenta la siguiente notación:

- ∞ La entrada a un algoritmo supervisado es un conjunto de n instancias de entrenamiento.
- ∞ Cada instancia X es un elemento del conjunto $F_1 \times F_2 \times F_3 \times \dots \times F_m$, donde F_i es el dominio de la característica i -ésima.
- ∞ Dada una instancia, el valor de la característica X_i en esa instancia es x_i .
- ∞ El objetivo a predecir es Y , siendo su valor denotado por y .
- ∞ S_i será el conjunto de todas las características excepto X_i , y se denota a s_i como el valor en todas las características en S_i . Por lo tanto, $S_i = \{X_1, X_2, \dots, X_{i-1}, X_{i+1}, \dots, X_m\}$
- ∞ p es la medida de probabilidad en el espacio $F_1 \times F_2 \times F_3 \times \dots \times F_m \times Y$

Se entiende por relevancia de una característica al grado en el que ésta influye en la obtención del resultado final. En [45] los autores definen a una característica X_i como relevante para el concepto Y si X_i aparece en cada fórmula booleana que representa Y . En caso contrario, la característica es irrelevante. Este concepto de relevancia es definido suponiendo que todas las características y sus etiquetas son booleanas y que no existe ruido en los datos.

En [46], se define que una característica X_i es relevante para el objetivo Y si existe una pareja de ejemplos A y B en el espacio de ejemplos, tales que A y B difieren solamente en su asignación a X_i y $Y(A) \neq Y(B)$. Esto quiere decir es que la característica X_i es relevante si existe algún ejemplo en el espacio de instancias para el cual, si se cambia el valor de X_i , afectaría al valor del concepto objetivo. Esta definición es en teoría buena, pero si se da que en una muestra dos ejemplos tienen los mismos valores x_i para la característica X_i (es decir, que existe duplicidad en los datos), entonces sería imposible determinar que la característica X_i es relevante.

En [46][47][48][49][50][51] se detalla más sobre el estudio de los conceptos de relevancia, clasificando las características en tres categorías disjuntas: fuertemente relevantes, débilmente relevantes e irrelevantes. Además, en [46] se detalla expresamente los conceptos de relevancia para una distribución o muestra, que acota mejor el problema que se comenta en el apartado anterior. En [47] se define que X_i es fuertemente relevante si existe algún x_i, y, s_i para los que $p(X_i=x_i, S_i=s_i) > \theta$ y tales que

$$p(Y=y | X_i=x_i, S_i=s_i) \neq p(Y=y | S_i=s_i)$$

La relevancia fuerte implica que una característica es indispensable de manera que no puede ser eliminada sin pérdida de precisión en la predicción. Según esta definición, X_i es fuertemente relevante si la probabilidad de la variable objetivo Y puede cambiar si se elimina conocimiento sobre X_i .

Por otro lado, también en [47], se detalla que una característica X_i es débilmente relevante si no es fuertemente relevante y existe un subconjunto de características S_i' de S_i para el cual existe algún x_i , y y s_i' con $p(X_i=x_i, S_i'=s_i') > 0$ y tales que

$$p(Y=y | X_i=x_i, S_i'=s_i') \neq p(Y=y | S_i'=s_i')$$

Una característica débilmente relevante puede contribuir a la precisión en la predicción, algo que una característica irrelevante nunca puede hacer, por definición.

Para concluir, se consideran irrelevantes todas aquellas características que no son fuertemente ni débilmente relevantes.

Así, se interpreta el espacio de características como el formado por una serie de características irrelevantes y relevantes, dividiéndose estas últimas en fuertemente relevantes y débilmente relevantes. Según [45], la selección óptima de características comprende a todas características fuertemente relevantes y algunas de las débilmente relevantes, seleccionando estas últimas según la capacidad de predicción que se aporte con el menor número de ellas (Figura 6). Hay otros ejemplos de aproximaciones de selección de características que incluyen todas las características relevantes, tanto las fuertemente relevantes como las débilmente relevantes, como Relief [52] y RelieveD [49].

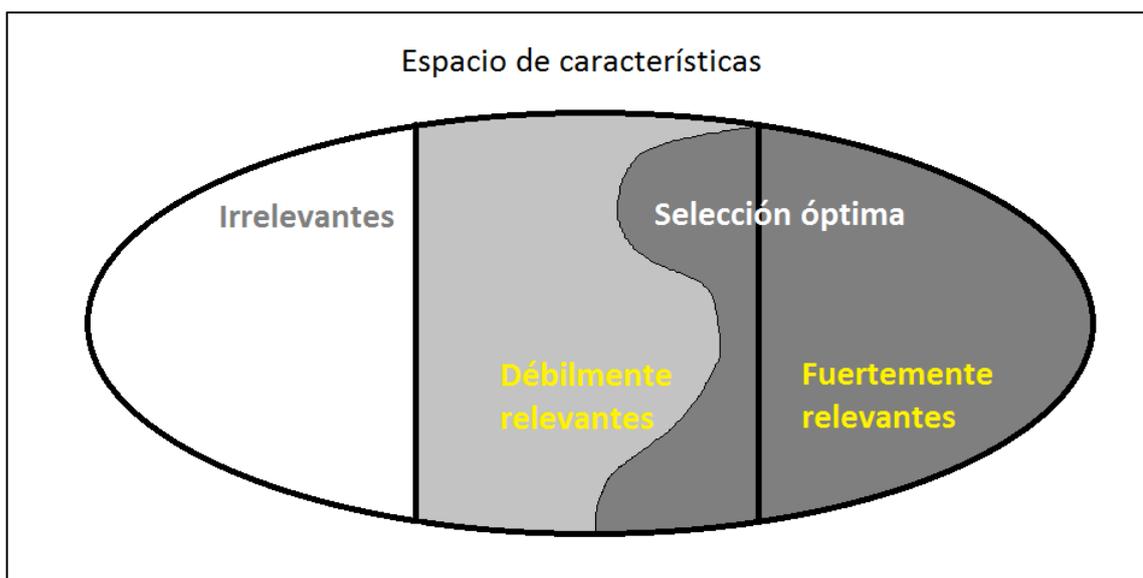


Figura 6. Representación del espacio de características según el tipo.

MÉTODOS DE EXTRACCIÓN Y SELECCIÓN DE CARACTERÍSTICAS

En esta sección se analizarán los métodos de selección y extracción de características existentes que han tenido más relevancia en el campo del aprendizaje automático y que han sido utilizados en esta memoria. Antes de pasar a comentar cada uno de los métodos que se aplicarán en este estudio, se describe brevemente el esquema general de todo método de selección de características y se procede a tipificarlos según diferentes criterios.

Todos los métodos tienen en común una serie de fases o esquema a seguir, siendo la forma de abordar cada una de estas fases lo que distingue a unos de otros. Este esquema de cuatro puntos fue introducido por Langley en [48].

- ∞ **Punto de comienzo.** Se trata de definir en el momento inicial el conjunto de características con el que se parte. Se puede partir del conjunto vacío, del conjunto de todas las características, o de un conjunto que incluya al azar o siguiendo algún criterio un número de ellas.
- ∞ **Tipo de búsqueda.** Es la forma en la que el algoritmo seleccionará nuevos ejemplos. Cabe mencionar, aunque en siguientes secciones se describirá con más detalle, la búsqueda exhaustiva, que recorrerá todas las posibles combinaciones de características, la búsqueda al azar, que elegirá características de forma aleatoria, o cualquier método que vaya seleccionando características según diferentes criterios.
- ∞ **Estrategia de evaluación.** Será el modo en que se valora cómo de buena o mala es una solución seleccionada por la búsqueda en el paso anterior. Según el modo de aplicación, que se comenta en el siguiente apartado, la estrategia de evaluación será diferente: desde el mismo algoritmo de aprendizaje final que vaya a usarse (en este caso, redes neuronales artificiales) a otras técnicas estadísticas o heurísticas.
- ∞ **Criterio de parada.** Una vez el algoritmo haya seleccionado un subconjunto de características, éste ha de detenerse según un criterio dado, siendo éste diferente dependiendo del método. Mientras que algunos se detienen en función del número de características seleccionado, otros lo hacen cuando se alcanza un valor predeterminado en un umbral [52] o simplemente cuando el tiempo de ejecución supera un valor dado [53].

A la hora de clasificar los diferentes métodos de selección de características, se tienen en cuenta diferentes criterios:

1. Según la solución

- a. Óptimos: Se llama métodos óptimos a aquellos que obtienen siempre la mejor solución entre las posibles. Un método óptimo de búsqueda en la selección de características es *Branch & Bound* [54].

- b. Subóptimos: Los métodos subóptimos no siempre obtienen la mejor solución, aunque generalmente son mejores que los óptimos computacionalmente hablando.

2. Según modo de aplicación

- a. Tipo filtro[55][56][57][58]: Separa la selección de características del algoritmo de aprendizaje, seleccionando así características independientemente del algoritmo de aprendizaje. Esta definición fue dada por John, Kohavi y Pflieger en [47]. Los métodos tipo filtro son los más usados y al ser independientes del método de aprendizaje, pueden ser usados con cualquiera de ellos. Además, se ha demostrado que es computacionalmente más rápido.
- b. Tipo envolvente(wrapper)[59][60]: Usa la precisión predictiva de un algoritmo de aprendizaje predeterminado para determinar la bondad de un subconjunto de características. Su principal ventaja es que permite obtener el subconjunto de características en el entorno real en el que serán finalmente evaluadas.
- c. Tipo inmerso (embedded): El método de selección de características está incluido en el algoritmo de aprendizaje como parte no separable.

3. Según el tipo de búsqueda (se incluye una definición de los diversos métodos de búsqueda más adelante en este capítulo).

- a. Completa o exhaustiva: Se comprueban todos los subconjuntos de variables posibles con un método de evaluación dado. Siempre encuentra la solución óptima, pero es computacionalmente el método más caro.
- b. Heurística: Búsqueda apoyada por algún tipo de conocimiento del problema (método Branch & Bound [54])
- c. Aleatoria: Búsqueda en la que cada paso siguiente se obtiene al azar entre los restantes. El método de Montecarlo[61] se considera de este tipo.

4. Según la estrategia de evaluación

- a. Evaluando las características de forma individual. Ocurre cuando un método evalúa las características según su capacidad predictiva individual. Sin tener en cuenta posibles interacciones entre características, suele obtener una lista ordenada de las mismas.
- b. Evaluando las características de forma conjunta. Tiene en cuenta la capacidad predictiva de conjuntos de características, pudiendo detectar dependencia entre variables.

5. Según orden de inclusión de características

- a. De arriba abajo (*backward*): Se van eliminando variables del conjunto inicial, que contiene la totalidad de las mismas. Generalmente, no se llegan a reelegir variables que ya han sido descartadas. Un ejemplo del uso de este tipo de métodos sería incluyendo el tipo de búsqueda “*Selección secuencial hacia atrás*” o la “*Selección secuencial hacia atrás generalizada*”.
- b. De abajo arriba (*forward*): Se parte del conjunto vacío y se van añadiendo variables. Hay variantes, en las que se pueden ir descartando variables que ya han sido seleccionadas. Un ejemplo del uso de este tipo de métodos sería incluyendo el tipo de búsqueda “*Selección secuencial hacia adelante*” o la “*Selección secuencial hacia adelante generalizada*”.

A. Análisis de Componentes Principales

El Análisis de Componentes Principales es un método de extracción de características. Se trata de un método estadístico que busca relaciones entre los datos para reducir el número de entradas (componentes del vector de entradas) en la medida de lo posible. El método, introducido por Jolliffe en [62], usa una transformación ortogonal para convertir un conjunto de características posiblemente correladas en un conjunto de valores no correlados llamados componentes principales, siendo este conjunto de menor tamaño al inicial. El método más utilizado para hacer esto es el método de la covarianza, descomponiendo en vectores propios (eigenvectores) la matriz de la covarianza. Las etapas principales que recorre este método son las siguientes:

1. Cálculo de la media de cada una de las dimensiones de los datos.
2. Cálculo de la matriz de covarianza.
3. Cálculo de los vectores y valores propios (eigenvalues) de la matriz de covarianza.
4. Elección de los componentes principales.
5. Obtención del nuevo conjunto de datos.

B. RELIEF

Presentado por Kenji Kira y Larry A. Rendell en 1992[52], este conocido método de selección de características de tipo filtro evalúa los atributos individualmente y calcula una medida estadística de cómo una característica separa dos clases. Devuelve una lista ordenada de atributos que superan un determinado umbral de relevancia. En un esfuerzo por reducir el coste computacional de calcular esta estadística, los ejemplos para hacer estos cálculos son seleccionados al azar, para lo que usa el algoritmo de Montecarlo. Este método evita la búsqueda heurística, requiriendo un tiempo lineal en el número de características dadas y el número de instancias de entrenamiento, independientemente del concepto objetivo a ser aprendido. Aunque no encuentra necesariamente el menor conjunto de

características, el tamaño tiende a ser pequeño pues sólo son seleccionados los atributos estadísticamente relevantes.

En Relief, una instancia es representada mediante un vector compuesto de p valores de características, denotando \check{S} al conjunto de instancias de entrenamiento de tamaño n . F denota al conjunto de características dadas $\{f_1, f_2, \dots, f_p\}$. Una instancia X es por tanto representada mediante un vector p -dimensional (x_1, x_2, \dots, x_p) donde x_j denota el valor de la característica f_j de X . Relief actúa muestreando repetidamente una instancia y considerando el valor del atributo dado para la instancia más cercana de la misma clase (*near-hit*) y de una clase diferente (*near-miss*). Ambos valores son calculados mediante la distancia Euclídea p -dimensional. Una vez obtiene las tres instancias (una instancia, su *near-hit* y su *near-miss*), utiliza una medida estadística para calcular cómo una característica individual es capaz de separar dos clases. De esta forma obtiene un vector de pesos por cada característica denominado "Relevancia", de forma que las variables cuya relevancia superan un umbral son propuestas.

```

Relief ( $\check{S}, m, \tau$ )
  Separar  $\check{S}$  en  $\check{S}^+ = \{\text{instancias positivas}\}$  y  $\check{S}^- = \{\text{instancias negativas}\}$ 
   $W = (0, 0, \dots, 0)$ 
  Para  $i=1$  hasta  $m$ 
    Elige al azar una instancia  $X \in \check{S}$ 
    Elige al azar una instancia positiva cercana a  $X$ ,  $Z^+ \in \check{S}^+$ 
    Elige al azar una instancia negativa cercana a  $X$ ,  $Z^- \in \check{S}^-$ 
    Si  $X$  es una instancia positiva
      Entonces near-hit =  $Z^+$ , near-miss =  $Z^-$ 
      Sino near-hit =  $Z^-$ , near-miss =  $Z^+$ 
    Fin Si
    actualiza-peso( $W, X, \text{near-hit}, \text{near-miss}$ )
  Fin Para
  Relevancia =  $(1/m)W$ 
  Para  $i=1$  hasta  $p$ 
    Si relevancia $_i \geq \tau$ 
      Entonces  $f_i$  es una característica relevante
      Sino  $f_i$  es una característica irrelevante.
    Fin Si
  Fin Para
Fin Relief

actualiza-peso( $W, X, \text{near-hit}, \text{near-miss}$ )
  Para  $i=1$  hasta  $p$ 
     $W_i = W_i - \text{diff}(x_i, \text{near-hit}_i)^2 + \text{diff}(x_i, \text{near-miss}_i)^2$ 
  Fin Para
Fin actualiza-peso

```

Dados unos datos de entrenamiento \check{S} , un tamaño de la muestra m y un umbral de relevancia τ , Relief detecta aquellas características que son estadísticamente relevantes al concepto objetivo. Se asume que la escala de cada

características es nominal, booleana o numérica (entero o real). Las diferencias entre los valores de cada característica de dos instancias X e Y se calculan, para valores numéricos, que es el caso que aquí se trata, de la siguiente manera:

$$diff(x_k, y_k) = (x_k - y_k) / nu_k$$

Las diferencias entre los valores de cada característica de dos instancias X e Y se calculan, para valores booleanos y nominales, de la siguiente manera:

$$diff(x_k, y_k) = \begin{cases} 0 & \text{si } x_k \text{ e } y_k \text{ son iguales} \\ 1 & \text{si } x_k \text{ e } y_k \text{ son diferentes} \end{cases}$$

El algoritmo Relief para valores booleanos o nominales elige una muestra de un número m de tripletas compuestas de una instancia X, su vecino más cercano con un valor igual del concepto objetivo (*near-hit*) y su vecino más cercano con un valor diferente del concepto objetivo (*near-miss*).

En [52], Relief es comparado con otros dos métodos, FOCUS[63] e ID3[64] en la aplicación a dos conocidos problemas (representación de números en pantalla LED y dominio de paridad). En ambos casos Relief representa una mejora en tiempo de aprendizaje y precisión predictiva del concepto objetivo.

Aunque inicialmente Relief sólo era válido para problemas de clasificación de dos clases, posteriormente se han propuesto diferentes extensiones para hacerlo extensible a otra clase de problemas de clasificación. Kononenko propone extensiones de RELIEF en [65]. En esta aportación, las diferentes propuestas (desde Relief-A hasta Relief -F) permiten a RELIEF manejar problemas con más de dos clases y con datos que incluyan ruido o estén incompletos.

C. Ganancia de Información

Este método, incluido en Weka[66] e implementado por Mark Hall [67], evalúa el valor de un atributo midiendo la ganancia de información respecto a la clase:

$$Ganancia_Información(clase, atributo) = H(clase) - H(clase | atributo)$$

siendo $H(clase)$ la entropía (incertidumbre) de la clase y $H(clase | atributo)$ la entropía de la clase conocido el valor del atributo.

D. Ratio de la Ganancia

Evalúa el valor de un atributo midiendo el ratio de ganancia con respecto a la clase:

$$Ratio_Ganancia(clase, atributo) = (H(clase) - H(clase | atributo)) / H(atributo)$$

siendo, al igual que en el caso de la ganancia de la información, $H(clase)$ la entropía (incertidumbre) de la clase y $H(clase | atributo)$ la entropía de la clase conocido el valor del atributo.

E. Selección de subconjuntos de características basadas en correlación

Evalúa un subconjunto de atributos teniendo en cuenta la habilidad individual de predicción de cada característica, junto con el grado de redundancia entre ellas. Este algoritmo ha sido descrito en [57], y el autor expone que su motivación fue crear un algoritmo que fuera capaz de trabajar con problemas continuos y discretos, ya que la mayoría de métodos existentes de tipo filtro tratan problemas solamente discretos.

Se eligen preferentemente subconjuntos de atributos que están altamente correlacionadas con la clase y cuya intercorrelación sea baja, siendo mejor en ocasiones a Relieff, uno de los métodos con los que en [57] es comparado.

F. Evaluación de la consistencia de los subconjuntos de características

Se trata de un método tipo filtro que evalúa el valor de un subconjunto de atributos mediante el nivel de consistencia en los valores de la clase cuando los ejemplos de entrenamiento se proyectan en un subconjunto de atributos.[68]

Parte de la base de que la consistencia de cualquier subconjunto no puede ser nunca menor que la del conjunto completo de atributos, de ahí que la práctica habitual sea utilizar este método en conjunción con una búsqueda aleatoria o exhaustiva, lo cual busca el subconjunto más pequeño con consistencia igual a la del conjunto completo de los atributos.

G. Método de mínima redundancia – máxima relevancia

Este método fue desarrollado en el año 2003 por Hanchuan Peng, Fuhui Long y Chris Ding y se explica ampliamente en [42]. Debido a la dificultad de aplicar directamente la condición de máxima dependencia, se implementa un criterio equivalente llamado “minimal-redundancy – maximal-relevance criterion (mRMR)”, para una selección de características incremental inicial. Se presenta un algoritmo de dos etapas de selección mediante la combinación de mRMR y otros selectores de características más específicos (por ejemplo, *wrappers*).

En [42] se lleva a cabo una extensa comparación experimental del algoritmo y otros métodos que utilizan tres diferentes tipos de clasificadores (Bayes-Naive, máquina de vector soporte, y el análisis discriminante lineal) y cuatro conjuntos de datos diferentes. Los autores afirman que los resultados confirman que mRMR implica una mejora en la selección de características y en la precisión de la clasificación.

Existen dos esquemas destacables de este método, MID y MIQ, Mutual Information Difference y Quotient, respectivamente. MIQ es más preciso y es definido ampliamente en [69].

H. Tipo Wrapper con conjunto de entrenamiento

Evalúa los subconjuntos de atributos en los datos de entrenamiento o en un conjunto de prueba independiente, utilizando un clasificador. Se considera de interés

aplicar un método de selección de características tipo Wrapper en el que el clasificador sea, como se viene utilizando, una red *neuronal*. [49]

I. Tipo Wrapper con validación cruzada

La validación cruzada o cross-validation, es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. Consiste en repetir y calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar cómo de preciso es un modelo que se llevará a cabo en la práctica [70][49].

En esta tesis se usa el método “validación cruzada de K iteraciones” (o K-fold cross-validation) [71]. En él, los datos de muestra se dividen en K subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba y el resto (K-1) como datos de entrenamiento. El proceso de validación cruzada es repetido durante k iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado. Este método es muy preciso puesto que se evalúa a partir de K combinaciones de datos de entrenamiento y de prueba, pero aún así tiene una desventaja, y es que, a diferencia del método de retención, es lento desde el punto de vista computacional. En la práctica, la elección del número de iteraciones depende de la medida del conjunto de datos. Lo más común es utilizar la validación cruzada de 10 iteraciones (10-fold cross-validation).

MÉTODOS DE BÚSQUEDA USADOS EN SELECCIÓN DE CARACTERÍSTICAS

A continuación se describen los métodos de búsqueda más característicos utilizados en los algoritmos de selección de características. Todos ellos han sido utilizados al menos en una ocasión en los métodos de selección de características cuya aplicación se acaba de comentar.

A. Búsqueda exhaustiva

Consiste en explorar todo el espacio de búsqueda sin excepciones. Esta es la más costosa de todas computacionalmente, pero permite obtener siempre una solución óptima.

Se basa en la evaluación del criterio de optimalidad para todas posibles combinaciones de variables y seleccionar la combinación que contribuye mejor a la discriminación de la variable objetivo c . El número de posibles subconjuntos sería:

$$n_d = \frac{p!}{(p-d)! d!}$$

Si tuviéramos 25 variables ($p=25$) y quisiéramos encontrar la combinación de 15 variables ($d=15$) que mejor obtiene c , deberíamos probar entre 3.268.760 combinaciones de las 15 variables.

B. Búsqueda al azar

Búsqueda aleatoria en el espacio de subconjuntos de atributos. Si no se define un punto de partida, se comienza con un grupo aleatorio y devuelve el mejor conjunto encontrado. Si se proporciona un conjunto de partida, este tipo de búsqueda elige sólo los subconjuntos iguales o mejores que el inicial.

La búsqueda al azar es útil cuando hay muchas maneras diferentes de proceder, pero es difícil determinar una buena entre ellas. También es útil cuando en problemas de selección de características determinísticos el algoritmo se queda en un óptimo local.

Los métodos al azar de búsqueda son de dos tipos: Las Vegas (siempre da una salida correcta, pero invierten más tiempo en ella) y Montecarlo (pueden dar una salida incorrecta, pero la ejecución es más rápida).

C. Primero el mejor

Busca el espacio de subconjuntos de atributos a través del método Greedy mejorado con posibilidad de backtracking. Para controlar el nivel de backtracking, se configura el número permitido de nodos consecutivos que no mejoran el resultado. Puede empezar con el conjunto vacío y buscar hacia adelante, o empezar con el conjunto de todos los atributos y buscar hacia atrás, o incluso empezar en cualquier punto intermedio y buscar en ambos sentidos (considerando todas las adiciones y sustracciones de atributos individuales posibles en un punto dado).

D. Mejor N individual

Dado que valora cada característica de forma individual, es decir, es un método escalar, es el más simple y el que peor resultado obtiene normalmente. Se le da un valor individual a cada característica y se ordenan según este valor:

$$J(x_1) \geq J(x_2) \geq \dots \geq J(x_p)$$

Posteriormente se selecciona el número de características que se desee que mejor valor individual tengan. Este método funciona especialmente bien cuando las variables no están correlacionadas.

E. Selección secuencial hacia adelante (SFS) y hacia adelante generalizada (GSFS)

También llamado el método de adición de conjuntos[72], es un procedimiento de búsqueda de abajo arriba que añade variables a un conjunto inicial, generalmente vacío. En cada iteración se añade la variable que maximiza el criterio de selección. El algoritmo se detiene cuando no hay ninguna variable que mejore el criterio de selección o cuando se ha llegado al máximo número de variables seleccionado.

La selección secuencial hacia adelante generalizada es similar al método SFS, pero en vez de ir añadiendo variable individualmente, se añade a las d variables ya elegidas, un número r de variables en cada iteración entre las $n-d$ variables restantes. Para elegir estas r variables, se evalúa cada grupo de r variables posible:

$$\binom{n-d}{r}$$

Se elige el que maximiza la función de coste. Aunque implica más coste computacional, se obtienen mejores resultados que en la SFS ya que tiene en cuenta la posible relación estadística entre variables.

F. Selección secuencial hacia atrás (SBS) y hacia atrás generalizada (GSBS)

Es el método análogo a SFS, pero haciendo una aproximación de arriba abajo[73]. Se parte del conjunto de todas las variables y se van eliminando una por una hasta que queda solamente el número de variables deseado. La principal desventaja respecto de SFS es que es computacionalmente más costoso, ya que se evalúan grandes conjuntos de variables.

La selección secuencial hacia atrás generalizada es a SBS, lo que GSFS es a SFS. Es decir, elimina grupos de variables del conjunto inicial, que contiene todas las variables, hasta obtener el conjunto de variables deseado.

G. Branch and Bound

Es un proceso de arriba abajo óptimo y no implica búsqueda exhaustiva [54]. Se empieza construyendo un árbol en cuyo nodo superior se consideran todas las variables y en cada rama se van quitando variables sucesivamente. Se basa en la siguiente propiedad de la selección de características:

$$Si X \subset Y \Rightarrow J(X) < J(Y)$$

Esto quiere decir, que si un conjunto de variables X (por ejemplo, las variables 1, 4, 5 y 6) está contenido en otro conjunto Y de variables Y (las variables 1,2,3,4,5 y 6), el valor de la función a optimizar $J(X)$ será menor que $J(Y)$. De esta forma pueden dejarse inexploradas ramas del árbol si ya se tiene un valor mejor de la función J con un menor número de variables en otra rama.

H. Otros métodos de búsqueda utilizados

- **Selección “añadir l - eliminar r ”:** Es similar a SFS y SBS, permitiendo, a su vez, algo de *backtracking*[74]. Se añaden l variables al conjunto usando SFS (partiendo de un conjunto vacío) y posteriormente las peores r variables se eliminan usando SBS.
- **Selección generalizada “añadir l - eliminar r ”:** Como el método anterior, permite *backtracking* añadiendo l variables al conjunto y eliminando r variables, pero usando los métodos GSFS y GSBS.

- **Método de búsqueda flotante:** Se llama métodos de búsqueda flotante a la selección secuencial flotante hacia adelante (SFFS) y la selección secuencial flotante hacia atrás (SBFS)[75]. Se trata de métodos que aplican la selección “añadir l – eliminar r ” en los que se permite que los valores de l y r cambien durante el proceso. La palabra “flotante” se refiere precisamente al hecho de que l y r cambien de valor.
- **Subset Size Forward Selection:** Se trata de una extensión de la selección secuencial hacia adelante. Lleva a cabo una validación cruzada interior, donde se hace una búsqueda lineal hacia adelante en cada uno de los grupos, con el objetivo de determinar el tamaño óptimo del subconjunto de características. Finalmente, se realiza otra selección secuencial hacia adelante con el grupo elegido y todos los datos.
- **Scatter Search:** Realiza una búsqueda dispersa a través del espacio de subconjuntos de atributos[76]. Empieza con una población de características significativas y diversas y para cuando el resultado es mejor que un umbral dado o no hay más mejora.
- **Rank Search:** Utiliza un evaluador de atributos o de subconjuntos de ellos para hacer un ranking de todos los atributos[77].
- **Ranker:**Elabora un ranking de los atributos mediante la evaluación individual de los mismos. Este tipo de búsqueda se usa en conjunción con evaluadores de atributos individuales, como ReliefF o GainRatio.
- **Race search:** Mide el error de la validación cruzada de diversos subconjuntos de atributos[78]. Se usa conjuntamente con un clasificador de subconjuntos.
- **Greedy Stepwise:** Lleva a cabo una búsqueda greedy hacia adelante o hacia atrás a través del espacio de atributos. Puede comenzar con un conjunto de todos los atributos, ninguno o cualquier subconjunto de ellos arbitrario, y para cuando la suma o extracción de algún atributo no resulta en una mejora.
- **Búsqueda genética:** Búsqueda heurística que imita el proceso de la evolución natural [39], generando soluciones a problemas de optimización usando técnicas como cruce, herencia, mutación y selección.

1.5 SOFTWARE

En esta tesis se ha utilizado diverso software para apoyar los contenidos y experimentos desarrollados. A continuación se detallan las aplicaciones y códigos utilizados.

WEKA

Weka es un entorno software de minería de datos y aprendizaje automático [66] desarrollado por la Universidad de Waikato, Nueva Zelanda. En él se encuentran implementados la mayoría de métodos de Selección de Características utilizados en este trabajo.

MATLAB

Esta herramienta de software matemático [79] se ha utilizado para implementar la Red Neuronal Artificial desarrollada en este trabajo y descrita en el capítulo III, así como para la ejecución del método de Análisis de Componentes Principales del capítulo IV.

MRMR

Método de Selección de Características cuyo código desarrollado en C++ y Matlab está disponible en [80].

**CAPÍTULO II: TAXONOMÍA DE LAS
APLICACIONES DE TÉCNICAS SOFT
COMPUTING EN EL RUIDO
AMBIENTAL**

Cuando se aborda una actividad de investigación que se pueda considerar novedosa, la primera tarea a realizar debe ser, sin duda, la revisión de la literatura científica para encontrar trabajos relacionados, de una forma u otra, con los objetivos que se pretenden conseguir.

Si en la revisión se encuentra un gran número de trabajos ya publicados, se suele reducir el ámbito de análisis. Por el contrario, si hay pocos trabajos publicados, se puede ampliar un poco el ámbito de la búsqueda.

En muchos casos es habitual o una buena contribución, resumir la búsqueda realizada, ya que ayuda a futuros investigadores, que necesitarán ya menos tiempo para la búsqueda de publicaciones relacionadas.

En este capítulo se aborda un repaso del estado del arte sobre el uso de técnicas de Soft Computing, especialmente Redes Neuronales Artificiales, en el estudio y análisis de ruido ambiental, más concretamente ruido urbano. Aunque inicialmente se consideró el estado del arte sobre Ruido y RNA, se ha ampliado el

estudio a técnicas de Soft Computing, porque se han encontrado pocas contribuciones.

Para llevar a cabo esta revisión, en primer lugar se incluye una taxonomía en la que se relacionan las referencias a los diferentes trabajos encontrados. Según el tratamiento del ruido y la técnicas Soft Computing empleados, se han encontrado un número mayor o menor de artículos, lo que ha llevado a conocer qué campos están más o menos explorados, y cuáles son los puntos en los que las investigaciones en este ámbito deberían incidir.

Esta revisión de la literatura actual se realiza dividiendo las publicaciones encontradas, por un lado, según la técnica Soft Computing que usan, y por otro, en tres tipos según el tratamiento del ruido ambiental. También se incluyen trabajos sobre la predicción del flujo de tráfico, al ser éste uno de los parámetros más influyentes en el ruido ambiental urbano.

Tipo de Análisis del Ruido	Técnica de Soft Computing
Predicción del Nivel de Ruido	Redes Neuronales
Predicción del Grado de Molestia	Algoritmos Genéticos
Clasificación del Ruido	Técnicas Difusas
	Modelos Ocultos de Markov

La descripción más detallada de los trabajos encontrados, se ha limitado, por cuestiones de interés, a los trabajos que utilizaban RNA. Se ha realizado una descripción más amplia cuando el trabajo usa RNA sobre el problema de la Predicción del Nivel de Ruido, que forma el tronco básico del trabajo propuesto.

2.1 SOFT COMPUTING Y ANÁLISIS DEL RUIDO

En la revisión de la literatura se han buscado especialmente trabajos que utilicen técnicas de Soft Computing para realizar cualquiera de los tratamientos que generalmente recibe el ruido ambiental. El principal es, de acuerdo con los objetivos de esta memoria, la predicción del nivel de ruido, pero también se consideran otros aspectos del ruido, como la molestia que provoca en las personas o la clasificación de los tipos de ruido que son generados. También se analiza el estudio del flujo de tráfico, principal causante del ruido ambiental, sobre todo urbano. A continuación se describen brevemente cada uno de los tratamientos mencionados.

Predicción del nivel de ruido: Se entiende la forma de obtener a priori el nivel de presión sonora, dadas unas características del entorno.

Predicción del grado de molestia: Como se ha visto en el capítulo I, el ruido es el sonido no deseado. Esta calidad de no deseado implica un grado de molestia en los seres humanos, que puede llegar a causar alteraciones en el sueño, ansiedad, y otros estados adversos de diversa consideración. La predicción del grado de molestia trata de dar a priori el nivel de desagrado que las personas que sufren un determinado nivel de ruido van a experimentar.

Clasificación del ruido: Se pretende obtener una tipificación del ruido, distinguiendo, por ejemplo, entre sonidos de automóviles, de claxon, de ambulancias, etc.

De estos tres aspectos del análisis del ruido se han encontrado 41 trabajos diferentes que son abordados desde cuatro técnicas Soft Computing: Redes Neuronales Artificiales, Técnicas Difusas, Algoritmos Genéticos y Modelos Ocultos de Markov.

Predicción de tráfico: Es una forma de dar a conocer, con anterioridad, el flujo de automóviles que existirá en un determinado punto del núcleo urbano o en vías fuera del territorio metropolitano, a partir de unos parámetros del entorno.

La predicción del flujo de tráfico con técnicas Soft Computing aporta en nuestra taxonomía otros 15 trabajos adicionales.

Como paso previo a una descripción detallada de las publicaciones, en la Tabla 5 se muestra una taxonomía de las mismas. Se pretende con ello dar una visión general de lo que se va a comentar.

La mayor parte de trabajos usan redes neuronales artificiales o técnicas difusas en la predicción del grado de molestia, del nivel de ruido y de flujo de tráfico. En cambio, para clasificación, los modelos ocultos de Markov han sido los más usados. Se han encontrado escasas referencias a algoritmos genéticos, y éstas están casi en su totalidad relacionadas con la predicción del nivel de presión sonora. Concretamente en cuanto a redes neuronales, se encuentran referencias para todos los ámbitos, especialmente en la predicción del nivel de ruido y de flujo de tráfico.

En las siguientes secciones se describen los trabajos relacionados, detallando los que aplican redes neuronales a la predicción de ruido, que es el tema de interés prioritario en este trabajo.

Tabla 5. Taxonomía de las publicaciones encontradas

		Predicción del nivel de ruido	Predicción del grado de molestia	Clasificación	Predicción del flujo de tráfico urbano
Técnica Soft Computing	Redes Neuronales Artificiales	[81], [82], [83], [84], [85], [86], [87], [88]	[89], [90], [91], [92], [93]	[94], [95], [96], [97]	[98], [99], [100], [101], [102], [103], [104], [105], [106], [108], [109], [110]
	Técnicas difusas	[111], [112], [113], [114], [115], [116]	[91], [114], [117], [118], [119], [120], [121], [122], [123], [124], [125], [126], [127]	[128], [129]	[98], [102], [103], [107], [108], [130], [131]
	Algoritmos Genéticos	[113], [117], [121], [126], [132], [133]			[104]
	Modelos Ocultos de Markov			[96], [97], [134], [135], [136], [137]	

2.2 REDES NEURONALES ARTIFICIALES Y RUIDO URBANO

Las redes neuronales se han utilizado hasta el momento en los tres aspectos del análisis de ruido urbano considerado. Se han encontrado 8 trabajos para la predicción del nivel de ruido, 5 sobre la molestia del ruido y 4 para clasificar ruidos. Además, otros 13 trabajos relacionan las redes neuronales con la predicción del flujo de tráfico.

Ahora se describen, con cierto detalle, los trabajos encontrados, según el aspecto de análisis del ruido a que se aplican y en el orden de su publicación.

A. PREDICCIÓN DEL NIVEL DE RUIDO

En [81] se define un sistema de predicción de ruido urbano usando sobre todo redes neuronales, que describe un completo modelo de predicción de ruido basado en una red neuronal tipo PMC. Usando solamente tres variables de entrada (número de vehículos, altura de los edificios y anchura de la calle) y 9 neuronas en la capa oculta, se construye una RNA que sigue la estructura de la Figura 7.

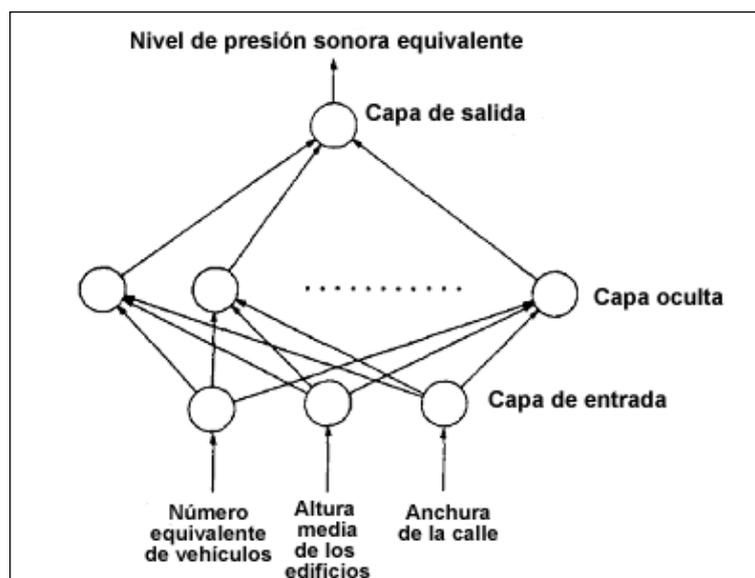


Figura 7. Estructura del modelo PMC de [81]

Para evaluar los resultados, se considera que el máximo error permitido es de ± 2.5 dB. Los resultados de este primer test se incluyen en la Tabla 6, donde se observa que, en los datos de Catania, el valor del nivel de presión sonora medio para el que la identificación fue positiva es de aproximadamente 80 dB, mientras que el valor para predicción negativa fue de 76 dB. Debido a esto, la predicción sólo ha sido positiva en un subconjunto de los datos de carreteras utilizados, precisamente cuando el nivel de presión sonora a ser identificado es alto.

Tabla 6. Resultados de la primera prueba de [81]

TABLE 1a			
City of Palermo			
Correct Prediction	Range	Average	Variance
Road Width	10–22 meters	16.367 meters	0.001856
Noise Measured	60.9–82.5 db	72.7 db	0.002805

TABLE 1b			
City of Palermo			
Wrong Prediction	Range	Average	Variance
Road Width	3–28.6 meters	14.7 meters	0.003365
Noise Measured	55.3–89.2 db	70 db	0.003188

TABLE 2a			
City of Catania			
Correct Prediction	Range	Average	Variance
Road Width	10.5–20 meters	14.23 meters	0.001678
Noise Measured	76–81.7 db	79.8 db	0.000404

TABLE 2b			
City of Catania			
Wrong Prediction	Range	Average	Variance
Road Width	6.5–28 meters	13.18 meters	0.002064
Noise Measured	63–86.3 db	75.8 db	0.000696

Dado que no se consideran resultados satisfactorios, se separa la variable de número de vehículos en tres: número de automóviles, número de motocicletas y número de camiones. De esta forma se obtiene un modelo con 20 neuronas en la capa oculta, que da mejores resultados. Para comprobar la mejoría respecto de la primera aproximación, se hicieron pruebas con datos muy similares a los anteriores. Los datos para esta segunda prueba se incluyen en la Tabla 7.

Tabla 7. Resultados de la segunda prueba de [81]

TABLE 3			
City of Palermo	Range	Average	Variance
Road Width	3–28.6 meters	15.9 meters	0.002168
Noise Measured	58.9–82.5 db	73.7 db	0.002101

TABLE 4			
City of Catania	Range	Average	Variance
Road Width	6.5–20 meters	13 meters	0.001427
Noise Measured	72.2–84.2 db	78 db	0.000434

Para demostrar la bondad de este nuevo modelo, se consideran varias zonas de Italia y se comparan las salidas del mismo con las de modelos clásicos como Burgess[138], Josse[139] y CSTB[140]². Puede verse en la Figura 8 uno de los ejemplos, con objeto de mostrar la calidad del resultado.

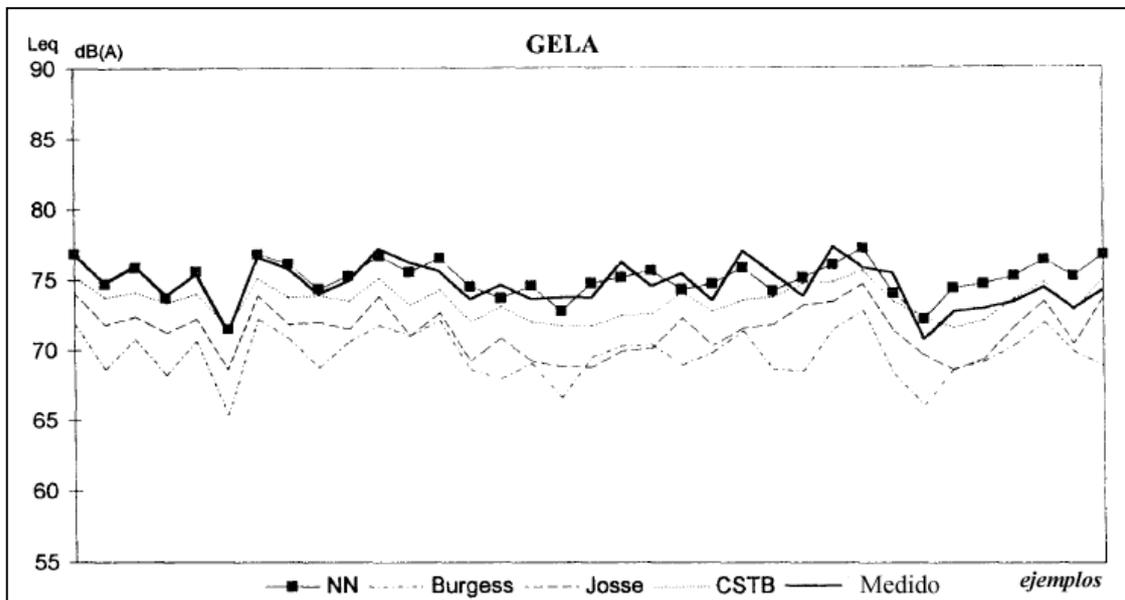


Figura 8. Resultados de la red neuronal de [81]

² Estos modelos no se han indicado en el capítulo I porque no se usan en la actualidad.

Un límite de este modelo neuronal es su dependencia de las medidas acústicas, ya que los sonómetros son colocados junto a la carretera y captan algo más que sonidos de los vehículos (sirenas u otros eventos que no corresponden con el tráfico normal), y las variables de entrada consideradas no tienen en cuenta estos fenómenos. Por lo tanto, al calcular el nivel sonoro equivalente, se obtendría un valor superior al real, causado por otro tipo de elementos que no son los considerados (número de vehículos, altura de los edificios y anchura de la calle). Por esto, se añade una red tipo LVQ (Learning Vector Quantization) a modo de filtro de los datos.

Así, se propone una arquitectura basada en 2 niveles: en el primero, la red LVQ filtra los datos, eliminando las medidas que son consideradas erróneas, mientras que en el segundo nivel, la red BPN predice el nivel de presión sonora. La arquitectura completa puede observarse en la Figura 9.

El funcionamiento del filtro LVQ fue probado con datos deliberadamente alterados, y los resultados fueron muy buenos. En la gráfica de la Figura 10 se representan los ejemplos en el eje de abscisas y la diferencia en decibelios (normalizado de 0 a 10, con lo que 0.25 serían 2.5 dB) entre el valor de nivel sonoro real y alterado en el eje de ordenadas. La función escalón está en 0, cuando el filtro ha dado el nivel sonoro por correcto, y en 1, cuando lo considera incorrecto. La otra curva representa la diferencia entre el L_{eq} real y el alterado. Es claro que el filtro es útil, pues datos de L_{eq} medidos (los alterados en este caso) que distan más de 3dB del valor real son desechados por el filtro.

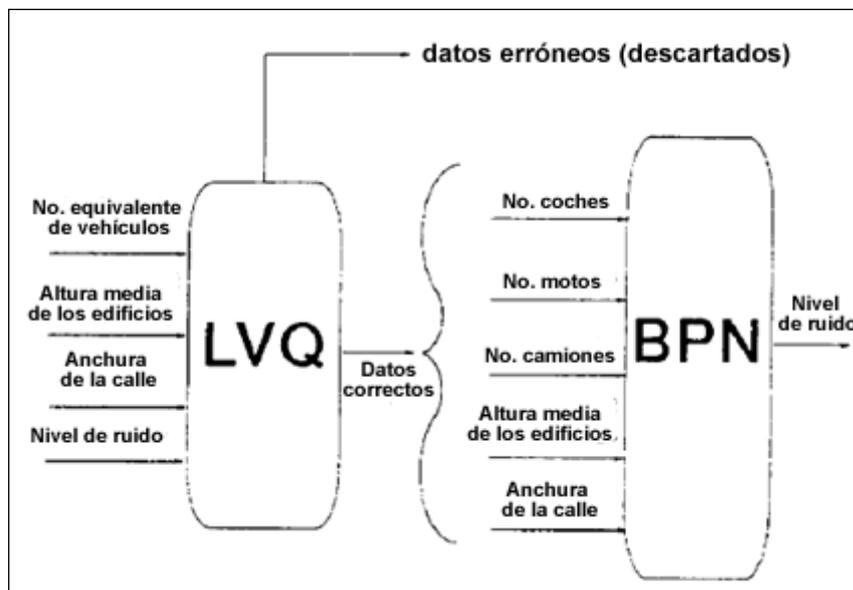


Figura 9. Arquitectura final del modelo [39]

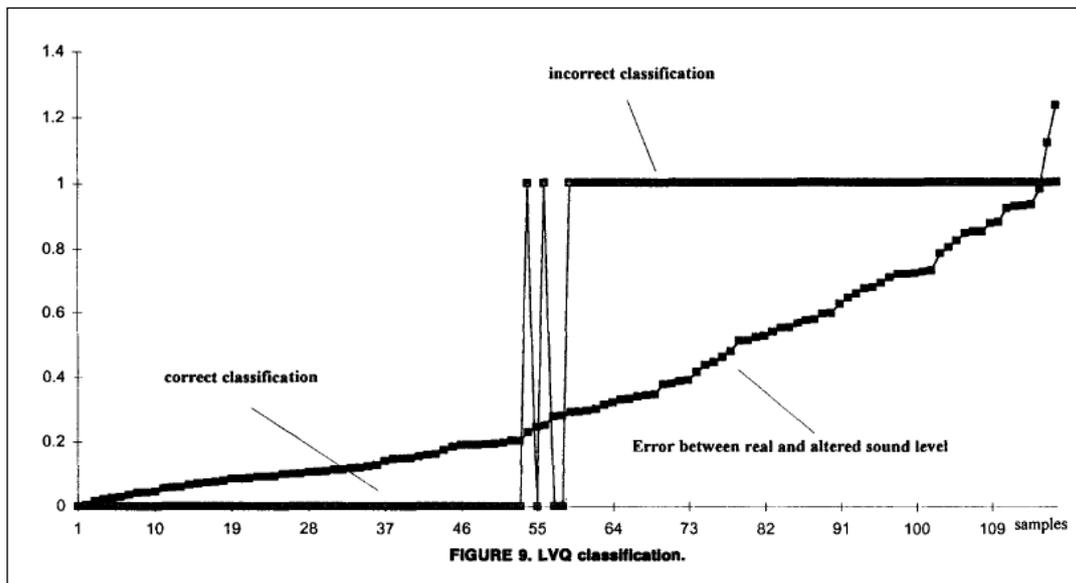


Figura 10. Clasificación del filtro LVQ de [39]

Usando una red neuronal tipo “perceptrón multicapa” y con tomas de ruido en 16 puntos diferentes en un campus universitario de Turquía, en [82] se estudia si se cumple la normativa turca que permite un máximo de 45 dB en lugares con ambiente de estudio y clases. Se construye una red neuronal tipo perceptrón multicapa con siete entradas: posición del punto de medida, distancia de la fuente al punto de medida, velocidad y dirección del viento, temperatura del aire, humedad relativa y hora del día. La salida es el nivel de ruido equivalente.

Se obtuvo un total de 319 patrones de datos, de los cuales la mitad fueron usados para el conjunto de entrenamiento y la otra mitad para el conjunto de test. Los resultados fueron positivos según los autores, ya que en los conjuntos de test y entrenamiento se obtuvieron niveles de presión sonora bastante similares a los medidos. Esto puede observarse en las Figuras 11 y 12.

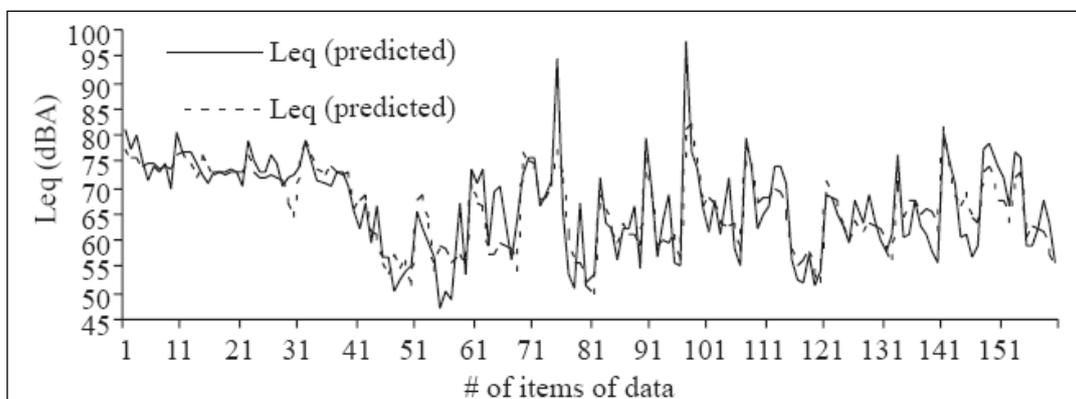


Figura 11. Resultados del conjunto de entrenamiento de [82]

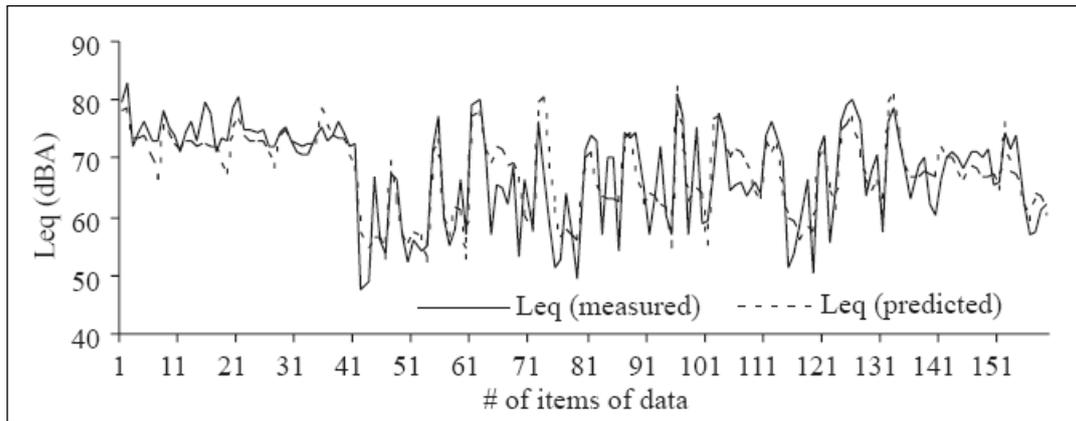


Figura 12. Resultados de conjunto de test [82]

Si se representa gráficamente tanto el nivel de ruido medido y el calculado conjuntamente, se aproximan linealmente ambas medidas según las gráficas de la Figura 12. Como observamos, se obtiene un coeficiente de correlación de 0.72 y 0.68 para el conjunto de entrenamiento y el de test, respectivamente. También se ve en la Figura 13, que la red se comporta algo peor para los niveles de L_{eq} altos, por encima de 80 dB.

En la Figura 14 se puede ver la distribución de los datos con los que se hicieron las pruebas en [82]. El número de datos de más de 80 dB es muy bajo, hecho al que se debe que la red no sea capaz de predecir correctamente estos ejemplos.

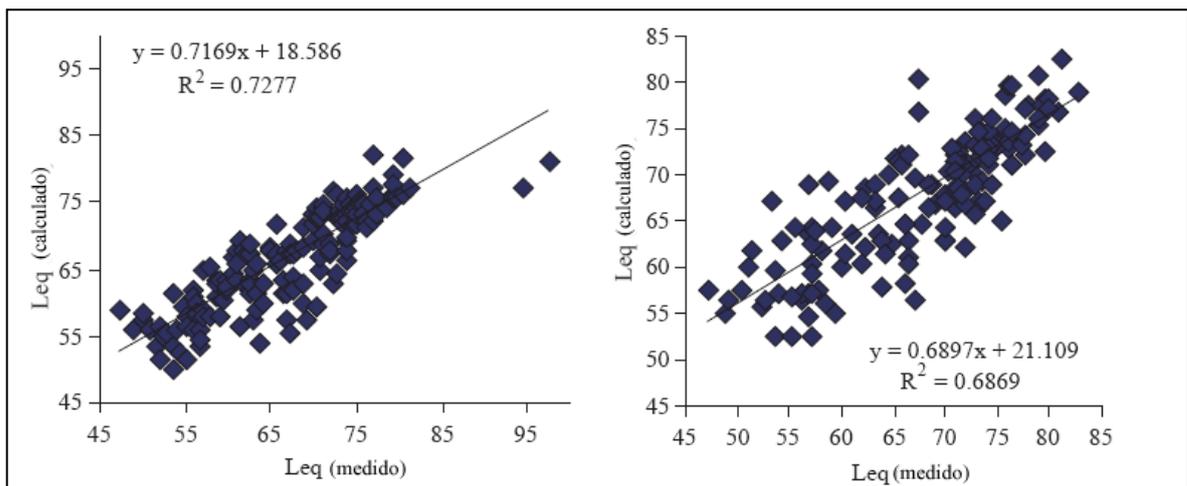


Figura 13. Comparación nivel sonoro medido - calculado de [82]

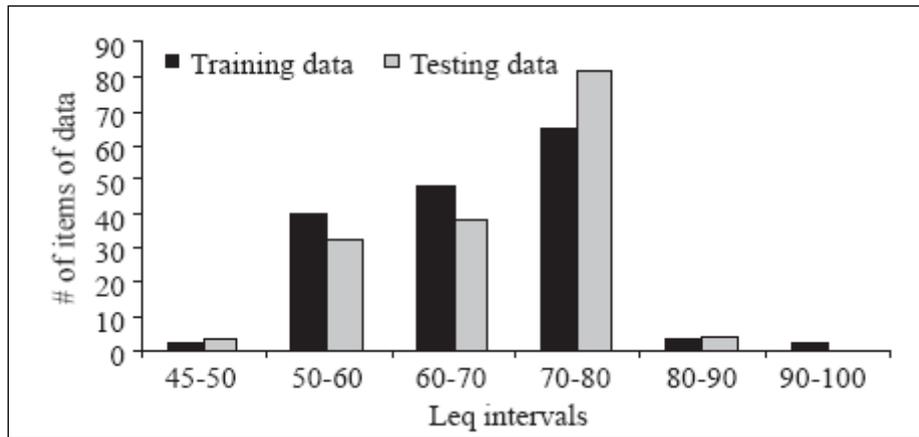


Figura 14. Distribución de los datos de [82]

Como conclusión, en [82] se expone la bondad de las redes neuronales para este tipo de problemas, pero se menciona que es necesaria una base de datos de ejemplos más amplia para aprovechar al máximo el potencial de esta técnica.

En [83], los autores han desarrollado una red neuronal multicapa con el algoritmo backpropagation. Las entradas que el modelo recibe son: composición del tráfico (buses, camiones, motocicletas, bicicletas, etc.), anchura de la calzada y distancia del receptor de ruido a la calzada. Las salidas son cinco: L_{eq} , L_{10} , nivel de polución del sonido (LNP), índice del ruido de tráfico (TNI) y ruido climático (NC). Para comprobar el resultado, se muestran varias gráficas con el análisis de regresión lineal de los resultados según el valor analizado y el tipo de flujo de tráfico. Éstos varían mucho, obteniendo un coeficiente de regresión lineal de entre 0.47 (para L_{10} en flujo de tráfico ininterrumpido) y 0.91 (para LNP en flujo de tráfico interrumpido, Figura 16), aunque se mueve generalmente en valores alrededor de 0.80 (Figura 15).

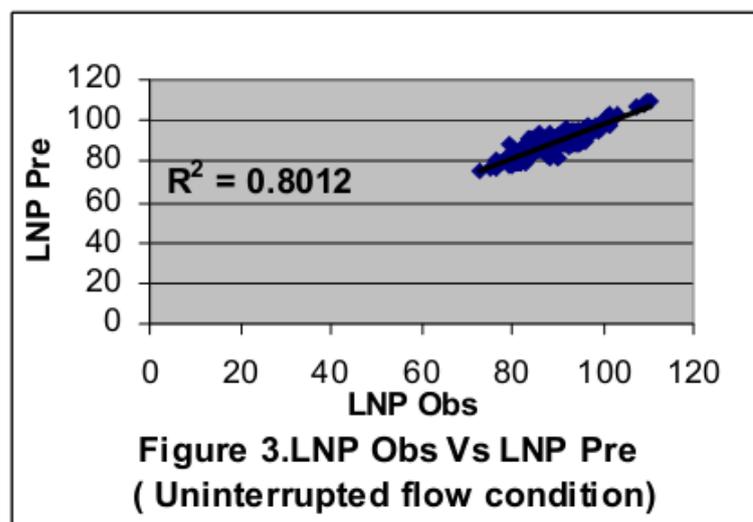


Figura 15. Comparación de valor medido y calculado para LNP en tráfico ininterrumpido en [83]

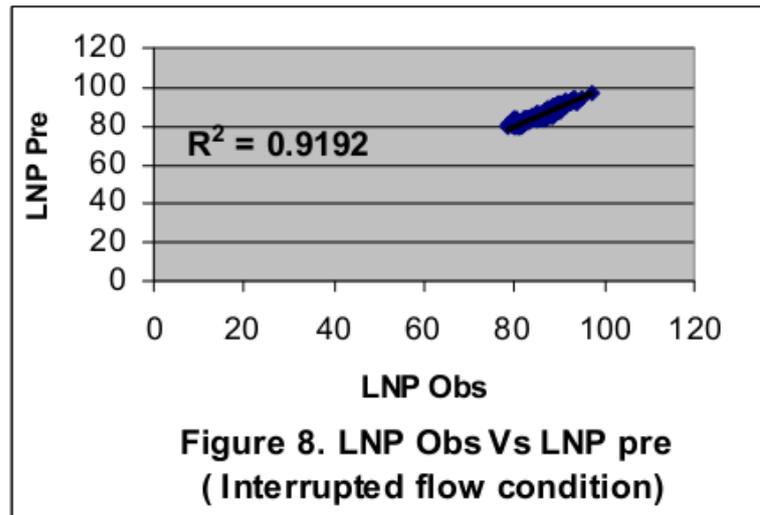


Figura 16. Comparación de valor medido y calculado para LNP en tráfico interrumpido en [83]

Los autores consideran el error cuadrático medio el mejor indicador de la precisión de la predicción, valor para el que obtienen los mejores resultados (Tabla 8 y Tabla 9).

En general, como se observa en las Tablas 8 - 9, la red neuronal utilizada en [83] devuelve unos resultados más precisos para el tráfico interrumpido.

Tabla 8. Indicadores de los resultados para tráfico ininterrumpido en [83]

Statistical parameters	L10	Leq	LNP	TNI	NC
Correlation Coefficient	0.69	0.81	0.90	0.88	0.84
Root mean square Error	2.55	2.20	2.57	5.75	1.91
Ave.% Error	2.31	2.18	2.21	4.56	10.0
Standard deviation	6.53	4.85	6.64	33.2	3.67
T-test	1.51	-0.87	1.06	0.76	2.49
Root mean square Error	2.55	2.20	2.57	5.75	1.91

Tabla 9. Indicadores de los resultados para tráfico interrumpido en [83]

Statistical parameters	L10	Leq	LNP	TNI	NC
Correlation Coefficient	0.89	0.94	0.96	0.95	0.94
Root mean square Error	1.01	0.74	1.09	2.46	0.74
Ave.% Error	0.93	0.81	0.91	2.15	4.47
Standard deviation	1.02	0.55	1.20	6.11	0.56
T-test	2.23	5.35	1.82	0.37	0.19
Root mean square Error	0.69	0.81	0.90	0.88	0.84

En [84] la predicción hecha por una red neuronal es comparada con los resultados del modelo predictivo inglés, CoRTN [17]. Se analizan datos del ruido tomados en una carretera de Teherán (Irán). Concretamente, el valor que se utiliza es $L_{Aeq,1h}$ (nivel de ruido en una hora de duración).

La red neuronal de tipo perceptrón multicapa (Figura 17) creada para la predicción del nivel sonoro tiene cinco entradas (flujo de tráfico, porcentaje de vehículos pesados, velocidad de los vehículos, pendiente de la calzada y ángulo de visión en la toma), cuatro neuronas en la capa oculta y una salida ($L_{Aeq,1h}$). Se dispone de 50 datos, separados para entrenamiento, validación y test en 30, 10 y 10 respectivamente, siendo el comportamiento de la red neuronal para los mismos el que puede verse en la Tabla 10.

Tabla 10. Resultados de la red neuronal de [84]

Subsets	$L_{Aeq,1hr}$ ranges dB(A)		Absolute prediction errors dB(A)			
	Model predictions	Field measurements	Mean	Standard deviation	Minimum	Maximum
Training	73.9–81.1	82.6–81.2	0.9	0.6	0.0	2.5
Testing	74.4–81.1	73–82.4	0.6	0.6	0.0	1.4
Holdout	74.1–80.3	72.1–79.1	0.9	0.8	0.0	2.4

Tabla 11. Test estadísticos para comparar la red neuronal y CoRTN en [84]

Asymp. Sig ^a (two-tailed)	Monte Carlo Sig ^b (two-tailed)	Wilcoxon Z ^c	Z _{0.975} ^d
0.401	0.460	-0.840	1.96

^a Asymptotic 2-tailed statistical significant.

^b Monte Carlo 2-tailed statistical significant, based on 40,000 sampled tables.

^c Z-score (Measure of the test).

^d Critical number (two-tailed).

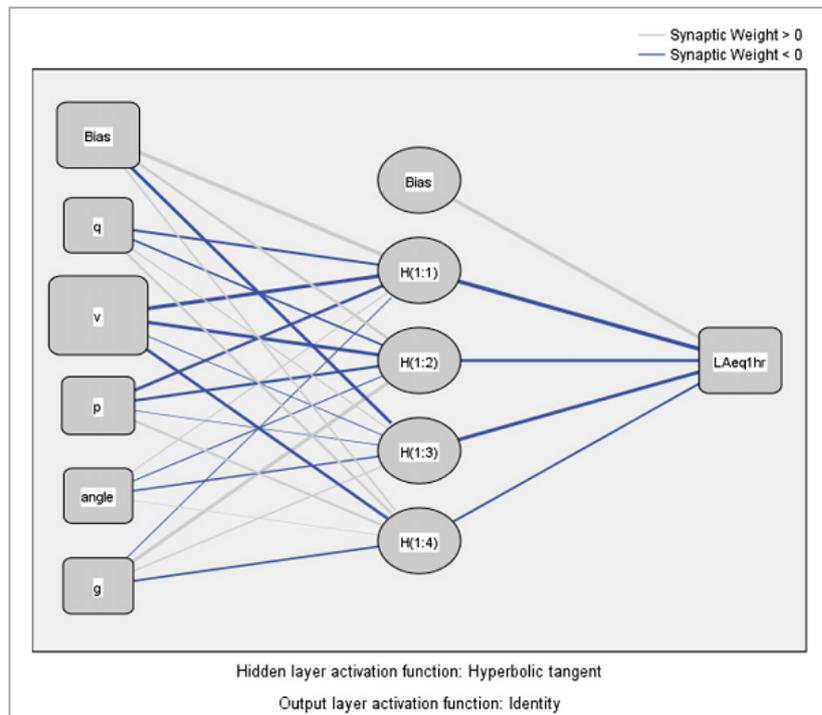


Figura 17. Red neuronal utilizada en [84]

En la comparación de la red neuronal con el modelo CoRTNse concluye finalmente que no hay diferencia significativa estadísticamente con un 95% de nivel de confianza entre los resultados de la predicción de ambos, según diferentes test estadísticos (ver Tabla 11). Los autores concluyen que la red neuronal no aporta una mejora a la predicción que ya hace CoRTN para estos datos.

En [85] se presenta un modelo de red neuronal de tipo perceptrón multicapa. Tras recolectar un total de 768 datos de ruido de la ciudad india de Lucknow en diez localizaciones de diferente naturaleza (zona residencial, comercial, poco ruidosa y de nivel alto de tráfico), la red recibe un total de 24 variables de entrada. Estas 24 variables son realmente 12 diferentes, ya que se hace distinción entre ambos sentidos del tráfico y todas están relacionadas únicamente con el tráfico (velocidad y número de vehículos). Las medidas son tomadas desde la misma distancia a la fuente de ruido (10 metros).

Tras la comprobación de diferentes modelos de red neuronal, la estructura de red neuronal que finalmente es elegida contiene una capa oculta con 10 neuronas y cuya variable de salida es L_{eq} . La red es entrenada durante 2000 épocas y el algoritmo de aprendizaje usado es Levenberg-Marquardt.

Tabla 12. Variables de entrada de la red neuronal de [85]

S.N.	Variables for near side of road (Ingoing flow)	S.No.	Variables for far side of road (outgoing flow)
1	Number of vehicles (Car/Jeep/Van)	13	Number of vehicles (Car/Jeep/Van)
2	Number of vehicles (Scooter/Motorcycle)	14	Number of vehicles (Scooter/Motorcycle)
3	Number of vehicles (LCV/Minibus)	15	Number of vehicles (LCV/Minibus)
4	Number of vehicles (Bus)	16	Number of vehicles (Bus)
5	Number of vehicles (Truck)	17	Number of vehicles (Truck)
6	Number of vehicles (3-Wheeler)	18	Number of vehicles (3-Wheeler)
7	Average speed of the vehicles (Car/Jeep/Van)	19	Average speed of the vehicles (Car/Jeep/Van)
8	Average speed of the vehicles (Two wheeler)	20	Average speed of the vehicles (Two wheeler)
9	Average speed of the vehicles (Mini bus/LCV)	21	Average speed of the vehicles (Mini bus/LCV)
10	Average speed of the vehicles (Bus)	22	Average speed of the vehicles (Bus)
11	Average speed of the vehicles (Truck)	23	Average speed of the vehicles (Truck)
12	Average speed of the vehicles (3-Wheeler)	24	Average speed of the vehicles (3-Wheeler)

Sus resultados son muy precisos (Figura 18), obteniendo un coeficiente de correlación de 0,94 para la predicción de L_{eq} . Tras realizar un t-test, se obtiene que los niveles medidos y los predichos no difieren significativamente al 95%.

En [86], también una red neuronal de una capa oculta es desarrollada para predecir el nivel de ruido de la autopista más concurrida de la zona de Agra (Agra-Firozabad, en India). En este caso, la red sólo recibe 3 entradas, también directamente relacionadas con el tráfico rodado: volumen de tráfico, velocidad del tráfico y número de vehículos pesados. Los niveles de ruido fueron tomados en 4 momentos diferentes del día en 88 puntos diferentes de la mencionada autopista,

llevando a cabo un total de 95 tomas de datos. Se relata el proceso de optimización de la estructura de la red, usando el 60% de los datos para entrenamiento, el 20% para validación cruzada, y el resto para test. La red neuronal por la que finalmente optan los autores es una red con una sola capa oculta de 10 neuronas que utiliza los algoritmos backpropagation (Levenberg-Marquardt). Se presentan varias gráficas de la evolución del error cuadrático medio en el entrenamiento (Figura 19) y en la validación según las épocas, pero no se citan los resultados de la red.

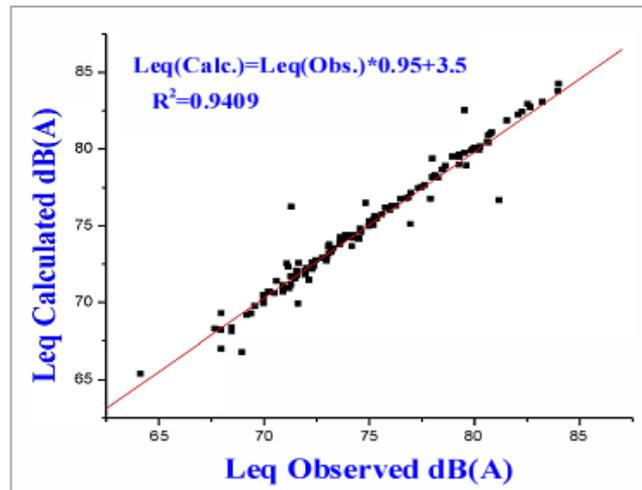


Figura 18. Resultados de predicción de Leq en [85]

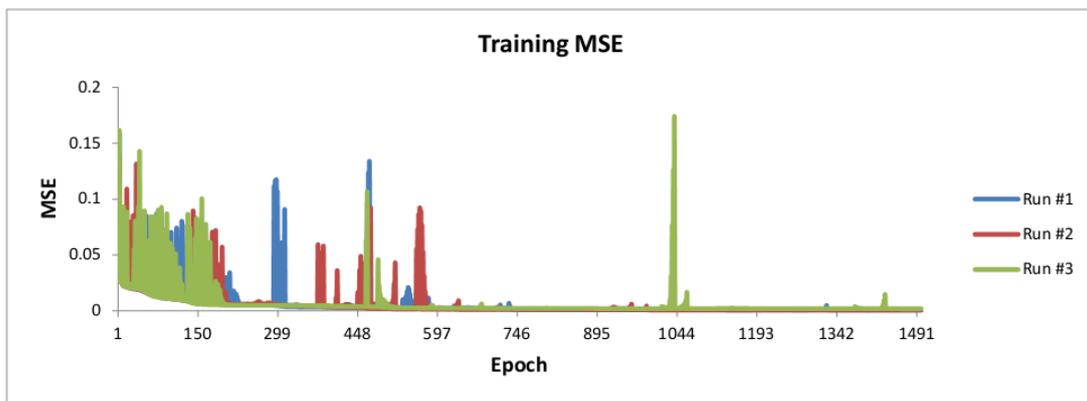


Figura 19. Evolución del error cuadrático medio en el entrenamiento de la red neuronal [86]

En [87] se utiliza una red neuronal de regresión general (GRNN, [141]). Los datos se toman en 14 localizaciones de la ciudad de Villa S. Giovanni (Italia), distinguiendo entre 5 tipos de vehículos (motocicletas, vehículos a motor, vehículos ligeros y pesados de transporte y buses) y tomando medidas de los parámetros L_{Aeq} , L_{10} , L_{50} y L_{90} . Finalmente se obtienen un total de 154 datos.

En un primer paso, se eligen 3 modelos matemáticos clásicos para determinar la bondad de los mismos y poder compararlos con los resultados que la red neuronal

obtenga: Burgess [138], García y Bernal [142] y una versión revisada del modelo CoRTN[143]. El coeficiente de correlación de la predicción hecha por los modelos clásicos seleccionados está entre 0,76 y 0,78 (Figura 20).

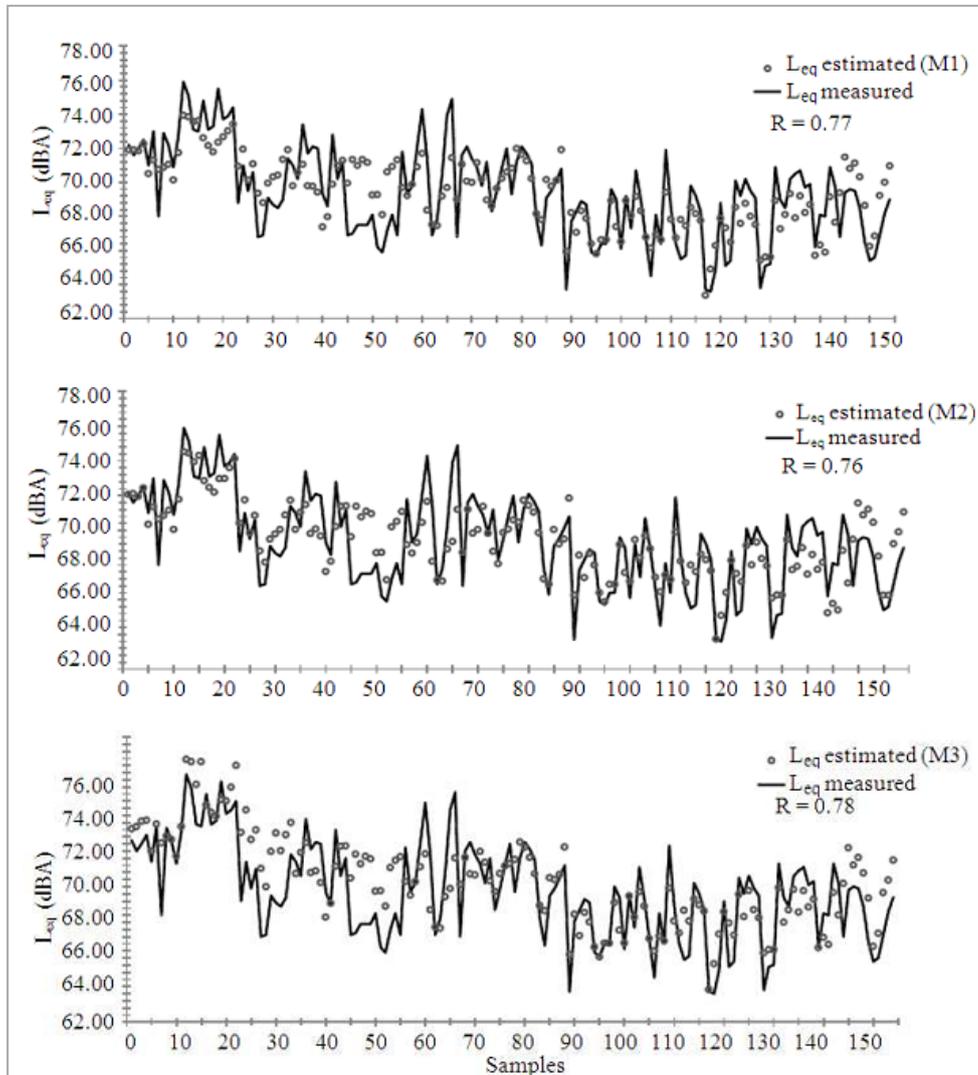


Figura 20. Comparación del L_{Aeq} medido y estimado por los diferentes modelos clásicos en [87].

Una vez comprobada la estimación de los modelos clásicos, el trabajo pasa a centrarse en la red neuronal de regresión general [141]. Ésta recibe 7 entradas: velocidad media de los vehículos, anchura de la calzada y flujo de los 5 tipos de vehículos comentados, siendo la salida L_{Aeq} . La red es entrenada usando conjunto de entrenamiento, de validación y de test. Los resultados obtenidos por la red son, según los autores, bastante satisfactorios (Tabla 13), aunque finalmente exponen que el disponer de más variables de entrada para la red (como las relativas al pavimento) haría más completo el modelo.

Tabla 13. Resultados de la ejecución de la red neuronal de [87]

Comparison between models and neural network		RMSE	Mean error	Correlation	Standard Deviation
GRNN Neural Network	training sets	0.643	0.317	0.996	0.44
	Verification set	1.254	1.030	0.904	1.249
	Test sets	1.361	1.121	0.874	1.478
	All data	0.941	0.514	0.938	0.928

En[88] también se utiliza una red neuronal de regresión general. Se comparan tres tipos de modelos predictivos de ruido urbano: una red neuronal de tipo PMC, una red neuronal de regresión general y un modelo estadístico de regresión lineal. Los resultados, en los que se detalla la media absoluta del error en entrenamiento y test (MAE_T y MAE_P) y el coeficiente de regresión lineal en ambos casos, se encuentran en la Tabla 14. Se usaron cinco variables como datos de entrada de los modelos: velocidad de los vehículos, número de vehículos por hora, porcentaje de vehículos pesados, anchura de la calzada y altura media de los edificios de la calzada. La salida es única: L_{Aeq} . Los 620 datos utilizados se tomaron en cuatro lugares cercanos a carreteras concurridas en la ciudad de Kuwait.

Tabla 14. Comparativa de los tres modelos predictivos usados en [88]

Neural Network Architecture	MAE_T	R^2_T	MAE_P	R^2_P	Smoothing factor
GRNN with Genetic adaptive	1.0	0.96	3.0	0.93	0.099
GRNN with Iterative adaptive	3.4	0.77	9.8	0.13	0.066
Optimum BPNN	5	0.83	4	0.87	
Regression	7.4	0.62	12.1	0.65	

En el caso de la red neuronal de tipo PMC, ésta fue configurada con un total de 30 neuronas en dos capas ocultas paralelas y entrenada durante 10.000 épocas.

Tabla 15. Resultados de la red neuronal de propagación hacia atrás en [88]

District/ Roadway	Noise Level (dBA)					
	Peak Hour			Off-Peak Hour		
	Measured	Predicted	Difference	Measured	Predicted	Difference
Gulf Rd.	81.4	80.5	-0.9	74.0	78.5	+4.5
S. Mobarak	75.3	77.5	+2.2	76.1	77.5	+1.4
Ghazali	80.2	82.4	+2.2	79.5	77.3	-2.2
Minawer	72.0	74.4	+2.4	72.1	74.1	+2.0

La red neuronal de regresión general está compuesta de una sola capa oculta y una sola neurona en la misma. Los autores prueban dos tipos de modelos adaptativos en este caso, genético e iterativo, comportándose el tipo genético mucho mejor, como se observa en la Tabla 14. En la Tabla 15 se muestran los resultados para

la red neuronal de propagación hacia atrás y en la Tabla 16 se detallan los resultados para el tipo iterativo genético.

Tabla 16. Resultados de la red neuronal de regresión general en [88]

District/ Roadway	Noise Level (dBA)					
	Peak Hour			Off-Peak Hour		
	Measured	Predicted	Difference	Measured	Predicted	Difference
Gulf Rd.	81.4	81.5	+0.1	74.0	74.5	+0.5
S. Mobarak	75.3	75.5	+0.2	76.1	76.5	+0.4
Ghazali	80.2	80.3	+0.1	79.5	79.7	+0.2
Minawer	72.0	72.2	+0.2	72.1	72.1	0

El modelo de regresión lineal ofrece resultados peores a los anteriores, como se observa en la Tabla 17.

Tabla 17. Resultados de la predicción del modelo de regresión lineal de [88]

District/ Roadway	Noise Level (dBA)					
	Peak Hour			Off-Peak Hour		
	Measured	Predicted	Difference	Measured	Predicted	Difference
Gulf Rd.	81.4	84.5	+3.1	74.0	81.5	+7.5
S. Mobarak	75.3	79.5	+4.2	76.1	87.5	+11.4
Ghazali	80.2	85.5	+5.3	79.5	86.3	+6.8
Minawer	72.0	79.0	+7.0	72.1	79.1	+7.0

Acorde a los resultados que se presentan en el artículo, los autores concluyen que el mejor método predictivo es el de red neuronal de regresión general.

B. PREDICCIÓN DEL GRADO DE MOLESTIA

Además de los efectos ya comentados, el ruido también puede causar deficiencias en los trabajadores que lo sufren, y para la predicción de esta deficiencia, en [91] usan un modelo neuro difuso. Una vez comprobado que los parámetros que más afectan a los trabajadores son la contaminación acústica, tipo de tarea y el tiempo de exposición a la misma, los autores proponen un sistema neuro-difuso usando la toolbox de lógica difusa de MATLAB y ANFIS (adaptive-network-based fuzzy inference system) [144]. Cuando ya está construido el modelo, se comprueba cómo el nivel de ruido afecta mayormente a la eficiencia, al realizar tareas más complejas, mientras que tareas más sencillas son completadas con éxito independientemente del nivel de ruido ambiental.

En la Figura 21 puede verse la estructura del modelo. En la primera capa se introducen los valores de entrada, en este caso son: el nivel de ruido, el tipo de tarea y el tiempo de exposición (duración de la tarea).

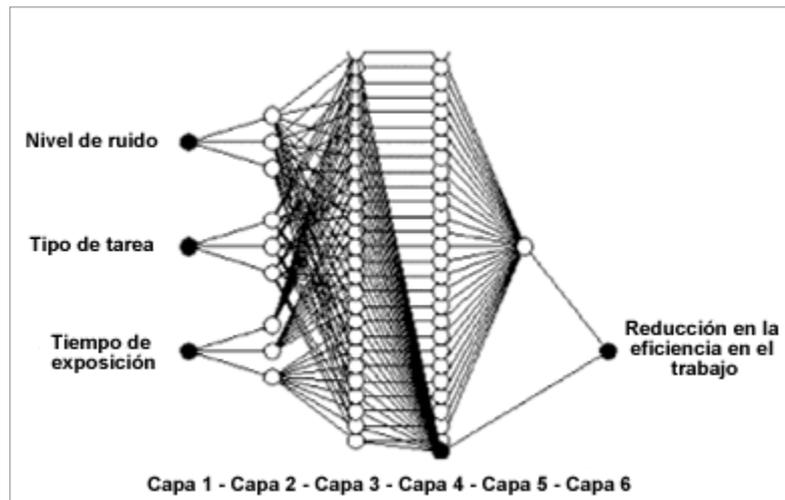


Figura 21. Estructura del modelo neuro difuso de [91]

La segunda etapa, o capa de “defuzzyficación”, describe la función de pertenencia de cada conjunto difuso de entrada. En la tercera capa o capa de inferencia, cada nodo es fijo y representa la parte “SI” (IF) o el antecedente de cada regla difusa. En la siguiente capa, llamada capa de normalización, se calcula el coeficiente de disparo de cada regla difusa. La quinta capa, o capa de salida, representa la parte “ENTONCES” o consecuente de la regla difusa en cuestión. Por último, en la sexta capa, se realiza la desfuzzyficación.

El modelo definido permite llegar a conclusiones como la siguiente: la reducción en eficiencia laboral es de un 37%, 70% y 85% para tareas simples, moderadas y complejas, respectivamente, cuando el nivel de ruido es de 105 dB y superior. Ésta y otras conclusiones han sido contrastadas con las dadas por los “Límites de exposición seguros para trabajadores del sector industrial (SLR)”, que, por ejemplo, no permite más de 85dB para un trabajo de 8 horas de duración. El modelo obtiene resultados concordantes, ya que, para 85 dB, no hay prácticamente reducción en el nivel de eficiencia laboral.

Por otro lado, en [93], se usa una red neuronal para predecir la molestia y otras opiniones subjetivas sobre el nivel sonoro recogidas en entrevistas durante las tomas de las componentes físicas del sonido. Las tomas de datos son llevadas a cabo en 6 países diferentes (China, Reino Unido, Grecia, Italia, Alemania y Suiza).

Los datos de entrada para las distintas redes neuronales valoradas son fundamentalmente los físicos (Phy), los psicológicos (Psy), los sociales (S) y los relativos al comportamiento (B), que se muestran en la Tabla 18. El objetivo es

predecir la evaluación subjetiva relativa al confort acústico y al nivel sonoro que manifiestan en las encuestas los ciudadanos o los turistas. También posteriormente se diseña un mapa de calidad del paisaje sonoro gracias a los datos recogidos.

Tabla 18. Factores medidos en la toma de datos de [93]

Attributes	Elements	Ref. No.	Attribute factors	Measures of the attributes
Explicit	Physical	Phy1	Season	1—winter; 2—autumn; 3—spring; 4—summer
		Phy2	Time of day	1—night > 21:00pm–8:59am; 2—evening: 18.00–20.59pm; 3—morning: 9.00am–11.59am; 4—afternoon: 15.00–17.59pm; 5—midday: 12.00–14.59pm
		Phy3	Air temperature	Measurement of air temperature: °C
		Phy4	Wind speed	Measurement of wind speed: m s ⁻¹
		Phy5	Relative humidity	Measurement of relative humidity: %
		Phy6	Horizontal luminance	Measurement of horizontal luminance: Klux (Europe); lux (China)
		Phy7	Sun shade	0—interviewee not standing in the sun; 1—interviewee standing in the sun
		Phy8	Sound pressure level	Measurement of sound pressure level: dB(A)
	Behavioral	B1	Whether wearing earphones	0—not wearing earphone; 1—wearing earphone
		B2	Whether reading or writing	0—neither reading nor writing; 1—either reading or writing
		B3	Whether watching somewhere	0—not watching anywhere; 1—watching somewhere
		B4	Movement status	1—sitting; 2—standing; 3—playing with kids; 4—sporting
		B5	Frequency of coming to the site	Scales 1–5; 1—first time; 5—every day
		B6	Reason for coming to the site	1—the equipment/services of the site; 2—children playing and social meetings; 3—business/meeting/break; 4—attending social events; 5—passing by
B7		Grouping: whether accompanied	0—none; 1—with 1 person; 2—with more than 1 person	
Implicit	Social/demographic	S1	Age	1 < 12; 2 = 12–17; 3 = 18–24; 4 = 25–34; 5 = 35–44; 6 = 45–54; 7 = 55–64; 8 > 65
		S2	Gender	1—male; 2—female
		S3	Occupation	1—students; 2—working people; 3—others (e.g., unemployed and pensioners)
		S4	Education	1—primary; 2—secondary; 3—higher education
		S5	Residential status	0—nonlocal; 1—local
		S6	Sound level experience at home	Scales –2 to 2, with –2 as very quiet and 2 as very noisy
	Psychological	Psy1	Site preference	0—not like the site for certain reasons; 1—like the site
		Psy2	View assessment	Scales –1 to 1, with –1 as negative and 1 as positive
		Psy3	Heat evaluation	Scales –2 to 2, with –2 as very cold and 2 as very hot
		Psy4	Wind evaluation	Scales –2 to 2, with –2 as stale and 2 as too much wind
		Psy5	Humidity evaluation	Scales –2 to 2, with –2 as very damp and 2 as very dry
		Psy6	Brightness evaluation	Scales –2 to 2, with –2 as very dark and 2 as very bright
		Psy7	Overall physical evaluation	0—not comfortable; 1—comfortable
		Psy8	Sound level evaluation	Scales –2 to 2, with –2 as very quiet and 2 as very noisy

Los autores concluyen que resulta complicado obtener un modelo general útil para todas las localizaciones propuestas, mientras que para un solo lugar se obtiene mejores resultados, pero su aplicación estaría más limitada. Según se resume finalmente en [93], los resultados en los conjuntos de test no son demasiado buenos (aunque mejores si se predicen en sitios de forma individual), ya que es difícil representar con un modelo computacional un comportamiento subjetivo (Tabla 19).

Trabajos similares son desarrollados por los autores de [93] en [92][89] y [90].

Tabla 19. Comportamiento de la red neuronal de [93] en la predicción del confort acústico

	Site	Network architecture				Results				
		Input variables	Output variable	Hide layer	Hide node	Test pattern	Coefficient		RMS	
							Train	Test	Train	Test
Model 1	13-Barkers Pool			2	6	100	0.60	0.56	0.119	0.121
	14-Peace Gardens	Phy3, 4, 5, and 8								
	15-Chuang Chun Yuan	B4 and 7								
	16-Xi Dan	S1, 4, 5, and 6								
	17-Century Square	Psy2, 4, 5, 6, 7, and 8								
	18-Nanjing Road Square		Subjective evaluation of acoustic comfort							
	19-Xu Jia Hui Park									
Model 2	13-Barkers Pool			1	4	24	0.80	0.63	0.089	0.117
	14-Peace Gardens	Phy5 and 8		1	3	25	0.76	0.58	0.010	0.125
	15-Chuang Chun Yuan	B7		1	4	30	0.71	0.59	0.082	0.124
	16-Xi Dan	S6		2	5	27	0.64	0.37	0.107	0.109
	Shanghai sites combined	Psy2, 3, 6, 7, and 8		1	2	15	0.71	0.35	0.085	0.136
	All seven sites combined			2	7	110	0.59	0.59	0.118	0.124

C. CLASIFICACIÓN

Las RNA son por lo general buenas en el proceso de clasificar, por lo que su uso parece adecuado para clasificar tipos de ruido.

En[94] se implementa una red neuronal de tipo PMC con una capa de entrada de 5500 nodos (5500 coeficientes de la transformada de Fourier para datos de diferentes sonidos), una capa oculta con 80 neuronas y dos salidas, con objeto de distinguir entre 5 sonidos diferentes: aire acondicionado en un taller, maquinaria de un taller, flujo de agua en un taller, voz masculina, paso de una motocicleta. Los experimentos diferentes fueron tres, cada uno de ellos intentando clasificar diferentes combinaciones de los mencionados sonidos. En un mismo experimento sólo se intentó clasificar un máximo de 4 de estos sonidos (las celdas sin datos de la Tabla 20 indican que no se clasificó ese sonido).

Los autores de [94] concluyen que cuando no está presente el sonido de la voz, la predicción es mucho mejor que en el resto de los casos. Los sonidos más difíciles de identificar fueron la voz humana y el paso de motocicletas.

Otro clasificador de ruido con el uso de redes neuronales es presentado en [95], esta vez diferenciando entre sonidos de aviones. La red neuronal tiene una arquitectura PMC con tres neuronas en la capa intermedia. Se obtuvieron un total de 600 registros, con mediciones hechas en aeropuertos noruegos; a la vez que se hicieron grabaciones de sonidos de trenes, autobuses, coches, etc. para alimentar a la red con ejemplos que describan al grupo no deseado en los cálculos.

Tabla 20. Porcentaje de sonidos mal clasificados para los 3 experimentos llevados a cabo en [94]

Exp	air-condition			machine			water			voice			motor bike		
	nd	w	%	nd	w	%	nd	w	%	n	w	%	n	w	%
1	3	0	7.5	18	0	45	6	0	15	28	0	70			
2	0	0	0	0	0	0	5	0	12.5						
3	0	0	0	0	0	0	1	0	2.5				1	5	55

La red recibe, con un total de 26 parámetros de entrada, vectores consistentes en valores medios en tiempo y frecuencia en 1/3 de banda de octava. Una vez entrenada la red, y dando como entrada un conjunto de datos completamente nuevo, los resultados han sido medianamente buenos, obteniendo un reconocimiento del 76% de los sonidos de hélices, un 90% de los sonidos de aviones y un 66% de los sonidos de motores.

En [96] se describe un “Sistema Automático de Reconocimiento de Ruido” (ANR Automatic Noise Recognition System) basado en Redes Neuronales Artificiales junto con Modelos Ocultos de Markov, y capaz de distinguir, gracias a una base de datos de 1000 registros sonoros, entre dos sonidos: claxon o motor de motocicleta. El proceso sigue el esquema de la Figura 22.

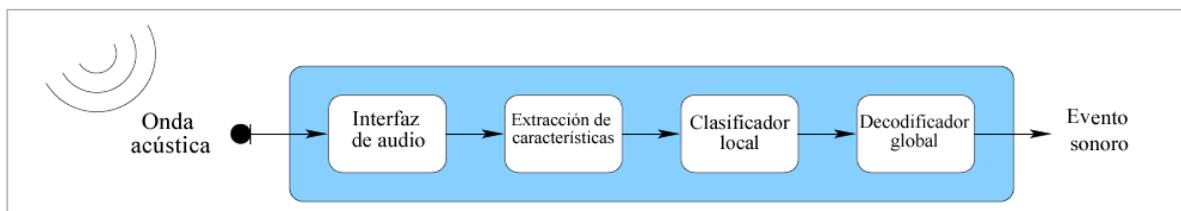


Figura 22. Esquema del Sistema Automático de Reconocimiento de Ruido [96]

En la primera fase, las ondas sonoras son captadas por un micrófono, y después, son convertidas en señal digital para que la información pueda ser procesada por un ordenador. Posteriormente se pasa a la segunda fase, de extracción de características, en la que, tras dividir la señal en marcos (de 60 ms), se procesa el espectro de cada uno de estos marcos. A partir de espectrogramas de potencia se

observan dos patrones diferentes, a los que se les aplica el algoritmo PLP [145] para obtener un mejor reconocimiento de los mismos.

Tras esta segunda fase, se pasa al clasificador local, el cual intenta extraer, de cada marco registrado, si pertenece al ruido de un claxon, de un motor de motocicleta o de ambos dados al mismo tiempo. Inicialmente se asigna, a cada vector de características, un estado. En este momento es cuando se hace uso de Redes Neuronales Artificiales, más exactamente del tipo perceptrón multicapa. La red tiene el objetivo de, dado un vector de características, estimar la probabilidad a posteriori $P(\text{estado}|\text{vector})$ para cada estado. La probabilidad más alta da el estado de un vector de características. Los coeficientes de la red se obtienen por entrenamiento supervisado.

Finalmente, el decodificador global tiene por objetivo buscar, entre las secuencias de vectores de características observados y las probabilidades de estado dadas, la secuencia de estados más adecuada. Para esta fase se usan Modelos Ocultos de Markov, donde los estados se suponen unidos formando una cadena de Markov. En el conjunto de prueba final se obtuvieron entre el 85% y el 95% de aciertos, lo cual valida los buenos resultados del sistema.

En este apartado también se incluye [97], ya que, aunque se basa en Modelos Ocultos de Markov, también incluyen redes neuronales. Se describe en la sección dedicada a Modelos Ocultos de Markov (2.3.C).

D. PREDICCIÓN DEL FLUJO DE TRÁFICO

Se incluyen también varias publicaciones sobre la predicción del flujo de tráfico urbano basada en redes neuronales artificiales. Estos artículos no tratan directamente el ruido como tal, no obstante, considerando que en la mayoría de los modelos de predicción de ruido el caudal de tráfico ha demostrado jugar un papel decisivo, es interesante incluir los más relevantes en este trabajo.

[98] aporta un modelo neuro-difuso para predecir el caudal de tráfico urbano. Desarrolla un enfoque para obtener relaciones entre los parámetros implicados en la caracterización del ruido ambiental producido por tráfico urbano y el número de vehículos. Se trata de una red neuronal difusa con la que se pretende aprovechar las cualidades de ambas técnicas (aprendizaje mediante ejemplos de las RNA y sencillez de lógica difusa) a la vez que evitar los inconvenientes de las mismas. El algoritmo de aprendizaje se deriva de una adecuada adaptación del algoritmo Backpropagation. La estructura del modelo neuro-difuso de [98] se observa en la Figura 23.

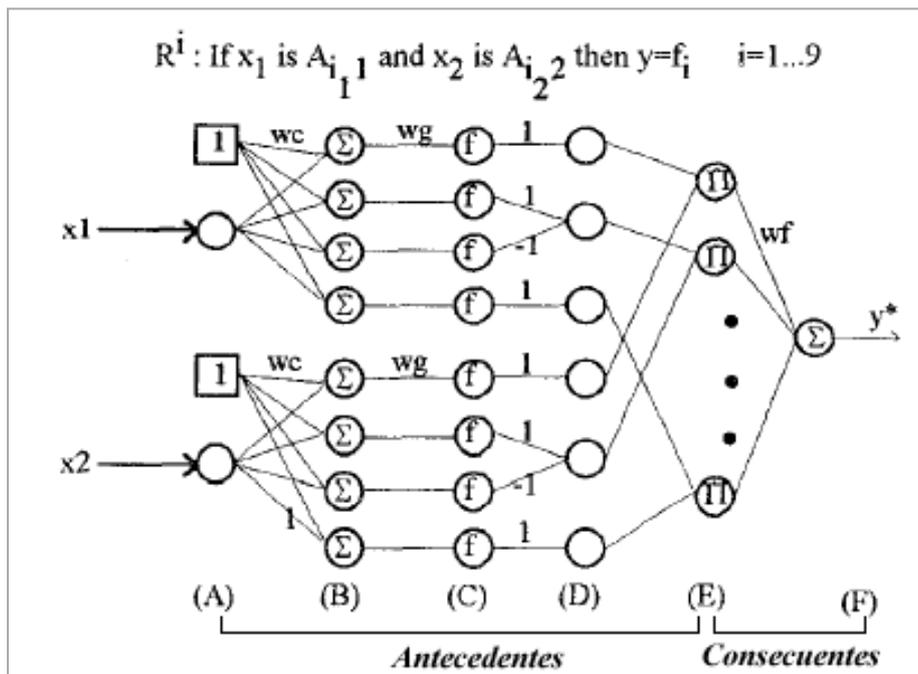


Figura 23. Estructura del modelo neuro-difuso de [98]

El caudal de vehículos se obtuvo mediante una red neuronal, en función de la anchura de la vía, la altura media de los edificios de la calle, y el nivel de presión sonora equivalente L_{eq} . Los datos para entrenar la red fueron obtenidos de la ciudad de Catania, en Sicilia. Los mejores resultados (no se encuentran explícitamente mencionados en [98]) fueron obtenidos para una red neuronal con 3 neuronas de entrada, 2 conjuntos difusos asociados a cada variable de entrada y 80 patrones de entrenamiento (de los 120 obtenidos).

Además del flujo de tráfico, en [99] se intenta predecir también la velocidad y la ocupación de la vía. El problema básico que se encuentra es la gran cantidad de entradas que puede tener el sistema, ya que se considera que la red será más efectiva cuantos más datos de periodos previos tenga disponibles. Además, se quiere trabajar en tiempo real, y la cantidad de datos a manejar hace complicado que sea así. Se debe tener en cuenta que estas dificultades no serían tales actualmente, ya que se dispone de máquinas más potentes que en el año 1997, cuando fue desarrollado este trabajo.

Finalmente, se opta por modelar el sistema con una red neuronal de tipo PMC, con 40 entradas y 10 neuronas en la capa intermedia. Los datos fueron tomados en autopistas en los alrededores de Utrecht, Holanda. Se hicieron predicciones a 5, 15 y 30 minutos, obteniendo mejor resultado a 5 minutos y cuando se predicen el flujo y la ocupación. Para analizarlos se utiliza como indicador el error cuadrático medio proporcional (RMSEP), logrando unas cifras en los siguientes rangos:

[0.18 0.28] para la predicción del flujo de vehículos

[0.15 0.27] para la predicción de ocupación

[0.31 0.42] para la predicción de la velocidad

Como conclusión final, se expone la imposibilidad de tener todas las entradas deseables en la red para la correcta predicción de los tres parámetros, por lo que se propone para trabajos futuros el uso de redes neuronales recurrentes. Esto, sumado a los malos resultados para la predicción de la velocidad (atribuido a condiciones de bajo flujo de vehículos), constituye el problema por el cual la red no ofrece los resultados esperados.

En [100] se propone un modelo de flujo de tráfico urbano basado en redes neuronales. Este modelo es parte de un sistema adaptativo en tiempo real de control de tráfico urbano (UTC) y se representa según la Figura 24. Este tipo de sistemas pueden estar enfocados a controlar el tráfico a nivel local (en un conjunto de cruces cercanos de forma separada) o a nivel global (teniendo en cuenta las interacciones entre los distintos cruces). Para intentar combinar las ventajas de ambos procesos, el modelado del flujo de tráfico se hace en dos pasos:

En primer lugar, el flujo del tráfico es modelado en un solo enlace (calle entrante en un cruce) señalizado por una red neuronal local. En general, se obtiene que el valor de la longitud de la cola en la calle entrante en un cruce depende de: su valor en el minuto anterior $-l_q(t+j)-$ y el número de entradas $-D(t+j)-$ y salidas $-A(t+j)-$ de la cola estimadas durante ese minuto. Este valor es calculado con una red neuronal de tipo perceptrón multicapa, cuatro neuronas en la capa oculta y con estos valores como entradas y la longitud de la cola como salida.

En segundo lugar, basándose en comunicaciones entre las redes neuronales locales, el flujo de tráfico es modelado sobre una gran red de cruces de calles. Las salidas de las diferentes redes neuronales de cada calle entrante son combinadas para obtener el flujo en un cruce.

Basado en datos simulados, el artículo concluye indicando el potencial que tienen las redes neuronales aplicadas al campo del modelado del flujo de tráfico urbano. Predicciones con un minuto de adelanto de las longitudes de las colas de los flujos de salida han sido obtenidas con una alta precisión, estando la media y la desviación estándar del error en los rangos [0.5 0.7] y [0.2 0.7] respectivamente. El artículo deja abierto el estudio dado el potencial de las redes neuronales para este caso, también debido a que se necesitan mejoras futuras.

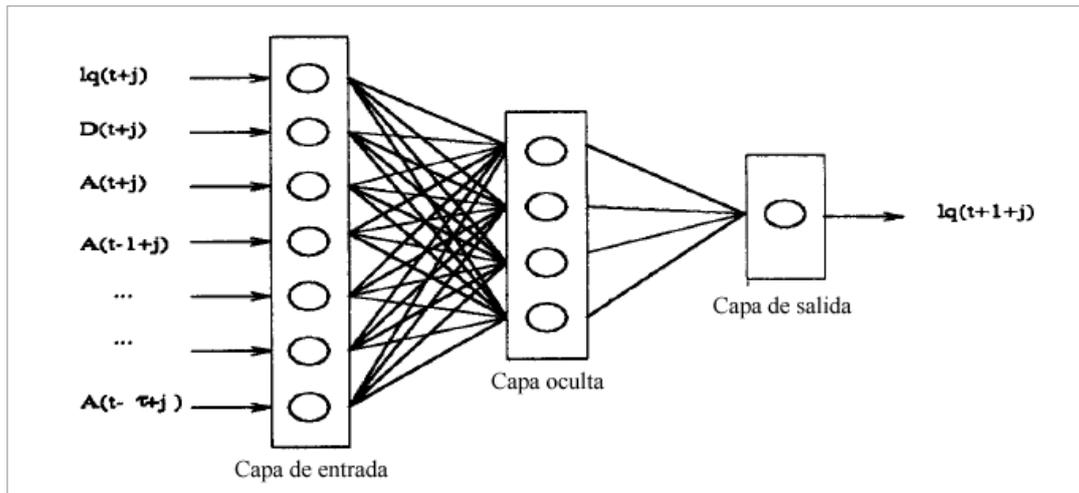


Figura 24. Estructura de la red de [100]

Por otro lado, en [101] se discute un modelo de red neuronal que fue desarrollado para la predicción a corto plazo de condiciones de tráfico en Queensland, Australia. La red utilizada es recurrente de tipo time-lag[146], aunque fueron probadas redes de tipos diferentes. Sirviéndose del error cuadrático medio para comparar, el mínimo obtenido para este tipo de red, con una media en los 3 conjuntos utilizados (de entrenamiento, de validación y de test), es de 0.0080.

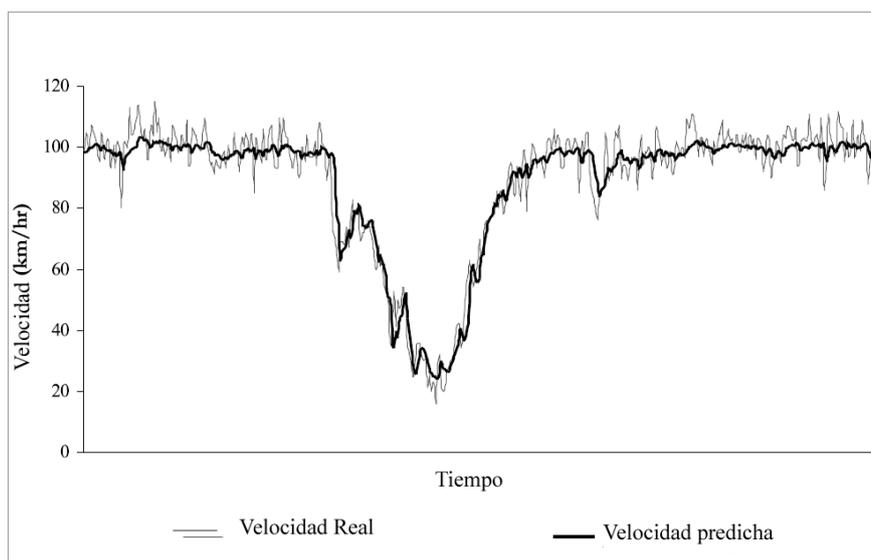


Figura 25. Velocidad predicha por la red con 20 seg de antelación [101]

La red fue entrenada para que fuera capaz de predecir resultados con hasta 15 minutos de antelación, aunque los resultados obtenidos indicaron que la red es capaz de predecir velocidad con hasta 5 minutos por delante de que se produzca el evento con un grado de precisión del 90- 94%. En cambio, los peores resultados se dieron cuando eran necesarios 10-15 minutos de antelación, obteniendo en este caso 88% y 84% de precisión respectivamente. Las Figuras 25 y 26 muestran las gráficas de la velocidad calculada y la real para tiempos de predicción de 20 y 60 segundos respectivamente.

Se parte de la idea de que sabiendo las condiciones actuales del tráfico y las demandas para usar la vía será posible predecir el estado futuro del tráfico en unos minutos. Para ello se construye una red neuronal de tipo PMC cuya estructura puede verse en la Figura 27.

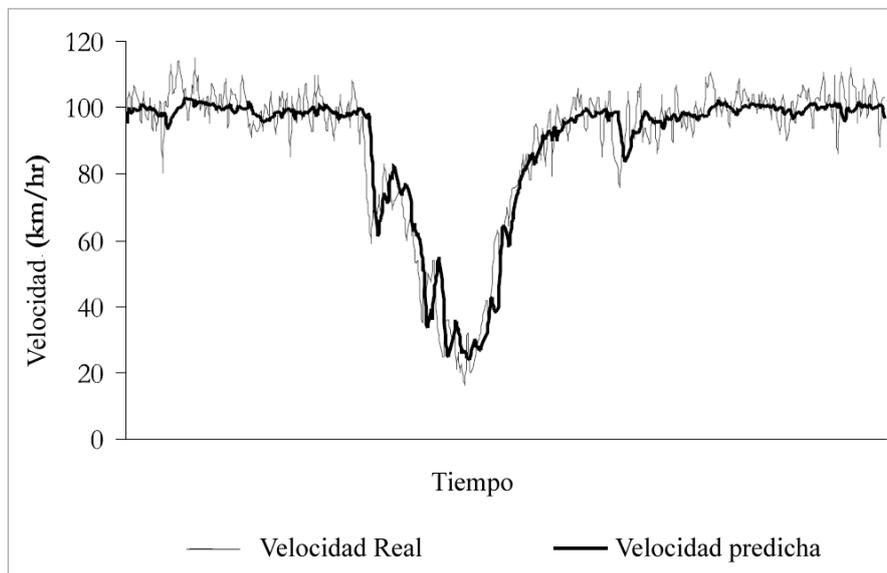


Figura 26. Velocidad predicha por la red con 60 seg de antelación [101]

Entre las conclusiones se considera relevante que, para la mejora de la predicción de la velocidad, haya siempre una buena estimación previa de la duración del trayecto entre dos puntos dados. Gracias a la estimación dinámica de la misma, los conductores pueden ser informados de retrasos o de la duración del trayecto hacia su destino. Los resultados fueron, según los autores, muy buenos, al igual que para la predicción de la velocidad. El máximo error cometido por la red fue para predicción de duración del trayecto con 15 minutos de antelación, para el cual se obtuvo un 7% de casos fallidos.

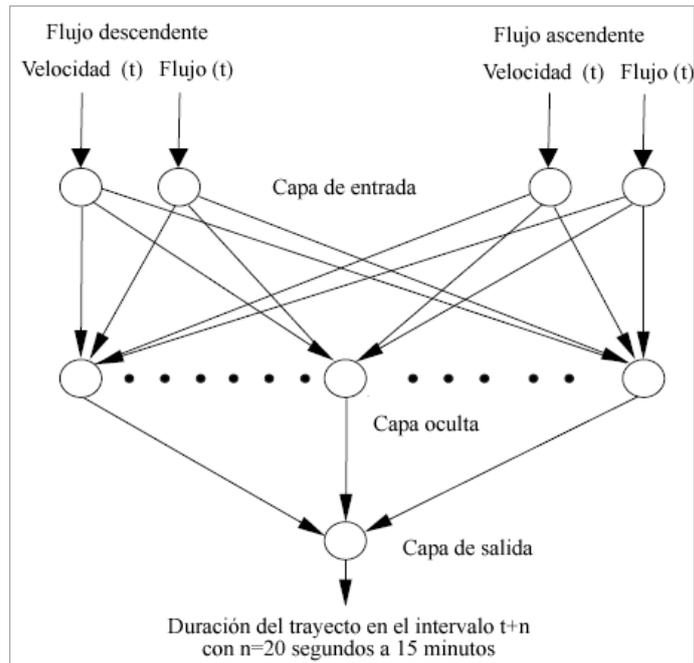


Figura 27. Red neuronal para predicción de la duración del trayecto de [101]

En [102] se desarrolla un modelo neuro-difuso para predecir el flujo de vehículos en una red de calles. El modelo consiste en dos módulos: una red de entrada (GN, Gate Network) y una red experta (EN, de Expert Network). La GN clasifica los datos de entrada en un número de clústeres usando un método difuso, y la EN se encarga de especificar la relación entrada-salida como en una red neuronal convencional. Es decir, mientras la GN agrupa patrones de tráfico de similares características en clústeres, la EN modela relaciones específicas de cada cluster. La estructura del modelo se detalla en la Figura 28.

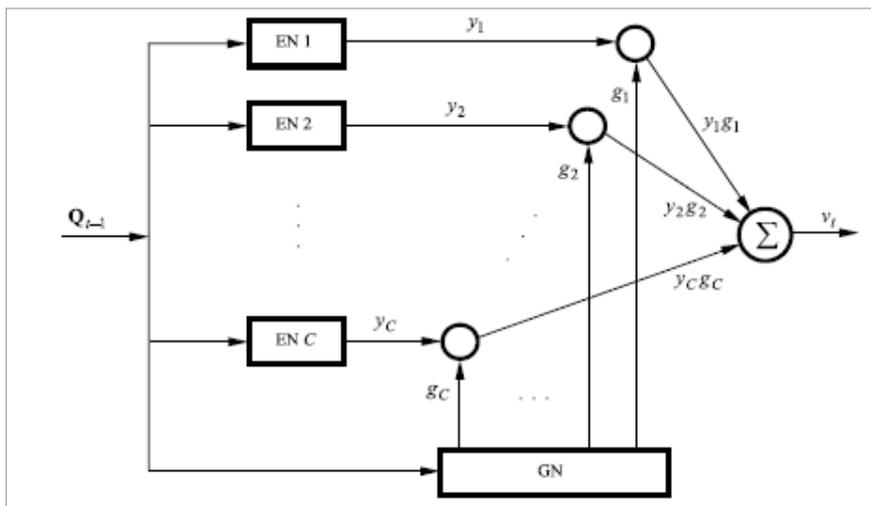


Figura 28. Estructura del modelo neuro-difuso de [102]

Los resultados han sido comparados con los dados por un modelo exclusivamente basado en redes neuronales [102], aventajando a éste el modelo presentado. Además de tener unos requerimientos de tiempo más bajos, presenta una mejora de entre el 23% y el 30% en la estimación de vehículos por hora en cinco puntos diferentes de la ciudad de Hong Kong.

Por otro lado, en [103], se comparan cuatro modelos diferentes de arquitecturas de redes neuronales: perceptrón multicapa, redes modulares, redes neuronales de Análisis de Componentes Principales y CANFIS (Sistema de inferencia neuro difuso coactivo). En cada modelo se ha investigado el efecto de introducir un componente de memoria a largo plazo además de los componentes a corto plazo en un rango de 5 a 20 minutos. Para asegurar un comportamiento óptimo, los modelos fueron probados bajo un amplio espectro de condiciones de tráfico, tanto espaciales como temporales.

Los datos fueron recogidos en una autopista de Florida, siendo la entrada a la red dividida en dos componentes: STM y LTM (memoria a corto y largo plazo, respectivamente). El componente STM se representó por la información espacio temporal de los 10 minutos anteriores y expresada en términos de medias de velocidad cada 5 minutos. El componente LTM se representó según el sello de tiempo de cada componente STM.

En [103] se concluye finalmente que ninguna topología es mejor que otra, aunque para el 33% de los casos, CANFIS dio los mejores resultados. El estudio muestra también que las configuraciones óptimas fueron más dependientes de LTM cuando el horizonte de predicción (tiempo de predicción) se incrementa (Figura 29).

[104] presenta una red neuronal para la predicción de flujo de tráfico que es seleccionada por una estrategia de optimización multicapa basada en algoritmos genéticos. Además, el algoritmo genético en cuestión es también utilizado para la representación del flujo de tráfico con características temporales y espaciales. Una vez seleccionada la red neuronal, se evalúa el comportamiento de la misma entrenándola con datos de tráfico, tanto de una como de múltiples variables, de una calle principal de Atenas, Grecia.

Los autores concluyen que la red neuronal que mejor captura en intervalos de tiempo el flujo de tráfico es el perceptrón multicapa (PMC) con optimización genética. Los resultados son, además, especialmente prometedores cuando los datos son de arterias de tráfico señalizadas muy congestionadas.

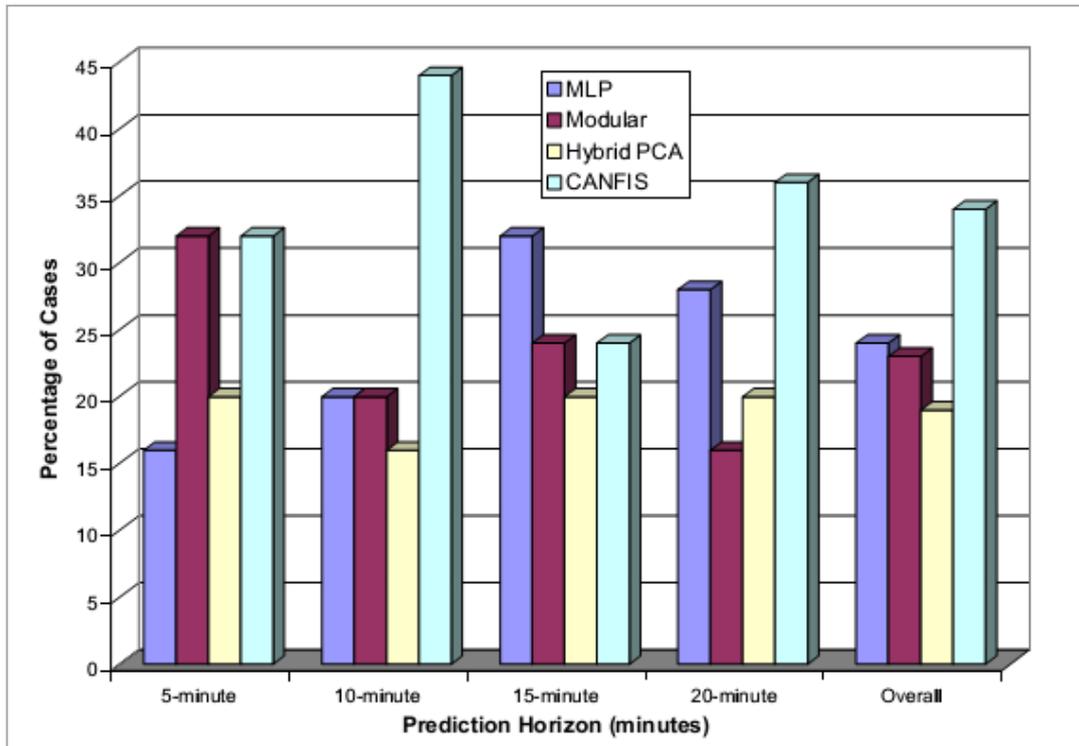


Figura 29. Frecuencia de redes óptimas según el horizonte de predicción en minutos basado en el error relativo medio absoluto [103].

Tabla 21. Resumen de resultados de [104]

Data	Model	Performance evaluation	Prediction					
			$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	
Univariate	Gradient descent	MAE	6	11	12	13	14	
		MRPE	8.54	16.76	18.74	20.36	21.71	
		r	0.95	0.84	0.80	0.77	0.75	
	Adaptive step size	MAE	6	11	12	13	14	
		MRPE	8.49	16.88	18.67	20.49	21.83	
		r	0.95	0.84	0.80	0.77	0.75	
	Scaled conjugate gradient descent	MAE	7	11	12	13	14	
		MRPE	8.71	17.01	18.86	20.66	22.02	
		r	0.95	0.84	0.80	0.77	0.74	
	ARIMA	MRPE	16	n/a	n/a	n/a	n/a	
	Multivariate	Gradient descent	MAE	7	11	11	12	13
			MRPE	8.39	16.66	18.31	19.99	20.39
r			0.95	0.85	0.82	0.79	0.78	
State-space		MRPE	13	n/a	n/a	n/a	n/a	

En [105] se presenta un modelo de red neuronal recurrente de ondícula (o wavelet) para la predicción del flujo de tráfico. Inicialmente se introduce al concepto de ondícula para añadir características extra como parámetros de traducción adaptada deseables en la predicción del flujo de tráfico. Posteriormente, se usa una función estadística de autocorrelación para la selección de las dimensiones óptimas

de entradas de la serie temporal del flujo de tráfico. El modelo (Figura 30) incorpora datos sobre el día de la semana y la hora, con lo que los autores indican que pueden predecir de esta manera el tráfico a corto y largo plazo.

Los autores concluyen mencionando que incluso con datos de entrenamiento limitados (672 datos), la precisión del modelo es buena considerando la naturaleza complicada del problema de la predicción.

En [106] se combina la predicción de las redes neuronales con un algoritmo heurístico adaptativo de asignación de “créditos” basado en la teoría de la probabilidad condicional y en el teorema de Bayes. Se diseñan y combinan en una sola red neuronal Bayesiana dos métodos de predicción: una red neuronal PMC y otra de base radial. Cada modelo predictivo tiene un valor calculado o “crédito” dado por el algoritmo de asignación, valor que depende de su comportamiento durante intervalos de predicción previos.

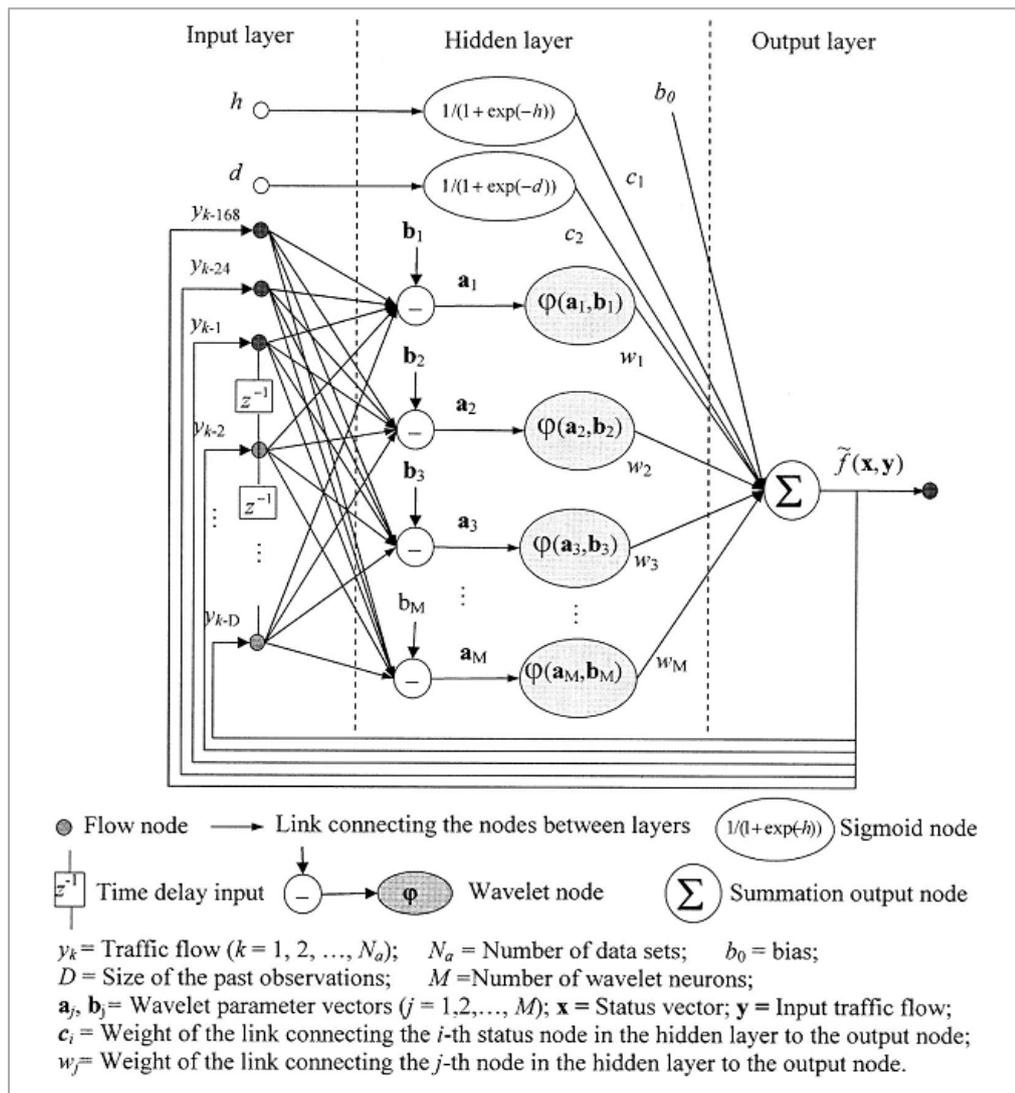


Figura 30. Modelo de red neuronal de ondícula de [105]

Tabla 22. Resultados de los distintos modelos predictivos de [106]

Time	BP neural network			RBF neural network			BCNN		
	MAPE (%)	VAPE (%)	Probability (error $\leq \pm 10\%$) (%)	MAPE (%)	VAPE (%)	Probability (error $\leq \pm 10\%$) (%)	MAPE (%)	VAPE (%)	Probability (error $\leq \pm 10\%$) (%)
Day 1	6.90	6.83	77.8	5.41	4.40	91.7	5.81	4.94	94.4
Day 2	6.99	7.03	83.3	6.99	6.24	77.8	6.34	6.37	80.6
Day 3	7.13	8.50	75.0	6.18	6.76	86.1	6.07	6.86	83.3
Day 4	7.27	6.77	75.0	6.06	5.99	86.1	6.20	5.41	88.9
Total	7.08	7.24	77.8	6.16	5.86	85.0	6.10	5.81	86.9

Los dos conjuntos de datos han sido recogidos en una autopista de Singapur, siendo uno de ellos utilizado para entrenar las dos redes neuronales, mientras que el otro para testear y comparar los resultados de los modelos individual y de forma combinada. Los resultados hacen que los autores concluyan que el modelo combinado se comporta mejor que la predicción de los modelos de forma individual (Tabla 22).

Para el entrenamiento y test de la red neuronal en [109], se disponía de un total de 130.701 datos en intervalos de 5 minutos cedidos por el ISBAK (organismo público de mantenimiento de carreteras de Turquía), que dispone de un complejo sistema de cámaras en 180 cruces de tráfico de la ciudad de Estambul. Se recogieron datos de la hora y fecha y el número de vehículos circulando en ese momento. El 18% de los mismos fue utilizado para entrenamiento y el resto para test, creando dos redes neuronales diferentes: una para la predicción del número de vehículos cada 5 minutos, la segunda para cada 60 minutos.

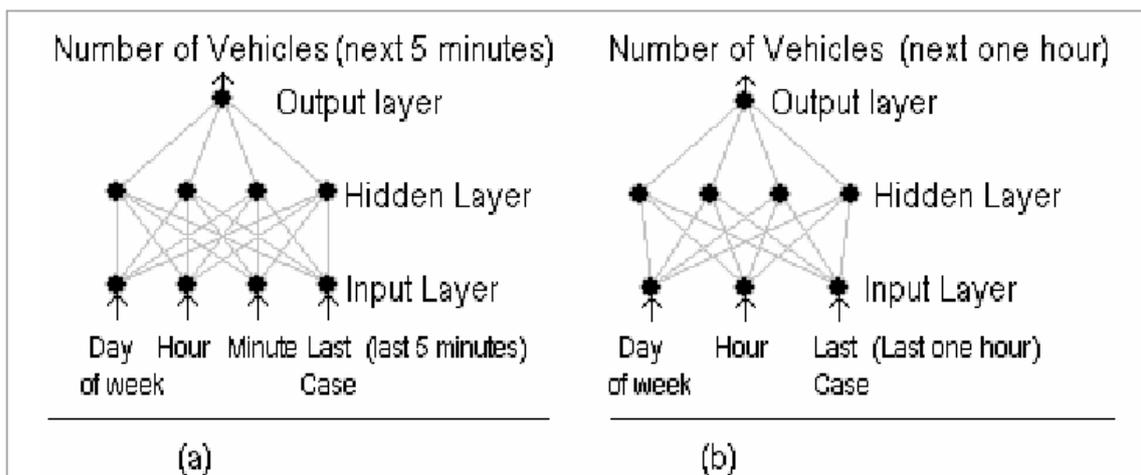


Figura 31. Redes neuronales para la predicción del flujo de tráfico en [109]

Los resultados obtenidos son buenos, obteniendo un coeficiente de correlación de entre 0.85 y 0.95 para los datos de test, no menos de 0.95 para los datos de entrenamiento. Los autores concluyen que disponer del parámetro del mes podría ser valioso para la mejora de los resultados.

En [110] se presenta una red neuronal basada en un método exponencial de suavizado para la mejora de la predicción del flujo de tráfico con otros métodos de redes neuronales usados previamente. El proceso de suavizado exponencial se realiza sobre los datos antes del entrenamiento de las redes neuronales. Los datos utilizados fueron recogidos en autopistas del oeste de Australia.

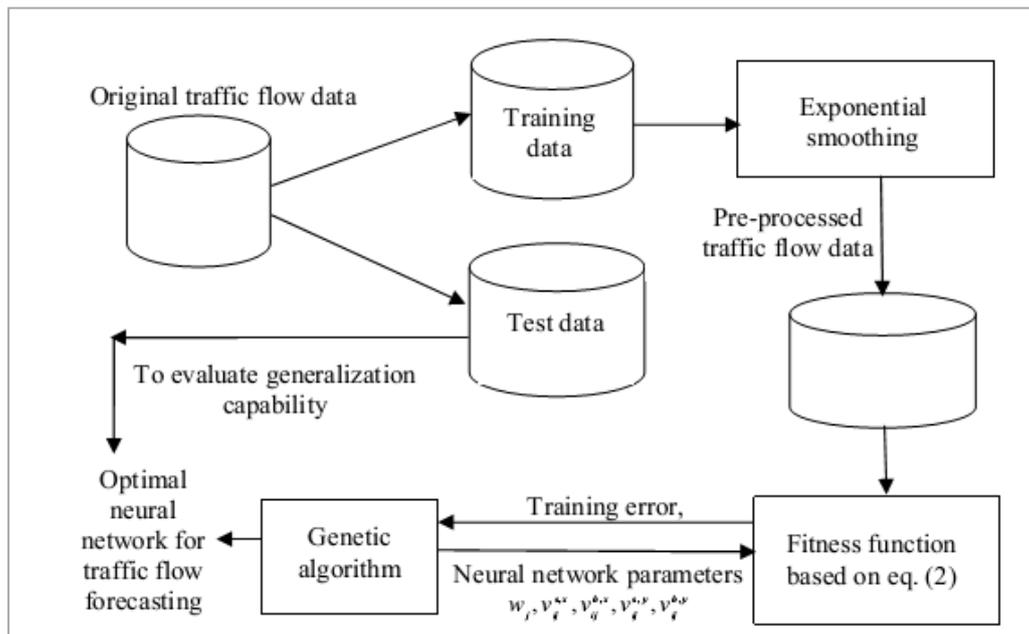


Figura 32. Flujo de desarrollo de la red neuronal con suavizado exponencial en [110]

Con el tratamiento previo de los datos, la red neuronal mejoró en una tasa del 20% los resultados de la red neuronal entrenada con los datos originales. También en cuanto a capacidad de generalización, la red entrenada con los datos tratados mejoró en un 8% los resultados de la primera.

Se incluiría asimismo en esta sección los artículos [107][108], que se describe en la sección 2.3.A por presentar sistemas neuro-difusos.

2.3 RUIDO Y OTRAS TÉCNICAS DE SOFT COMPUTING

Además de las RNA, se han usado otras técnicas de Soft Computing en el estudio y análisis del ruido. Con técnicas difusas se han localizado 6 trabajos para la predicción del nivel de ruido, 13 trabajos para el estudio de la molestia que causa el

ruido en personas, 2 para clasificar ruidos y 7 para la predicción del tráfico como principal componente del ruido urbano. Usando algoritmos genéticos, 5 trabajos versan sobre predicción del nivel de ruido y 1 para predecir el tráfico. Para clasificar ruidos se han localizado 6 trabajos que usan Modelos Ocultos de Markov. Se describen a continuación de forma resumida.

A. TÉCNICAS DIFUSAS

Los autores más prolíficos que han tratado con técnicas basadas en lógica difusa el modelado del ruido ambiental son Andy Verkeyn y Dick Botteldooren de la Universidad de Ghent, Bélgica ([117]-[127]). Su principal campo de estudio es la predicción del grado de molestia del ruido ambiental en las personas. Para ello usan modelos difusos en la mayoría de sus propuestas, buscando un conjunto de reglas difusas que describan el sistema.

PREDICCIÓN DEL NIVEL DE RUIDO

En [111], cuyos autores son en parte comunes a [81], se analizan datos de ruido de la ciudad de Catania (Italia). Se construye un modelo difuso que se compara con la función lineal siguiente:

$$L_{AeqT} = 38.8 + 15 \log(n_{eq}) - 10 \log(w).$$

Las comparaciones se hacen con dos conjuntos de datos: el que ha servido para la construcción del modelo difuso y otro conjunto que no se ha utilizado para ello. Según las conclusiones de los autores, el modelo difuso se comporta mejor que el modelo matemático, comparando según el error cuadrático medio (Tabla 23).

	Case A	Case B
Fuzzy model	1.25	0.86
Model of Eq. (3)	2.5	1.22

Tabla 23. Resultados comparativos del modelo difuso de [111] con un modelo matemático

En [112] se explica brevemente un método basado en lógica difusa para predicción de ruido urbano. Se describen tres fases principales:

1. Establecer las variables que se deben considerar en las premisas: se reducen a tres, número equivalente de vehículos, altura media de los edificios y anchura de la carretera.

2. Decidir el número de etiquetas lingüísticas de cada variable.
3. Definir el polinomio-ecuación asociado con la consecuencia de cada regla difusa.

Aunque no se dan resultados exactos, en la sección de conclusiones se indica que los resultados obtenidos por el modelo difuso se han contrastado con datos reales de medidas realizadas en los escenarios de las predicciones y han dado muy buenos resultados. Sin embargo, este modelo es de aplicación exclusiva, es decir, está especialmente diseñado para las condiciones particulares de una ciudad (su flujo de tráfico, la distribución de sus calles, la altura media de sus edificios, etc.).

Por otro lado, [115], presenta un modelo híbrido difuso y estocástico para evaluar el impacto del ruido de tráfico bajo incertidumbre, considerando como especialmente inciertos los parámetros siguientes: flujo de tráfico, velocidad del tráfico y componentes del tráfico. Estos parámetros son representados por distribuciones de probabilidad y posteriormente se llevan a cabo simulaciones de Monte Carlo para generar una distribución del ruido tras considerar otros factores influyentes. Conjuntos difusos, relaciones difusas binarias y métodos de análisis probabilísticos se aplican finalmente para identificar el impacto del ruido de tráfico tanto de manera cualitativa como cuantitativa.

En [116] se presenta un método difuso para estimar el volumen del ruido a partir de una combinación de ruidos en valores normalizados (A, B,C). Este método proporciona un indicador de volumen L_{fuz} , y es comparado con otros dos modelos de cálculo de volumen, concluyendo los autores que el modelo a evaluar devuelve resultados aceptables aproximados a los otros dos.

En este apartado también se incluye [113], que, al estar más relacionado con algoritmos genéticos, se va a definir en ese apartado concreto. Asimismo se incluye en el siguiente apartado la descripción de [114], pues versa no sólo sobre predicción de niveles de ruido, sino que se estudia directamente su influencia sobre la molestia sobre las personas.

PREDICCIÓN DEL GRADO DE MOLESTIA

Verkeyn y Botteldooren vienen dedicándose a medir el efecto del ruido en las personas desde el año 2000, aunque sus trabajos y resultados comienzan a aparecer en 2001. Se aborda el estudio creando un sencillo sistema basado en reglas difusas que predice las respuestas de la población sobre la molestia del ruido causado por

una carretera. Este modelo es revisado y mejorado en el resto de sus trabajos. También tratan temas más concretos en algunos de sus trabajos, como la generación de la función de pertenencia [125], el efecto en las personas de ruidos acumulados [121][122][118], olores [127], o el efecto del uso de la tierra (espacio utilizado para la actividad humana) [124].

En primer lugar, en 2001, presentan un modelo enfocado al ruido del tráfico de carreteras y trenes [117], que luego revisan y amplían en [121] y [126]. El modelo común en que se basan son una serie de bloques comúnmente encontrados en sistemas basados en reglas difusas, y responde a la estructura de la Figura 33.

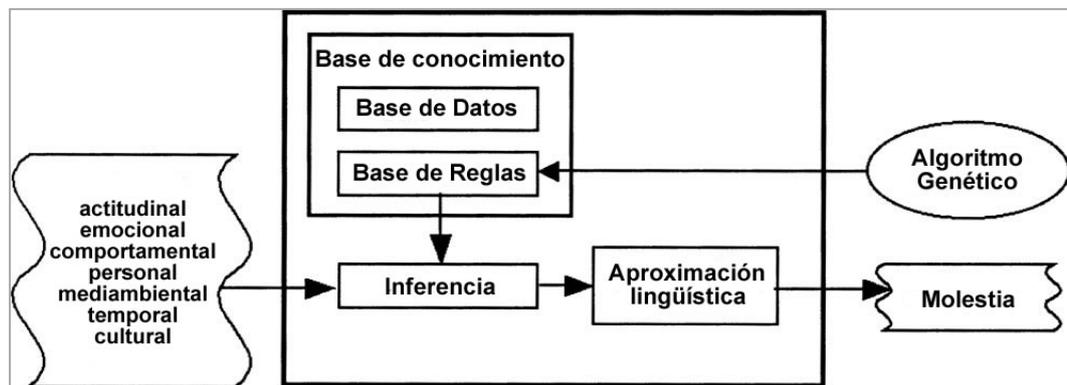


Figura 33. Modelo de predicción del grado de molestia [126][121][117]

Según muestra la Figura 33, los autores definen su modelo como una caja negra que produce como salida el grado de molestia predicho para los valores de entrada. Este modelo comprende:

- La Base de Datos, que contiene definiciones de los términos lingüísticos que serán utilizados en el modelo. Estos son: altamente molesto, moderadamente molesto, un poco molesto y nada molesto. Estos términos describen el grado de molestia de las personas encuestadas con respecto al ruido ambiental. Cada término lingüístico es representado por un conjunto difuso en un Universo U y su función de pertenencia.
- Una Base de Reglas que contiene un conjunto de reglas lingüísticas "IF-THEN", donde cada regla relaciona cada premisa con una conclusión de la forma: SI $X = A$ ENTONCES $Y = B$, donde X e Y son variables sobre los Universos U y V respectivamente, y A y B son distribuciones de posibilidad sobre U y V . En su primera versión, el modelo tenía 21 reglas.
- El motor de inferencia, usando Modus Ponens, donde B' , la distribución de la posibilidad inferida de Y , se calcula con la siguiente regla de inferencia:

$$(\forall v \in V) B'(v) = \sup_{u \in U} \min(A'(u), R(u, v))$$

R es la representación de la regla y A' es la entrada difusa dada a la regla.

- La fase de aproximación lingüística, en la que la distribución de posibilidad B' que resultó del motor de inferencia es asociada con uno de los cuatro términos lingüísticos de molestia. El algoritmo genético optimiza los pesos que se le dan a cada regla difusa (estos pesos dan un grado de certidumbre de cada regla de la base de reglas).

Para probar el modelo, en [117], se entrevistó a 2007 habitantes de Austria y los resultados fueron relativamente positivos, ya que además de que el cálculo explícito de la regla de representación restaba prestaciones de tiempo, sólo se consiguió un 40% de predicciones correctas. Aunque se trata de resultados de dudoso éxito, se ha considerado importante incluir esta descripción, ya que se trata del modelo que sentó las bases para sus sucesivos trabajos, en los que los resultados van mejorando considerablemente.

En [121], se hace un estudio más exhaustivo de otras razones por los que la molestia puede venir en aumento, ya que tienen en cuenta en su estudio, hacia qué lado dan las ventanas de la casa o la distancia a la fuente del ruido. Además, se aventuran a considerar, además del grado de molestia que el ruido causa a cada entrevistado, la medida en la que cada uno de ellos se considera "sensible" al ruido (a la que se refieren como indicadores indirectos). A partir de estos nuevos datos, obtienen reglas como "SI vive cerca carretera ENTONCES sensibilidad al ruido no es muy alta", reglas en la que se tienen en cuenta otros factores como la edad, tener niños en casa o la densidad de ocupantes de una casa, que pueden hacer más insensibles al ruido a las personas.

En [121] se obtienen resultados algo mejores que en [117], llegando a obtener un 45% de aciertos, comparando con resultados dados por métodos no difusos, supera en un 2%. Su método funciona de forma más precisa para predecir los términos lingüísticos extremos que para los intermedios.

En [126], los autores se centran en comparar los resultados obtenidos con diferentes bases de reglas difusas, llegando a la conclusión de que cuánto más difuso es el modelo (tiene de media más incertidumbre en su salida), mejor predice las respuestas de los entrevistados. En las Figuras 34 y 35 se muestra la distribución de los sujetos, donde se puede comprobar lo afirmado. El área de los círculos es proporcional al número de sujetos; en el eje de las Y se indica el porcentaje de predicciones falsas y en el de las X, el grado de no especificidad del modelo. En la gráfica de la Figura 34 se pueden observar los resultados de un modelo que es menos difuso, pero a menudo incorrecto (hay más falsos negativos), mientras que en la Figura 35 se refleja un sistema que es más difuso (tiene de media más incertidumbre en su salida), pero suele dar más a menudo la salida dada por el sujeto (por lo tanto, más aciertos).

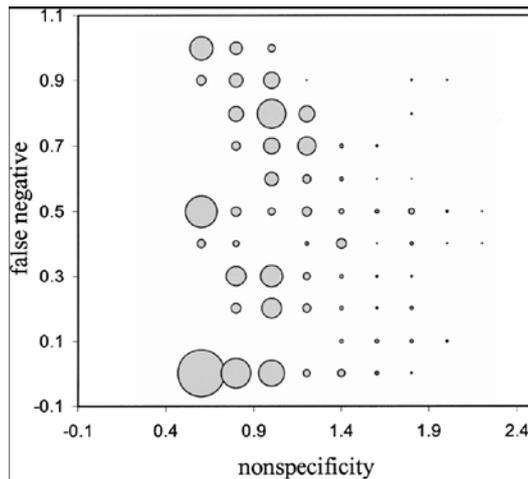


Figura 34. Distribución de los sujetos[126]

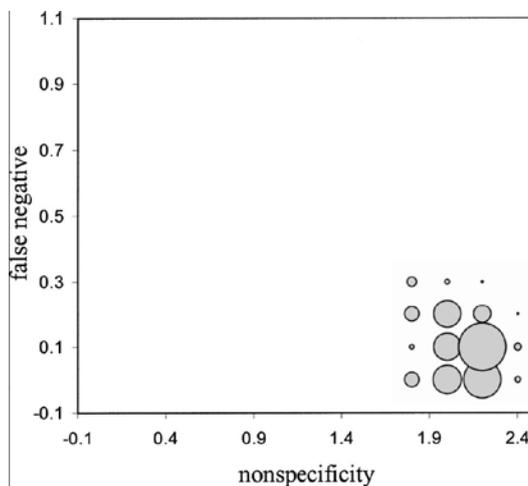


Figura 35. Distribución de los sujetos[126]

En [126], los resultados de cada categoría individualmente son notablemente mejores, llegando a obtener un máximo de un 20% de errores en la predicción, para algunas de las categorías de molestia del ruido (concretamente para un poco molesto). Pero, por otro lado, los resultados generales son algo peores, obteniendo 38.7% de acierto en la predicción de molestia por tráfico de ferrocarril (un 37.5% sin reglas que tienen en cuenta la sensibilidad, concepto introducido en [121]) y un 35.4% y 32.5% de aciertos respectivamente para tráfico de carreteras.

A partir de 2002, Verkeyn y Botteldooren ya disponen de un nuevo conjunto de datos de 3200 personas entrevistadas en la zona de Flandes, aunque siguen contando con los datos de Austria. En esta segunda encuesta, los entrevistados fueron interrogados sobre ruido, olor [127] y exceso de luz en el entorno de su residencia.

En ambas encuestas se aseguró que la muestra fuera significativa en factores demográficos como edad, género y provincia.

Respecto al cálculo del grado de molestia cuando hay varias fuentes de ruido diferentes y combinadas, en [118][122][123], se enuncian varias formas de calcularlo, entre las que destaca la forma denominada “el componente más fuerte”: el nivel de molestia es la mayor molestia causada por cualquiera de las fuentes. Probando con este y otros métodos, el “componente más fuerte” es el que da mejores resultados (un 55.5% de resultados correctamente predichos frente a 24.1%, 34.1%, 42.2% de las demás formas: vector suma, vector suma escalado y regresión lineal-respectivamente-). A partir del método escogido, se obtienen reglas para la base de reglas del tipo: SI molestia de una de las fuentes es EXTREMA ENTONCES molestia general es EXTREMA.

En la Figura 36 se observa el modelo difuso de agregación propuesto en [119] y [123] de diferentes fuentes de ruido. Las reglas se construyen mediante el proceso denominado del “componente más fuerte”, citado antes.

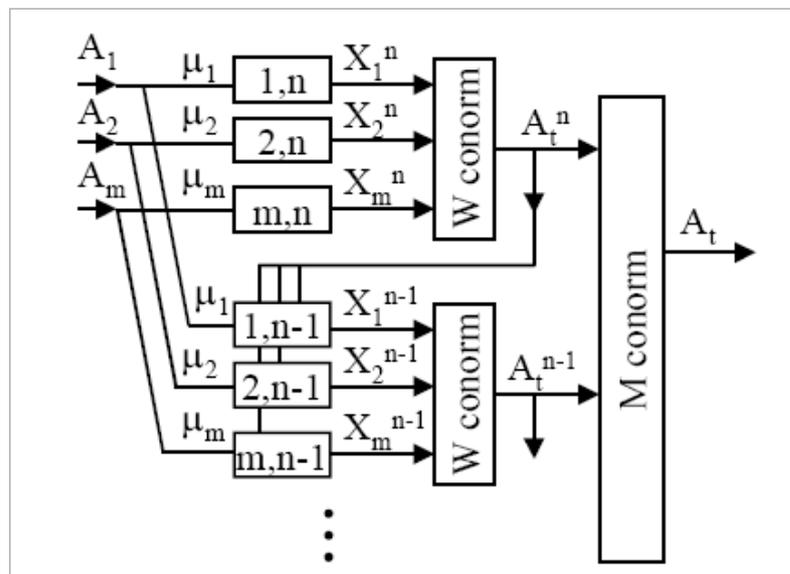


Figura 36. Modelo difuso de agregación [119][123]

El modelo contiene una regla para cada etiqueta ($i=1\dots n$ siendo n el mayor nivel de molestia), que describe un grado de molestia, representado por un conjunto difuso. Cada base de reglas contiene una regla para cada fuente de molestia. La característica más esencial del modelo es la base de reglas difusas encadenadas linealmente en la que disparar reglas de niveles altos de molestia prohíbe disparar reglas de niveles más bajos que el disparado. Dado que las reglas posibilísticas han de hacer procesos de recogida de datos, los resultados de cada regla son agregados

usando una conorma apropiada. Finalmente, los resultados de todas las bases de reglas se agregan usando una conorma para calcular la salida final. Tras este proceso, esta salida se asocia a una etiqueta lingüística para comparar con los resultados de la encuesta.

Se han desarrollado también dos estudios sobre la generación de funciones de pertenencia, en [120] y [125]. En [125] se describen varios métodos para la construcción de las funciones de pertenencias a partir de medidas tomadas de experimentos psicolingüísticos, poniendo especial atención a las interpretaciones inclusivas (primera función de pertenencia –para nada molesto- decreciente, y resto crecientes) y no inclusivas (primera función de pertenencia decreciente, última –totalmente molesto- creciente, e intermedias con forma de campana) de los términos lingüísticos. Obtienen cuatro términos lingüísticos -nada molesto, un poco molesto, moderadamente molesto y totalmente molesto- de los que obtienen las funciones de pertenencia que se observan en la Figura 37. Las líneas punteadas indican las curvas construidas inicialmente y las sólidas, las que finalmente adoptaron para evitar formas abruptas, dada la vaguedad de los términos.

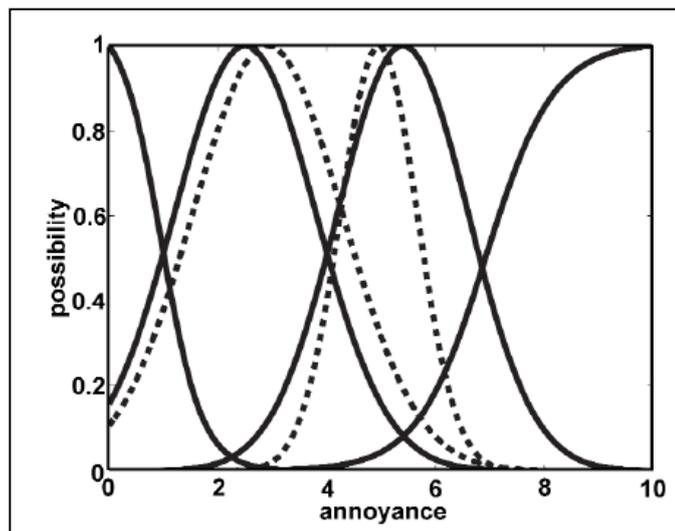


Figura 37. Funciones de pertenencia de los cuatro términos lingüísticos indicados en [125] y [120]

Por otro lado, en [120], se discute el modo matemático de traducir las etiquetas lingüísticas en nueve idiomas (incluido el español) sin que estas pierdan el significado original, ayudándose de la teoría de conjuntos difusos. Su objetivo en [120] es definir etiquetas que puedan ser usadas en encuestas en varios idiomas y que los resultados sean precisos y sobre todo comparables.

En [114] se propone un modelado de la molestia por ruido usando técnicas difusas. Los autores defienden el modelo al considerar importante poder definir las entradas y salidas como reglas SI-ENTONCES, de forma que la molestia es considerada como una función del nivel de ruido, su duración y el estatus socio económico de una persona. Los resultados del modelo propuesto son aplicables a

áreas urbanas de la India. Entre otras conclusiones interesantes, los autores destacan que la población de mayor estatus social es más susceptible al ruido.

En esta sección se incluye también [91], ya comentado en el apartado 2.2.B, ya que se trata de un modelo neuro-difuso.

CLASIFICACIÓN

En [128] se propone un clasificador de ruidos urbanos considerando un conjunto de características acústicas, obteniendo siete categorías: bus, coche, tren, construcción, personas hablando, fábrica y calle. Para ello se obtiene, en el siguiente orden:

1. Un conjunto de criterios para agrupar un gran grupo de ruidos ambientales en un conjunto reducido de clases de ruido con similares características acústicas: se usa un método de reconocimiento de patrones.
2. Un conjunto de ruidos de fondo junto con una arquitectura de clasificación multinivel: tras la fase de selección de características, se utiliza, para cada nivel de clasificación, un sistema difuso (en total 6), como se observa en la Figura 38.
3. Un conjunto de parámetros acústicos: finalmente, en un proceso de selección de características, a partir de 100 parámetros, se obtienen un total de quince que serán usados por los sistemas difusos a la hora de diferenciar las siete categorías sonoras.

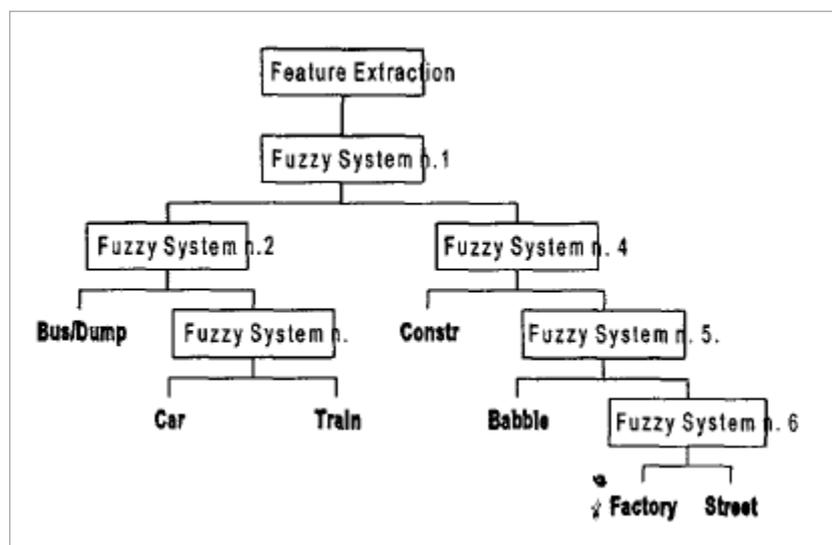


Figura 38. Arquitectura de clasificación multinivel de [128]

Los resultados del modelo son aceptables para la mayoría de los sonidos a distinguir, consiguiendo en 6 de ellos un porcentaje de aciertos superior al 76,4%. En

cambio, para reconocer el sonido de la calle no se obtienen los resultados esperados, cayendo la precisión al 58,3%.

En [129] se propone un clasificador de ruido ambiental basado en lógica difusa. Dicho clasificador usa parámetros que se extraen con una función de autocorrelación a corto plazo. Los ruidos que se reconocen son: metro, aeropuerto, coche, tren, restaurante y lluvia.

En la Figura 39 puede observarse el diagrama de bloques del modelo difuso utilizado en [129]. Los subsistemas difusos 1 a 3 poseen cada uno 6 reglas, una salida y 6 entradas con dos funciones de pertenencia cada una. En el bloque de decisión se aplica un algoritmo de eliminación. Si las diferencias entre los valores de pertenencia de las salidas son cercanas unas a otras, se vuelve a aplicar la lógica difusa a estos casos. La precisión del modelo varía desde el 62% al 90%.

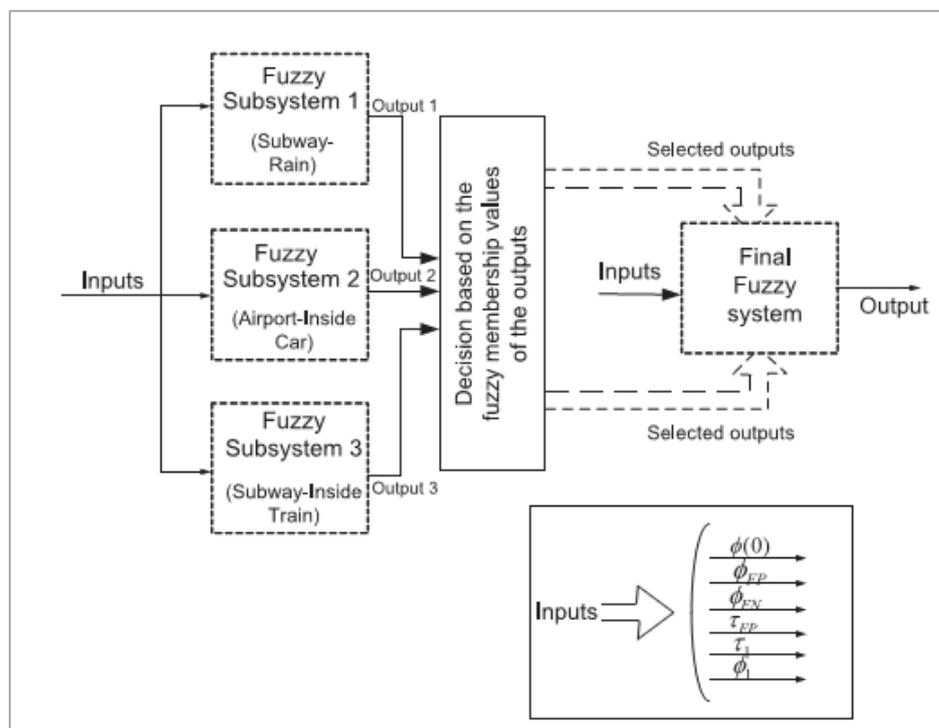


Figura 39. Diagrama de bloques del modelo difuso de [129]

PREDICCIÓN DEL FLUJO DE TRÁFICO

Incluyen modelos difusos para predecir el flujo de tráfico [130] y [131], ambos de los mismos autores. Se presenta en ellos un sistema adaptativo híbrido basado en reglas difusas (FRBS) para modelar y predecir a corto plazo el flujo de tráfico.

En [130] se emplea un algoritmo genético para modificar convenientemente los parámetros envueltos en las funciones de pertenencia. Para evaluar los resultados de su capacidad de predicción, se compara con el filtro de Kalman y el modelo ARIMA, siendo en la mayor parte de los casos más preciso el sistema basado en reglas difusas que presenta el artículo.

En [131], combinando las respuestas de un modelo basado en redes neuronales y otro basado en filtro de Kalman, el sistema difuso obtiene reglas para predecir el flujo de tráfico con una mayor precisión a los dos primeros (Figura 40).

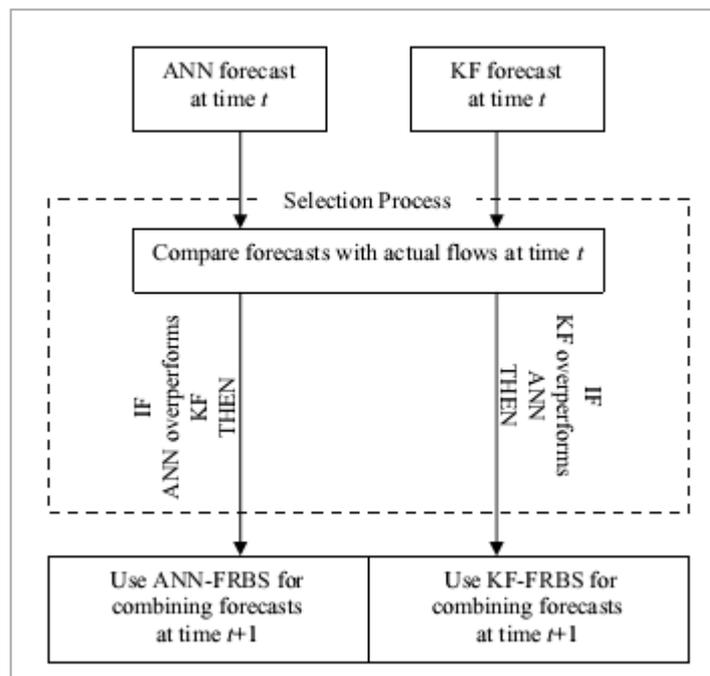


Figura 40. Proceso de combinado de modelos predictivos con un FRBS en [131].

Por otro lado, el sistema POPFNN-TVR (pseudo outer-product fuzzy neural network – true value restriction method) es presentado en los artículos [107][108] con el objeto de predecir el flujo de tráfico a corto plazo. Se trata de un sistema híbrido neuro-difuso que, perteneciendo a la clase de sistemas difusos que se construyen automáticamente usando datos numéricos, sus parámetros se ajustan usando técnicas de redes neuronales. Las funciones que realiza cada capa del sistema corresponden a inferencias hechas por el método TVR (Truth-Value-Restriction)[147].

Los resultados de predicción del flujo de tráfico son comparados con los devueltos por una red neuronal de tipo PMC (FFBP) (Tabla 24), siendo peores los

resultados de esta última. Además, el sistema parece comportarse también de forma aceptable cuando los datos son mezclados con un 10% y un 30% de ruido.

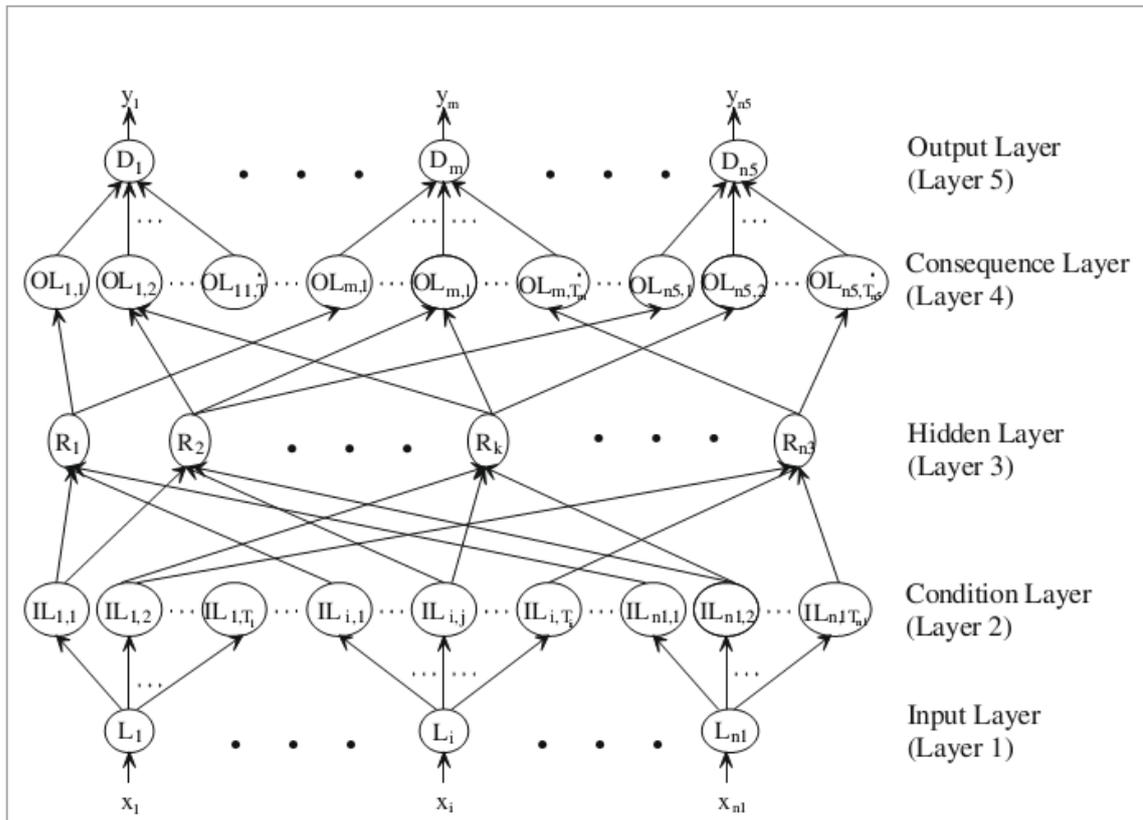


Figura 41. Estructura del sistema POPFNN-TVR de [107][108]

Tabla 24. Resultados del sistema POPFNN-TVR comparados con FFBP de [108]

Lanes	R^2 of speed		R^2 of density	
	FFBP	POPFNN-TVR	FFBP	POPFNN-TVR
1	0.298021	0.282159	0.746806	0.675296
2	0.62021	0.592842	0.79219	0.768371
3	0.10224	0.010045	0.838725	0.796968

En esta sección también se pueden incluir [98], [102] y [103] ya comentados en el apartado 2.2.D, ya que tratan de modelos neuro-difusos.

B. ALGORITMOS GENÉTICOS

Los trabajos encontrados que usan algoritmos genéticos son en su mayoría aplicaciones a la predicción del nivel de ruido, excepto [104], que se aplica al flujo de tráfico y que se describió en la sección 2.2.D por incluir una optimización genética en la red neuronal.

En [113] se expone un método basado en algoritmos genéticos con el que se busca la optimización de un sistema basado en reglas difusas para el modelado del ruido ambiental. El trabajo da ideas que contribuyen a la mejora de la predicción de resultados usando algoritmos genéticos.

A partir del caudal de vehículos, la altura y la anchura de la calle, la idea es obtener reglas difusas que describan la relación entre estos tres parámetros. El algoritmo de búsqueda ha evaluado doce modelos, siendo el que dio mejores resultados el "M440" (4 conjuntos difusos para el número equivalente de vehículos y la altura media de la calle, mientras que la anchura de la calle no aparece en ninguna premisa del sistema obtenido).

Para demostrar los resultados de su método, se compara en una gráfica el resultado de su sistema con el que daría la herramienta AFM (Adaptive Fuzzy Modeller). En esta gráfica (se muestra en la Figura 42) se observa, con línea de puntos, el original, con línea continua el sistema presentado en [113] y con "- ." el resultado de AFM. En el eje de abscisas está representado el tiempo, y en el de ordenadas, el nivel de presión sonora.

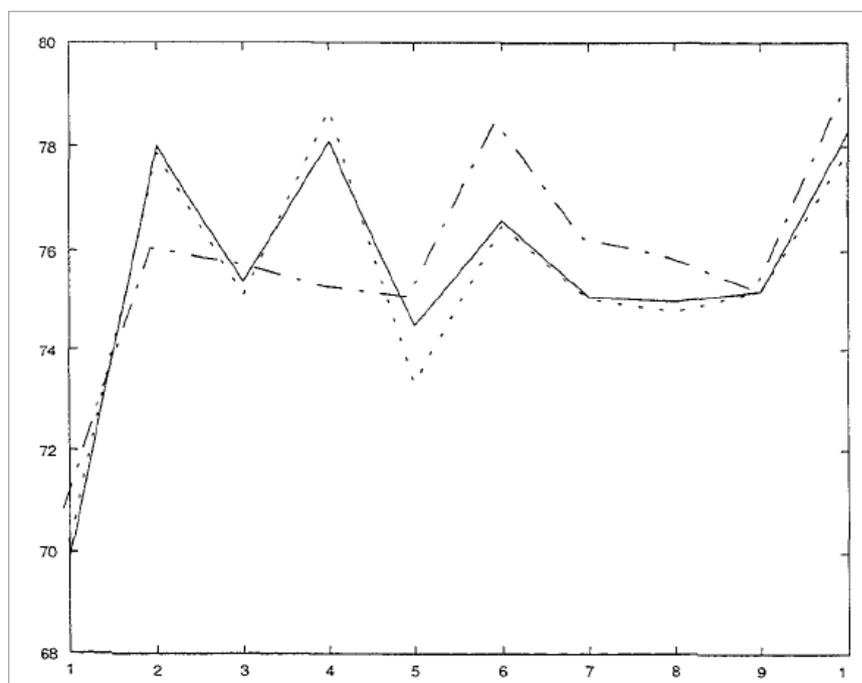


Figura 42. Gráfica comparativa para el modelo de [113]

Por otro lado, en [132], también se presentan dos modelos matemáticos de predicción del nivel de presión sonora optimizados con algoritmos genéticos. Al igual que en [133], uno de ellos es variante en el tiempo (TV) y el otro no (TIV). En este caso, se comparan los resultados de ambos modelos con un modelo estadístico [148], el modelo inglés CoRTN[17] y una versión de CoRTN revisado [149]. En la Tabla 25 se pueden apreciar los resultados de las comparaciones. Los resultados de los modelos presentados son variados, aunque en general obtienen mejores comportamientos que los de los otros métodos.

Tabla 25. Comparación de resultados de los distintos modelos de [132]

Method	Location	Minimum	Mean	Maximum	Standard deviation	Root-mean-square	Pearson R correlation
GA-TIV	P1	-8.66	-2.88	2.89	8.17	6.46	0.66
	P2	-9.72	-5.26	-0.81	6.30	6.89	0.28
	P3	-3.03	0.31	3.65	4.73	3.36	0.83
	P4	2.34	3.03	3.72	0.98	3.11	0.14
GA-TV	P1	0.68	1.08	1.48	0.56	1.15	0.96
	P2	-0.57	-0.12	0.32	0.63	0.46	0.81
	P3	2.42	3.96	5.50	2.18	4.25	0.76
	P4	3.28	8.21	13.14	6.97	9.58	0.30
Statistical model	P1	5.56	10.85	16.15	7.49	12.08	0.34
	P2	10.56	13.67	16.78	4.39	14.02	0.66
	P3	11.73	15.66	19.60	5.57	16.15	0.78
CoRTN	P4	18.15	20.17	22.20	2.86	20.28	0.62
	P1	8.80	13.96	19.11	7.29	14.88	0.33
	P2	14.70	17.47	20.23	3.91	17.68	0.66
	P3	14.58	18.80	23.02	5.97	19.27	0.78
	P4	21.28	23.79	26.31	3.55	23.93	0.63
Revised CoRTN	P1	2.69	8.18	13.67	7.76	9.85	0.34
	P2	8.56	11.61	14.65	4.30	12.00	0.65
	P3	8.94	13.28	17.62	6.14	13.97	0.78
	P4	15.89	18.34	20.80	3.47	18.51	0.63

En[133]se muestran dos modelos matemáticos para predecir niveles de ruido de tráfico, teniendo en cuenta el volumen de tráfico, la composición del mismo y la velocidad. También los vehículos fueron separados entre ligeros, medianos y pesados. El primero de los modelos se ha denominado TID (time independent), pues no se considera el efecto directo del tiempo en su cálculo, mientras que el segundo se denomina TD (time dependent), ya que sí lo hace. Algoritmos genéticos han sido utilizados para optimizar los modelos matemáticos, siguiendo el esquema de la Figura 43.

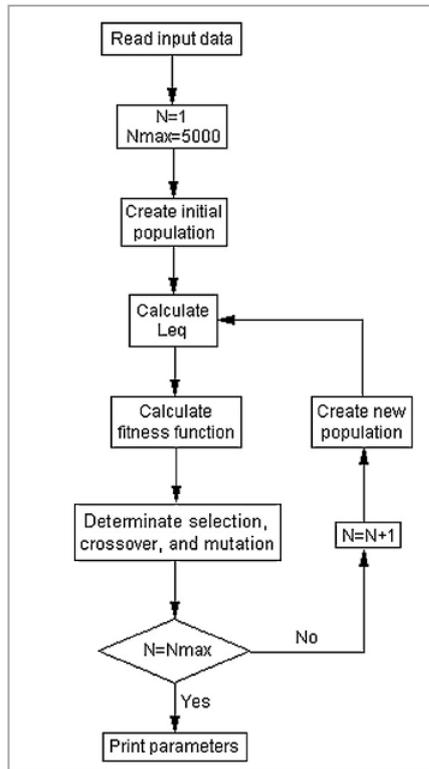


Figura 43. Diagrama de flujo del proceso de optimización mediante AG en [133]

Ambos modelos han sido evaluados con datos de ruido y presentan un coeficiente de correlación que varía según los datos entre 0.55 y 0.97.

También en este apartado también se incluyen los artículos [126][121][117], definidos en la sección 2.3.A, ya que se tratan de modelos difusos con algoritmos genéticos.

C. MODELOS OCULTOS DE MARKOV

Los Modelos Ocultos de Markov (o HMM del inglés Hidden Markov Models) es otra técnica incluida normalmente dentro del Soft Computing. Los trabajos encontrados son en su totalidad aplicaciones para la clasificación de ruidos.

En[134] y[135] se desarrolla un sistema clasificador de ruido basado en Modelos Ocultos de Markov. Ambas publicaciones describen cómo los HMM pueden ser usados para construir un sistema de reconocimiento de ruido ambiental basado en el análisis frecuencial en el tiempo de las señales sonoras. El sistema sigue la estructura de la Figura 44: un preprocesador proporciona la representación frecuencial en el tiempo de la señal sonora, que es posteriormente usada por un clasificador. El clasificador, que es el elemento basado en HMM, tomará una decisión

según la naturaleza de la fuente del ruido, de acuerdo a las características dadas por el preprocesador.

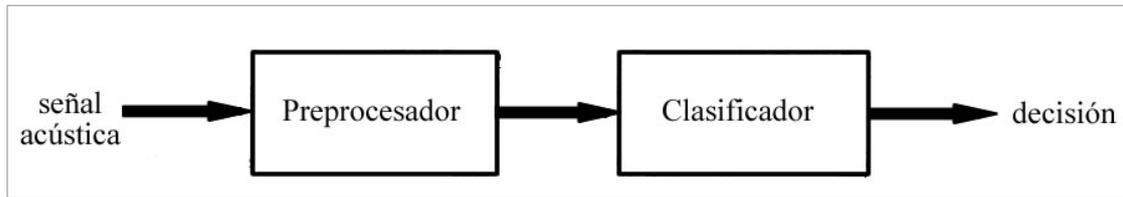


Figura 44. Estructura del sistema de [134] y [135]

El comportamiento del modelo es evaluado de forma experimental para la clasificación de cinco tipos de ruido: coche, ciclomotor, camión, avión y tren. Posteriormente es comparado con las respuestas de seis personas tras una prueba de laboratorio. Las personas elegidas fueron en su mayoría ingenieros con experiencia en control del ruido, por lo que se espera de ellos unos resultados superiores a la media. Clasificando con el mismo conjunto de datos, el grupo de personas elegidas fue superado por el sistema, que obtuvo un 95% de aciertos.

En [136] y [137], igualmente en el marco de los Modelos Ocultos de Markov, se describe un clasificador de 11 tipos de ruido diferentes, logrando unos niveles de precisión muy similares a [134] y [135]. Se trata de una topología de izquierda a derecha, y el número de estados, que normalmente varía según el número de datos de entrenamiento disponible y la duración de la señal a ser clasificada, varió entre 3 y 21 estados. El entrenamiento se hizo con sesenta ejemplos de ruido provenientes de once escenas diferentes, lo que da un conjunto de once modelos ocultos de Markov inicial. La precisión varía según el número de estados, subiendo hasta un 92% con 11, 13 y 15 estados, según se muestra en la Figura 45.

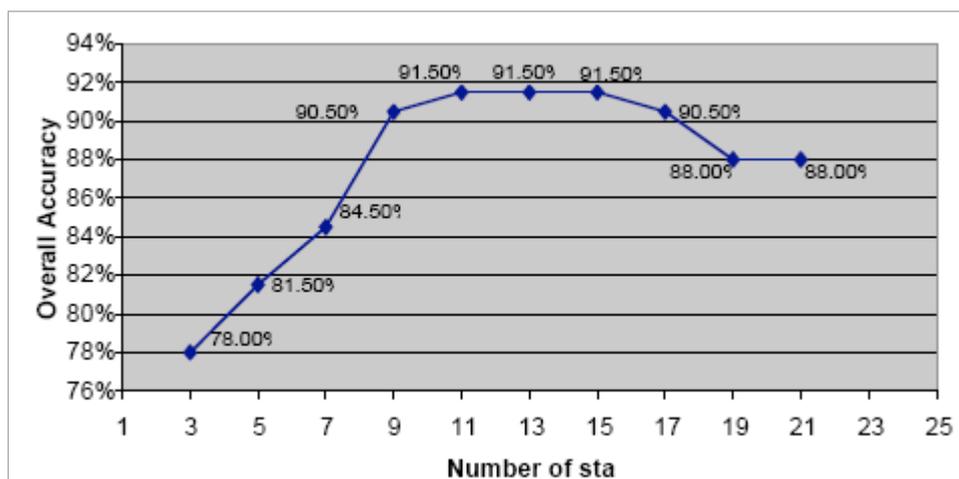


Figura 45. Evolución de la precisión según el número de estados en [136] y [137]

Tabla 26. Precisión del modelo propuesto en [136] y [137] según el tipo de sonido

Accuracy, %	Bar	Beach	Bus	Car	Football M.	Laundrette	Lecture	office	Rail Station	Street
Bar	85	0	15	0	0	0	0	0	0	0
Beach	0	100	0	0	0	0	0	0	0	0
Bus	0	0	95	0	5	0	0	0	0	0
Car	0	10	0	85	0	0	0	0	0	5
Football M.	0	0	0	0	100	0	0	0	0	0
Laundrette	0	0	0	0	0	100	0	0	0	0
Lecture	0	0	0	0	0	0	85	0	0	15
Office	0	0	0	0	0	0	0	100	0	0
Rail Station	0	0	0	0	0	0	0	0	90	10
Street	0	0	0	0	0	10	0	0	15	75
Overall accuracy: 91.5%										

Finalmente se apostó por 11 estados, dado que se minimiza el número de estados y se maximizan los aciertos en la clasificación. Los resultados obtenidos se observan en la Tabla 26.

Por otro lado, en [97] se presentan dos formas de mejora del comportamiento de sistemas automáticos de reconocimiento de ruido (ANR). Los dos métodos propuestos se aplican sobre muestras de ruido obtenidas en el entorno del hogar de doce familias.

En primer lugar, se propone la minimización del número de parámetros de un HMM clasificador de ruidos según un conjunto de datos de entrenamiento, con el objeto de reducir su complejidad. Para ello se usa el modelo Gaussiano de reducción [150]. La aplicación de esta mejora a sistemas de reconocimiento de ruido da excelentes resultados, que se observan en la Figura 46. Se dispone del modelo inicial, que es aquél con el cual se van a comparar el resto de modelos obtenidos con este método. En la gráfica se puede comprobar que se obtiene más precisión en el reconocimiento cuando el modelo ha sido adaptado y además se le ha aplicado una técnica de clustering.

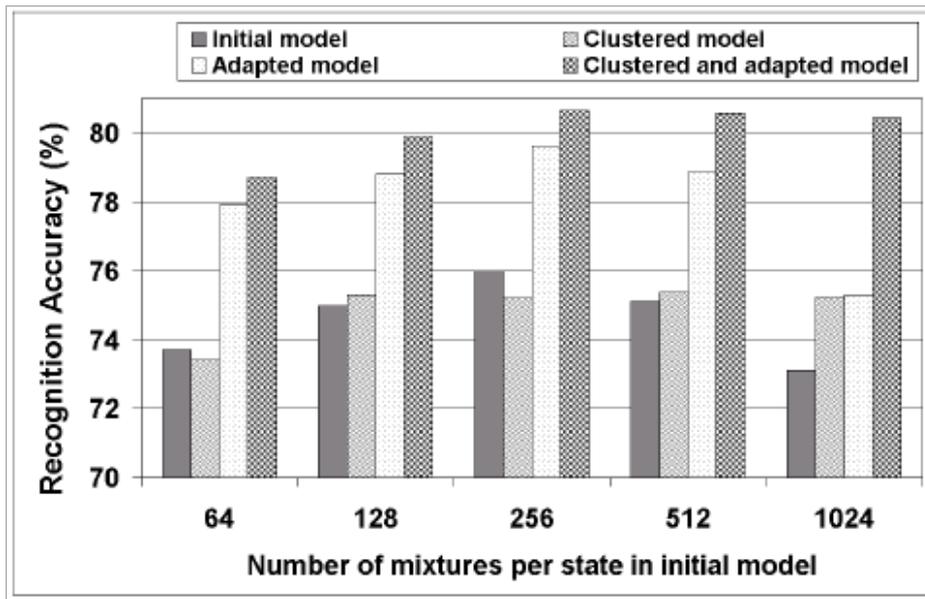


Figura 46. Comparativa del primer método de [97]

En segundo lugar, se pretende combinar los resultados de sistemas de reconocimiento diferentes, aplicando el método de la combinación de expertos (MoE, Mixture of Experts). Para ello se usan redes neuronales, que se aplican para combinar los resultados y tomar una decisión final sobre la categoría de la señal. La estructura de este sistema se presenta en la Figura 47.

En el experimento que comprueba este método, se obtuvo un sistema ANR para cada tipo de ruido a discriminar, cada uno de los cuales era el mejor en el tipo de ruido correspondiente. Como base para comprobar los buenos resultados se tomaron los mejores resultados de estos sistemas ANR. Posteriormente se construyeron dos tipos de MoE: el primero de ellos (A), con una red capaz de separar todos los ruidos; y el segundo (B), una red de discriminación en dos pasos, en la que inicialmente se separa sonido verbal (TV, personas) de no verbal (cocina, pasos) y después cada uno va a una red experta especialista en el tipo de sonido correspondiente, que hace la clasificación final.

Los resultados se aprecian en la Figura 48, siendo óptimos en el MoE B para el conjunto de entrenamiento (79.8% de precisión), y en MoE A (71.3% de precisión) para el conjunto de test.

En esta sección también se incluye [96], que fue descrito en el apartado 2.2.C, ya que también se basa en redes neuronales artificiales.

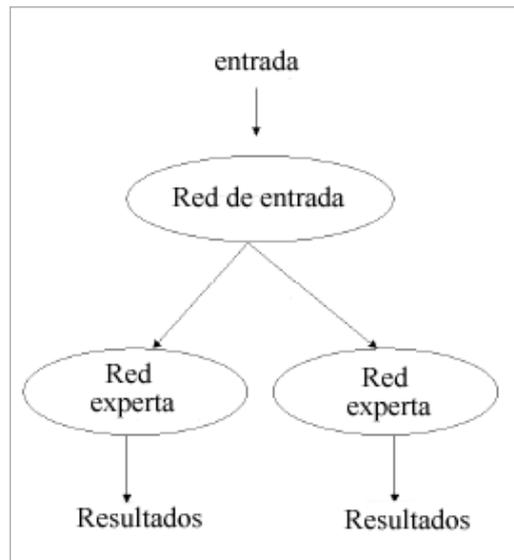


Figura 47. Estructura del Sistema MoE de [97]

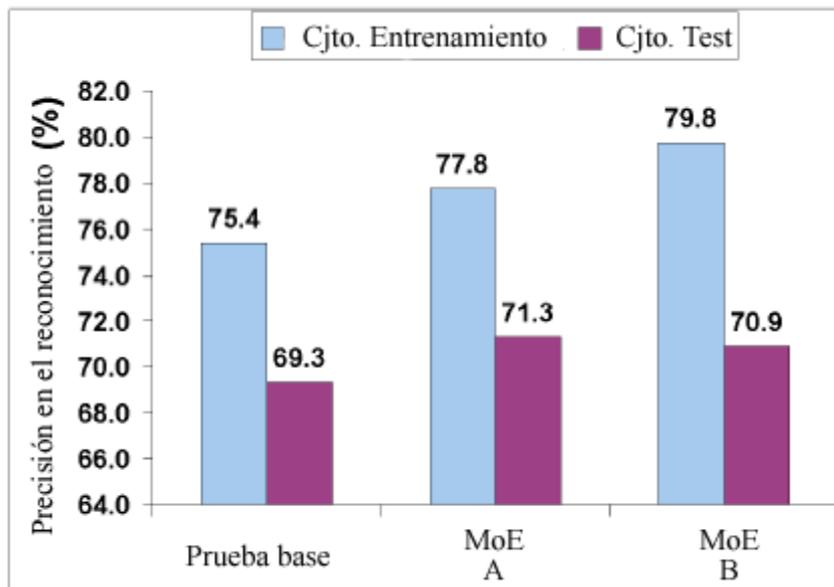


Figura 48. Comparativa del segundo método de [97]

**CAPÍTULO III: REDES NEURONALES
PARA LA PREDICCIÓN DEL RUIDO
URBANO**

Hoy día se utilizan algunos métodos como los descritos en el capítulo I, nombrados como métodos matemáticos porque utilizan una ecuación o función matemática, para la predicción del nivel de ruido urbano. Uno de los problemas que tienen estos métodos es que utilizan poca información ya que generalmente consideran sólo el flujo de tráfico, con algunas modificaciones en función de los valores de unas pocas variables de entrada, diferentes para cada método.

La precisión de estos modelos no suele ser muy alta, bien porque no consideran toda la información disponible, bien porque la función evaluadora no es suficientemente adecuada, aunque algunos mejoran cuando se hace una estimación de algunos de los parámetros a medida del sitio de medición del ruido.

Por ello, es necesario y/o conveniente considerar otras metodologías. Metodologías basadas en Técnicas de Soft Computing han obtenido buenos resultados en muchos y diferentes problemas, por lo que pueden ser adecuadas también para analizar el ruido urbano. Por ejemplo, las Redes Neuronales son una metodología que puede obtener buenos resultados en la predicción del ruido, ya que habitualmente se obtienen buenos resultados en la identificación de sistemas.

En este capítulo se plantea entonces la predicción del nivel de ruido urbano mediante Redes Neuronales Artificiales. Para ello, se realiza un proceso de obtención de datos de ruido urbano en la ciudad de Granada, en el que se incluyeron un total de 12 vías de diferentes características, y se obtuvieron un total de 289 registros (mediciones en diferentes momentos en cada vía), que incluyen 25 variables de entrada y como salida el nivel de ruido equivalente, medido en decibelios con un sonómetro.

Las 25 variables consideradas se obtienen mediante la unión conjuntista de las variables que utilizan los diferentes métodos matemáticos, y la división de algunas de ellas, según se considere la vía o el flujo, como ascendente o descendente. Entonces se procede a diseñar una estructura de RNA adecuada para la resolución del problema, considerando una RNA tipo Perceptrón Multicapa, con 25 entradas y una salida.

Se realizan varios experimentos con distintas estructuras de la RNA (considerando diferente número de neuronas ocultas en la única capa intermedia), combinando distintos algoritmos de entrenamiento, para seleccionar la mejor estructura de la RNA. Una vez elegida la estructura definitiva para la red neuronal, se entrena la misma con distintos conjuntos de 200 datos obtenidos aleatoriamente, y los otros 89 se utilizan para el conjunto de test. Completado este proceso, se analizan los resultados, comparando éstos con los obtenidos por los modelos predictivos de ruido urbano existentes.

3.1 PROCESO DE OBTENCIÓN DE DATOS

Antes de diseñar una red neuronal artificial para predecir el nivel de ruido urbano, es necesario disponer de un conjunto de datos suficientemente amplio y descriptivo del entorno que se quiere estudiar. En esta sección se describe el proceso de toma de datos y posterior tratamiento de los mismos.

Para obtener los datos de forma correcta, se ha colaborado con investigadores de los Departamentos de Física Aplicada e Ingeniería Civil de la Universidad de Granada, a los que se agradece, sobre todo, la toma física de los datos en las calles de Granada³.

HERRAMIENTAS

Aunque en realidad existe una amplia gama de instrumentos acústicos diseñados para realizar mediciones, el sonómetro es el único aparato que permite obtener directamente el valor L_{eq} . Uno de los componentes, sin duda importante, en

³Se le agradece al Dr. Ruiz Padillo de la UGR y su grupo de investigación.

el nivel sonoro es el ruido originado por el tráfico rodado, que varía a lo largo del tiempo, por lo que los instrumentos más útiles son los sistemas preparados para una medición continua de L_{eq} .

Para la obtención de los datos de entrada se han utilizado dos sonómetros, ambos del tipo I y del mismo modelo, el sonómetro integrador *2260 ObserverTM*, equipado con el software *BZ7219*, Que cumple las siguientes normas:

- IEC 60651 (1979) Tipo I, más Adendums 1 y 2.
- IEC 60804 (2000) Tipo I.
- IEC 61672 (Borrador de marzo de 2001) Clase 1.
- IEC 61260 (1995) Bandas de octava y bandas de 1/3 de octava Clase 0.
- ANSI S1.4-1983 Tipo I más Adendums
- ANSI S1.4A-1985.
- ANSI S1.43-1997 Tipo I.
- ANSI S1.11-1986 Bandas de octava y bandas de 1/3 de octava, Orden 3, Tipo 0-C, intervalo opcional.

TIEMPO E INTERVALOS

Las medidas objetivas de los niveles de sonido son una parte indispensable de cualquier programa de protección contra el ruido ambiental. Los niveles de ruido ambiental varían espacial y cronológicamente, ya que el ruido es a menudo impulsivo o puede contener tonos puros. Además, las molestias procedentes de fuentes de ruido externas – ladridos de perro, vuelo de aviones, niños jugando – deben tratarse de formas diferentes.

El resultado de una evaluación de un ruido nunca es una simple cifra como 77 dB. Es el valor de los parámetros o indicadores específicos obtenidos bajo unas condiciones conocidas y bien documentadas.

Se eligen diferentes intervalos de tiempo dependiendo de los objetivos perseguidos y de las características de la circulación. En la elección de los períodos de medida existen varias tendencias:

- Encontrar las horas de mayor tráfico y medir para obtener el valor medio de ese período.
- Medir durante el tiempo correspondiente al paso de, al menos, un cierto número de vehículos ligeros y/o pesados y considerar los resultados obtenidos como la energía sonora característica de la carretera.
- Medir durante largos períodos (más de 24 horas).

Lo ideal es medir el ruido durante el intervalo temporal de referencia completo. Como regla general, los períodos de medición deben ser tan largos como sea necesario para conseguir un buen conocimiento de la evolución del ruido durante un día, una semana o una estación, teniendo en cuenta las condiciones atmosféricas de la zona. A veces por determinadas razones se utilizan períodos más largos tales como

mediciones de un mes y un año. En dichos casos para obtener un historial de tiempo de los niveles de ruido, se utiliza un registro de valores obtenidos cada segundo, minuto o cuarto de hora. Sin embargo, las mediciones de larga duración pueden ser caras y difíciles de gestionar. Normalmente esas evaluaciones se basarán en tomar mediciones de muestras representativas y con ellas, extrapolar una visión completa y general.

De todo esto se extrae que el tiempo de medición no tiene límite. Sin embargo, como deben realizarse con frecuencia estudios y evaluaciones en cortos períodos de tiempo, la mayoría de las normas nacionales han fijado mínimos intervalos de tiempo para medir, que dependerán del propósito de la evaluación y suele estar entre quince minutos y una hora. En las mediciones para este estudio se han hecho durante un tiempo de entre 25 y 70 minutos, dependiendo del *tiempo de estabilización*, que ha oscilado entre 5 y 53 minutos. El tiempo de estabilización es el tiempo a partir del cual no tiene sentido seguir midiendo puesto que no van a obtenerse datos diferentes del nivel de L_{Aeq} , ya que se considera que se ha estabilizado. Sobre el estudio del tiempo mínimo de medición y del tiempo de estabilización se han publicado varios trabajos entre los que se destacan [8] y [151].

VARIABLES QUE SE MIDEN

Ya en el capítulo I se dio una definición de los elementos que influyen en el nivel de presión sonora y/o que han sido utilizados por algunos de los métodos analizados en el capítulo I. Son los siguientes:

- Caudal de vehículos.
- Velocidad de los vehículos.
- Distancia desde la fuente de ruido al oyente.
- Ángulo de visión.
- Tipo de vía.
- Anchura de la vía.
- Altura de la vía.
- Pendiente de la vía.
- Tipo de flujo del tráfico.
- Tipo de pavimento.

Dado que la ciudad de Granada es especialmente ruidosa y de una fisonomía muy particular, los expertos en ruido ambiental han recomendado tener en cuenta una serie de consideraciones especiales para el estudio:

- Se distinguirá entre el número de carriles ascendente y descendente, así como el estado del pavimento.
- Además de la anchura de la vía, se tendrá en cuenta la anchura de la calzada.
- Dada la gran influencia de la pendiente de la vía en la producción de ruido por parte de los automóviles, se diferenciará entre el caudal de vehículos ascendente y descendente para vehículos ligeros, pesados y motocicletas. Esto implicará 6 variables que engloban el caudal de vehículos en general.

- Se tendrá en cuenta la aparición de eventos anómalos, separados en: eventos anómalos relacionados con el tráfico y eventos anómalos no relacionados con el tráfico. Por otro lado, se tendrá en cuenta el número de vehículos con sirena observados durante la toma. Los eventos anómalos son eventos sonoros cortos y de gran intensidad, como se describió en la definición del modelo lineal multivariante granadino en el capítulo I.
- Las condiciones del entorno son también importantes, teniendo en cuenta la presencia de obras y el tipo de entorno (comercial y de ocio o no).
- Se considerará el período del día en el que la toma fue efectuada (día o noche), así como el tiempo de estabilización (descrito en el apartado anterior, Tiempo e Intervalos).
- Se descarta el ángulo de visión, dado que su influencia es despreciable.

Estas consideraciones dan lugar a 25 variables [152], mostradas en la Tabla 27.

Tabla 27. Variables de entrada consideradas por la red

1	Período del día	14	Aparición de eventos anómalos no relacionados con el tráfico
2	Entorno Comercial o de Ocio		
3	Presencia de Obras en la zona	15	Velocidad media de los vehículos
4	Tiempo estabilización	16	Pendiente de la calzada
5	Tipo de flujo	17	Número de carriles ascendente
6	Caudal Vehículos Ligeros Ascendente	18	Número de carriles descendente
7	Caudal Vehículos Ligeros Descendente	19	Tipo de pavimento
8	Caudal Motocicletas Ascendente	20	Estado del firme
9	Caudal Motocicletas Descendente	21	Tipo de vía
10	Caudal Vehículos Pesados Ascendente	22	Anchura de la vía
11	Caudal Vehículos Pesados Descendente	23	Altura media de los edificios de la vía
12	Número de vehículos con sirena	24	Anchura de la calzada
13	Aparición de eventos anómalos relacionados con el tráfico	25	Distancia de la fuente de ruido al receptor

PUNTOS DE MEDICIÓN

El número y situación de los puntos de medición necesarios para identificar el sonido medioambiental de una zona, depende del tipo de mediciones que se vayan a realizar. Los criterios de selección pueden resumirse en dos generales:

- Seleccionar puntos en donde sea probable que el ruido del tráfico pueda perturbar a la población.
 - Seleccionar los puntos que sean representativos de las diferentes situaciones y condiciones de la zona.

La toma de datos en la ciudad de Granada se llevó a cabo en las siguientes calles:

Camino de Ronda: En esta calle, caracterizada por el denso tráfico y tipo de vía en “U” (edificios a ambos lados), se hizo tomas en dos lugares diferentes, que se muestra en el mapa de la Figura 49.

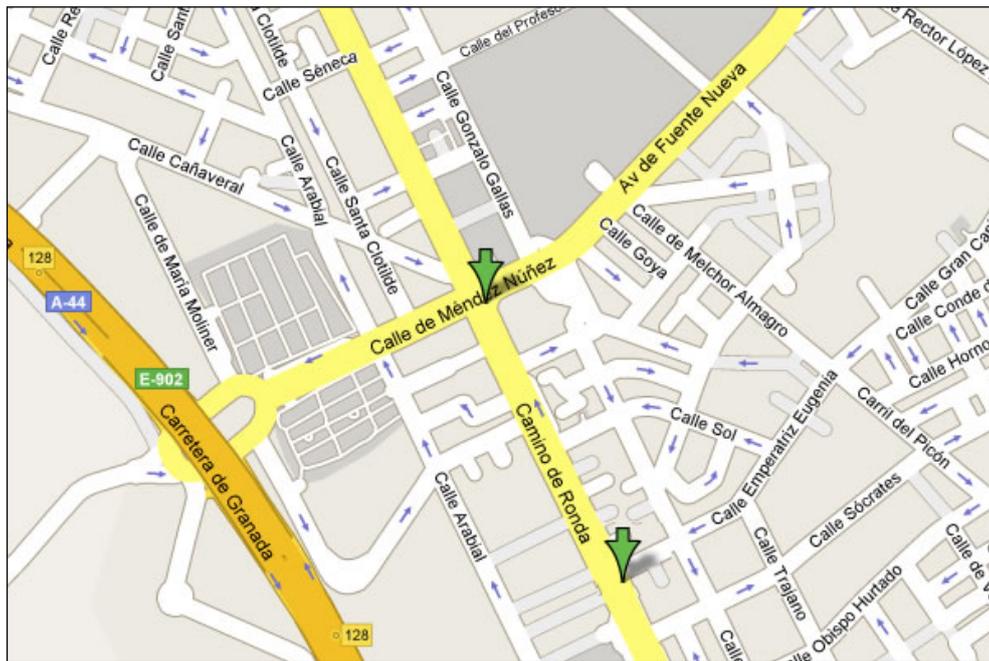


Figura 49. Camino de Ronda

Gran Vía: Vía en “U” del centro de la ciudad de Granada, que también sufre de denso tráfico de transporte público y motocicletas.

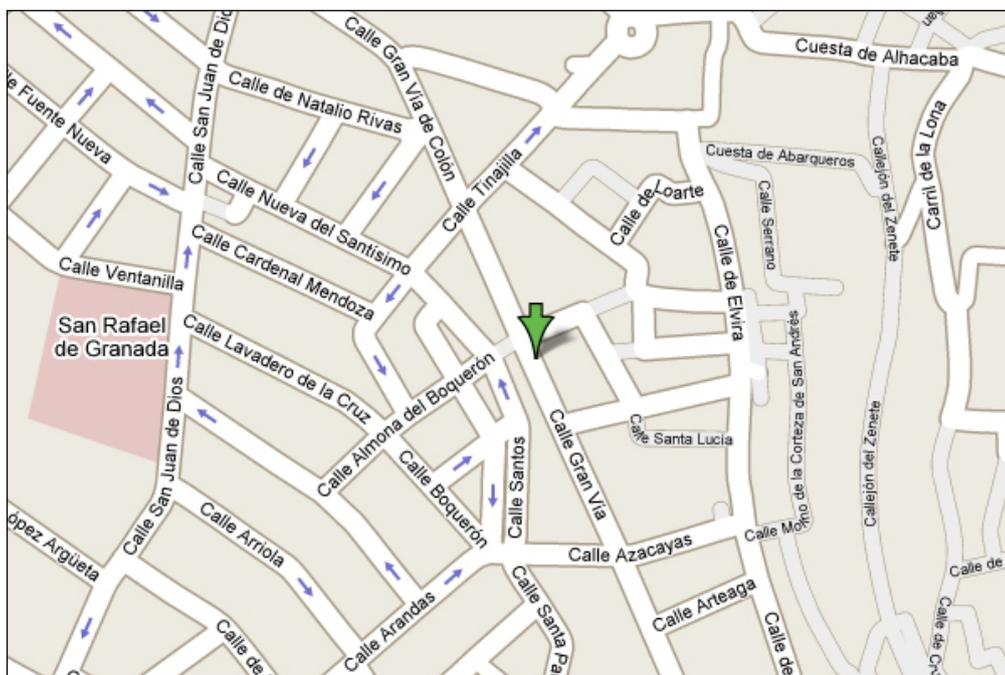


Figura 50. Gran Vía

Méndez Núñez: Vía en campo libre, ya está libre de edificios en los lugares en los que se realizaron las dos tomas. La primera toma se realizó junto a Plaza Einstein y la segunda en el cruce con calle Arabial.

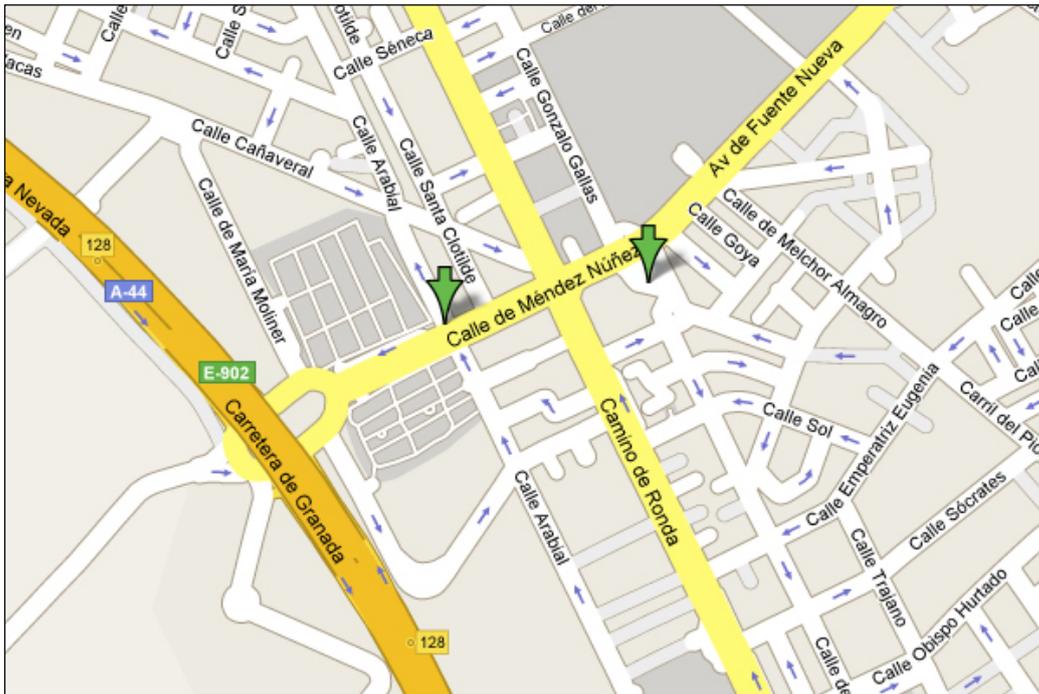


Figura 53. Méndez Núñez

Reyes Católicos: Vía del centro de la ciudad, típicamente en "U", que, al igual que Gran Vía, soporta tráfico de transporte público y motocicletas durante el día.

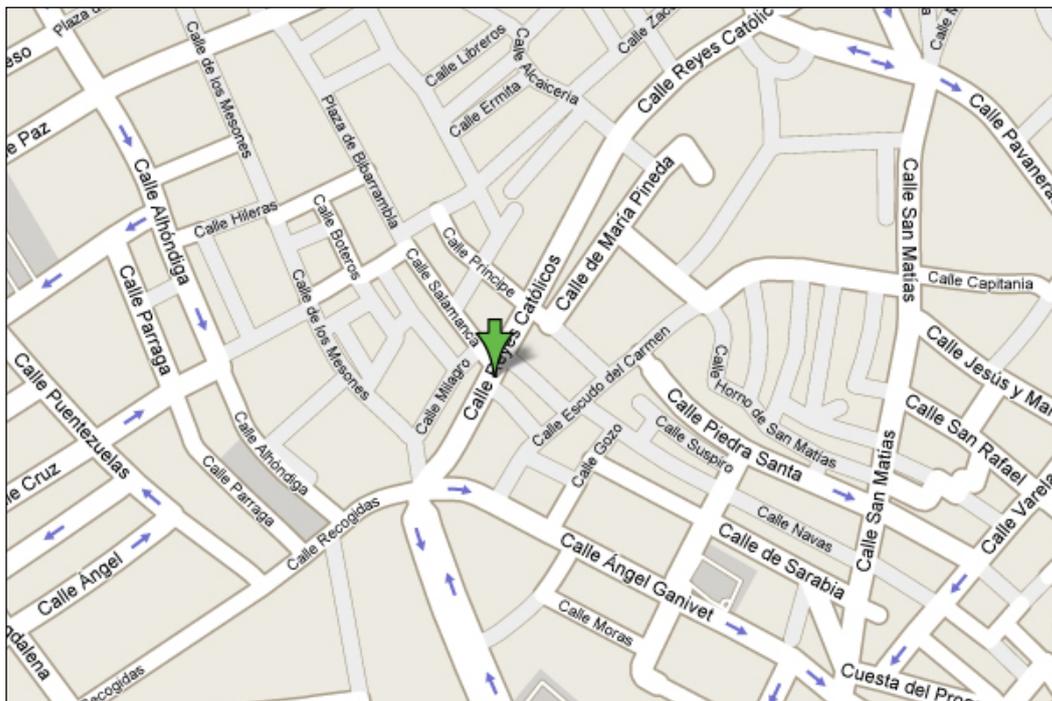


Figura 54. Reyes Católicos

Camino de las Vacas: Calle en campo libre que soporta cierto tráfico continuo fluido.

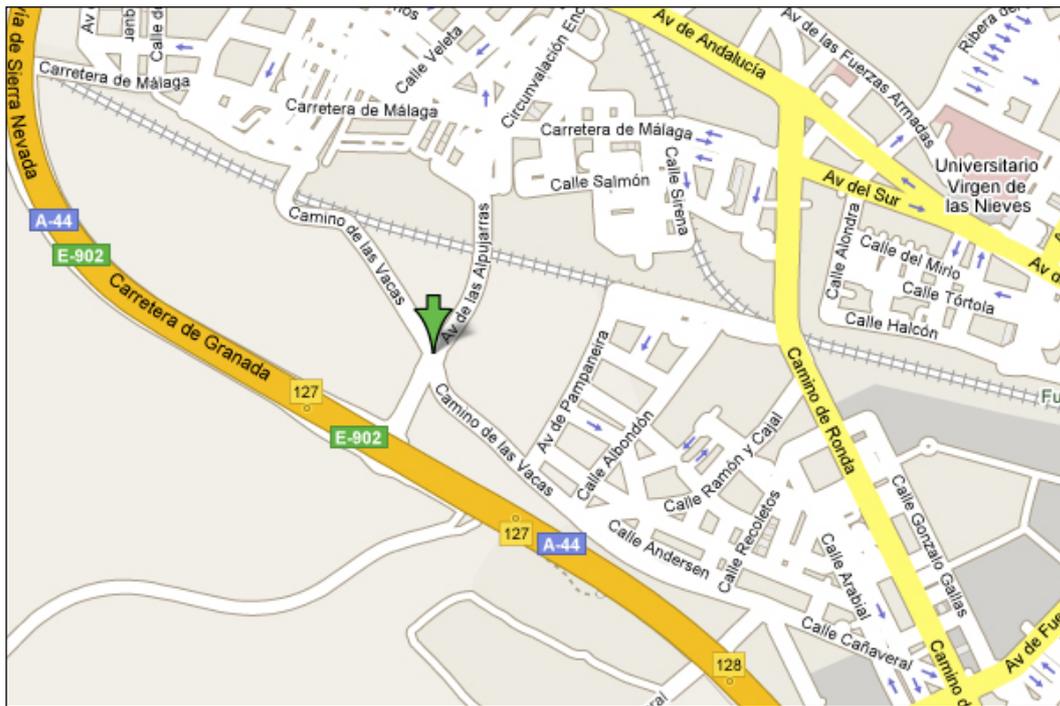


Figura 55. Camino de las Vacas

Avda. Doctor Olóriz: Avenida de varios carriles en un sólo sentido localizada en zona de hospitales de la ciudad. Se trata de una vía en “L”, en algunas zonas y en “U”, en otras.

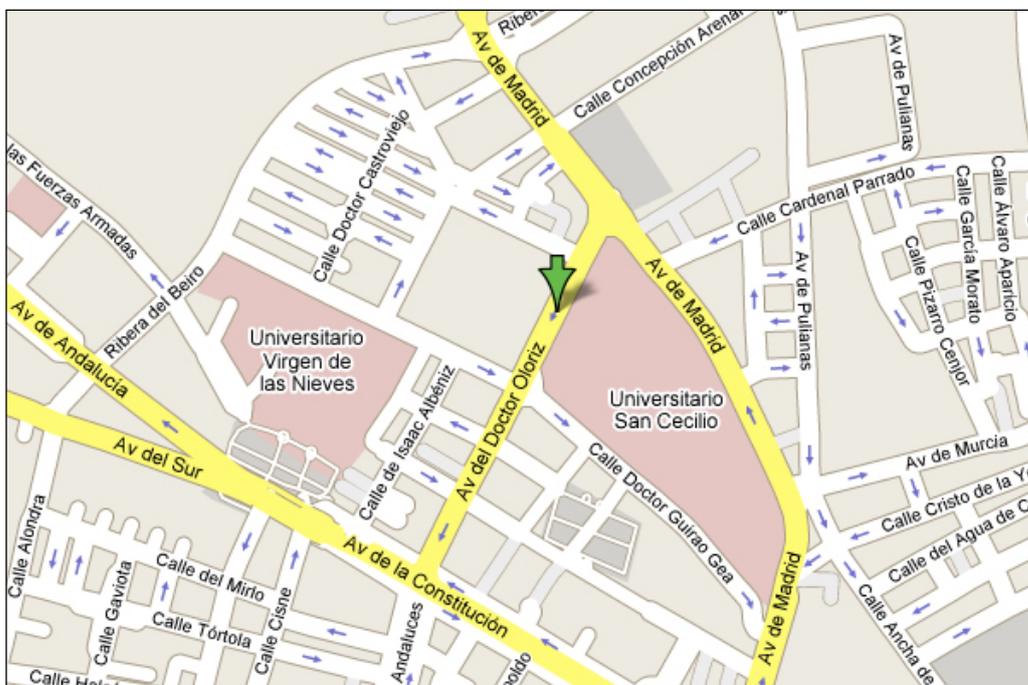


Figura 56. Doctor Olóriz

DATOS

Con la recogida de datos se obtiene un conjunto formado por 289 registros de datos, cada uno representado por 26 valores de las 26 variables consideradas, las 25 de la Tabla 28, que son variables de entrada al sistema y el L_{Aeq} medida por el sonómetro, como variable de salida.

3.2 RED NEURONAL ARTIFICIAL PARA LA PREDICCIÓN DE NIVELES SONOROS

Por su facilidad para identificar sistemas a partir de ejemplos, con aprendizaje supervisado y su capacidad de generalizar, se va a usar una red neuronal de tipo Perceptron Multicapa, con el algoritmo de aprendizaje Backpropagation. Hay diferentes variantes, entre las que se encuentra Levenberg-Marquardt, o el gradiente descendente con tasa de aprendizaje variable. La red neuronal ha sido construida con MATLAB (versión 7.11, tool neural networks), que dispone de múltiples funciones para el tratamiento de datos con redes neuronales.

Una de las 26 características de los registros disponibles, el nivel equivalente de ruido L_{Aeq} , nuestra variable objetivo, será la salida de la red, mientras que las otras 25 características se tomarán como factores influyentes en el nivel de ruido dado, es decir, como entradas para la red.

Además, la red ha de recibir los datos de entrada en un formato numérico adecuado; por ejemplo, variables lingüísticas como “presencia de eventos anómalos”, cuyos posibles valores son “SI” o “NO”, han de ir expresadas con valores numéricos.

Por otro lado, los rangos de las variables han de ser homogéneos, por lo que se hace conveniente normalizar los datos, es decir acotarlos en un rango de valores apropiado.

En la Tabla 28 se muestran las variables de entrada que se han tenido en cuenta y su correspondiente normalización, donde las variables 4, 6-12, 15-18 y 22-25 se transforman linealmente al rango $[0, 1]$. Las demás variables toman también valores en $[0, 1]$.

A su vez, era necesario normalizar la salida. Dado que el rango auditivo de una persona oscila aproximadamente entre los 20dB y los 120dB, se tomarán éstos como valores máximo y mínimo más adecuados con los que normalizar la salida al intervalo $[0, 1]$.

Tabla 28. Variables de entrada consideradas por la red y su normalización

	Variable	Significado	Rango de valores	Normalización
1	periodo	Período del día	Evening Day	0 1
2	entorno	Entorno Comercial o de Ocio	No Sí	0 1
3	obras	Presencia de Obras en la zona	No Sí	0 1
4	estabilización	Tiempo estabilización	[4.5, 52.8]	[0, 70]
5	flujo	Tipo de flujo	Flujo continuo en pulsos Flujo decelerado en pulsos Flujo acelerado en pulsos Flujo continuo fluido	0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8
6	caudalVLA	Caudal Vehículo Ligero Ascendente	[0, 34.8]	[0, 80]
7	caudalVLD	Caudal Vehíc. Ligeros Descendente	[0, 62.4]	[0, 80]
8	caudalMTA	Caudal Motocicletas Ascendente	[0, 16]	[0, 40]
9	caudalMTD	Caudal Motocicletas Descendente	[0, 14]	[0, 40]
10	caudalVPA	Caudal Vehículo Pesado Ascendente	[0, 8]	[0, 20]
11	caudalVPD	Caudal Vehic. Pesados Descendente	[0, 4.4]	[0, 20]
12	sirena	Número de vehículos con sirena	[0, 2]	[0, 10]
13	anomalosT	Aparición de eventos anómalos relacionados con el tráfico	No Sí	0 1
14	anomalos	Aparición de eventos anómalos no relacionados con el tráfico	No Sí	0 1
15	velocidad	Velocidad media de los vehículos	[0, 50]	[0, 80]
16	pendiente	Pendiente de la calzada	[0, 8.6]	[0, 15]
17	carrilesA	Número de carriles ascendente	[0, 2]	[0, 10]
18	carrilesD	Número de carriles descendente	[0, 4]	[0, 10]
19	pavimento	Tipo de pavimento	Asfalto liso Asfalto poroso Empedrado No tipo	0.2 0.4 0.6 0.8
20	firme	Estado del firme	Bueno Regular Malo Muy malo	0.2 0.4 0.6 0.8
21	vía	Tipo de vía	En U En L En J Campo libre	0.2 0.4 0.6 0.8
22	anchuraV	Anchura de la vía	[0, 32.2]	[0, 60]
23	altura	Altura media de edificios de la vía	[0, 31]	[0, 80]
24	anchuraC	Anchura de la calzada	[4.69, 22.53]	[0, 40]
25	distancia	Distancia de la fuente al receptor	[3.18, 16.1]	[0, 20]

PRUEBAS INICIALES

En primer lugar se consideraron las distintas opciones que MATLAB ofrece entre las variantes del algoritmo de retropropagación de errores. Para el método de aprendizaje, se consideró LEARNGD (gradiente descendente) y como función de entrenamiento, TRAINLM (Levenberg-Marquardt) y TRAIINGDA (gradiente descendente con tasa de aprendizaje variable).

Por otro lado se consideró como primera versión una Red Neuronal con una sola capa oculta. Inicialmente se han considerado diferentes números de neuronas ocultas, seleccionando 5 y 7 neuronas para el tratamiento de los datos.

Con estas elecciones, obtenidas con algunos experimentos previos, se procedió a hacer una prueba que permitiera decidir qué estructura de red usar. Para ello se crearon 4 redes, siguiendo las combinaciones:

TRAIINGDA con 5 neuronas en la capa intermedia

TRAIINGDA con 7 neuronas en la capa intermedia

TRAINLM con 5 neuronas en la capa intermedia

TRAINLM con 7 neuronas en la capa intermedia

De los 289 ejemplos disponibles, se utilizaron 200 para el entrenamiento y 89 para el test.

Para evitar resultados que pudieran ser considerados de buena/mala suerte, se hicieron 3 ejecuciones de cada una de las 4 redes, obteniendo así 12 ejecuciones en total a examinar. En la Tabla 29 se pueden ver los resultados de las ejecuciones.

Como se puede observar, el tipo de función de entrenamiento TRAINLM necesita menos épocas, y además ofrece mejores resultados.

Aunque en el caso de la red 2, ejecución 1, el error máximo para el conjunto de test sea algo alto comparado con el resto de los resultados, se ha querido estudiar el siguiente valor más alto, el cual se sitúa por debajo de 4. Lo mismo ocurre en la ejecución 3 de la red 4. En el caso del error que asciende a 4.9 dB, el siguiente error se encuentra, siguiendo la línea de las demás ejecuciones, en 3.04 dB. A pesar de esto, respecto a esta característica, es conveniente usar el algoritmo TRAINLM, dado que no es representativo que en uno de los 289 registros se produzcan estos fallos.

Por otro lado, y observando las medias de los errores, en las redes con función de entrenamiento TRAINLM (las redes 2 y 4), se ha obtenido una media del error inferior a las redes 1 y 3. Entre las redes 2 y 4, las medias son muy similares. Al calcular la media de las medias de cada ejecución por red, se obtiene un comportamiento algo mejor en la red 4, con 0.48 dB en el conjunto de entrenamiento frente a 0.52 dB.

	<i>Red 1</i>			<i>Red 2</i>			<i>Red 3</i>			<i>Red 4</i>		
Conjunto entrenamiento	200 registros											
Conjunto test	89 registros											
Número de entradas	25											
Número de salidas	1											
Neuronas capa intermedia	5			5			7			7		
Función entrenamiento	TRAINGDA			TRAINLM			TRAINGDA			TRAINLM		
Función aprendizaje	LEARNGD			LEARNGD			LEARNGD			LEARNGD		
Tasa de aprendizaje	Variable											
Ejecución	1	2	3	1	2	3	1	2	3	1	2	3
Épocas	100000	100000	100000	70	20	20	100000	100000	100000	21	25	10
Objetivo	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}
Performance conseguido	12×10^{-5}	13×10^{-5}	9×10^{-5}	2.8×10^{-5}	5×10^{-5}	5×10^{-5}	9.8×10^{-5}	9.3×10^{-5}	9.5×10^{-5}	3.7×10^{-5}	2.8×10^{-5}	5.8×10^{-5}
Error máximo conjunto entrenamiento	4.1 dB	6.9 dB	3.3 dB	2.7 dB	2.8 dB	2.7 dB	3 dB	3.5 dB	4.04 dB	2.7 dB	2.9 dB	2.9 dB
Error máximo conjunto test	4.1 dB	6.7 dB	3.9 dB	7.1 dB	2.9 dB	3.1 dB	3.9 dB	2.8 dB	3.6 dB	3.1 dB	4.3 dB	4.9 dB
Media conjunto entrenamiento	0.82 dB	0.8 dB	0.76 dB	0.55 dB	0.48 dB	0.53 dB	0.77 dB	0.72 dB	0.77 dB	0.45 dB	0.45 dB	0.55 dB
Media conjunto test	1.03 dB	1.05 dB	0.99 dB	0.75 dB	0.80 dB	0.74 dB	0.91 dB	1.06 dB	0.91 dB	0.72 dB	0.74 dB	0.81 dB
Media de las medias	0.79 dB			0.52 dB			0.76 dB			0.48 dB		
	1.02 dB			0.76 dB			0.96 dB			0.76 dB		
Desviación típica conjunto entrenamiento	0.73	0.81	0.61	0.48	0.45	0.45	0.61	0.63	0.67	0.41	0.40	0.49
Desviación típica conjunto test	0.86	1.15	0.92	0.67	0.69	0.70	0.79	0.75	0.77	0.67	0.74	0.81

Tabla 29. Resultados de las 12 ejecuciones para elegir una estructura de red definitiva

Además se observa, en los casos de las redes 2 y 4, una desviación típica baja, lo cual es interesante dado que la media del error es muy baja.

En las Figuras 61 y 62 se muestran las gráficas del nivel de ruido equivalente esperado junto con el calculado por la red, en ambos casos, de entrenamiento y de test. En ellas es directamente observable el buen resultado de este tipo de red con nuestro problema. La línea interior (más clara) indica $\pm 5\%$ respecto de 60dB (± 3 dB) y la exterior $\pm 5\%$ respecto de 80 dB (± 4 dB). Esto se ha planteado así debido a que prácticamente todas las medidas están entre 60 y 80 dB.

Todo lo que se acaba de comentar es decisivo para elegir definitivamente una estructura de red de tipo perceptrón multicapa (backpropagation), con función de entrenamiento TRAINLM, función de aprendizaje LEARNGD y 7 neuronas en la capa oculta.

Además, posteriormente se hicieron pruebas con 9 neuronas cuyos resultados, similares a los aquí expuestos, se presentaron en un congreso internacional [153]. En MATLAB se representaría según la Figura 63.

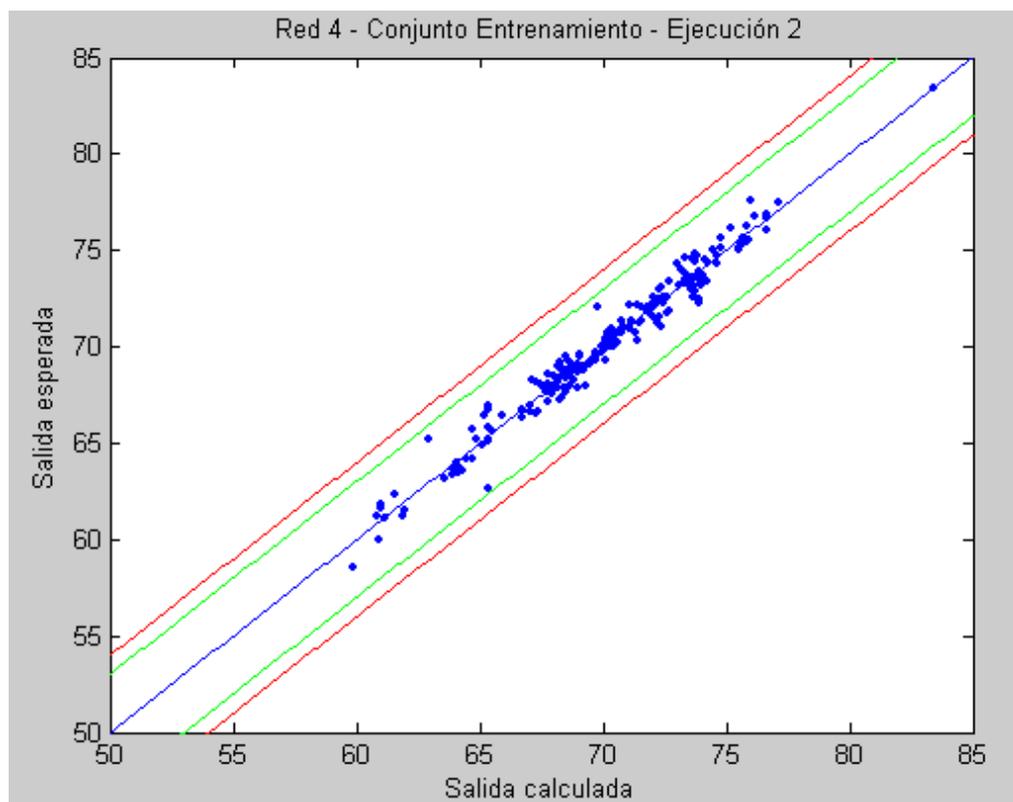


Figura 61. Gráfica comparativa del conjunto entrenamiento, red 4, Ejecución 2

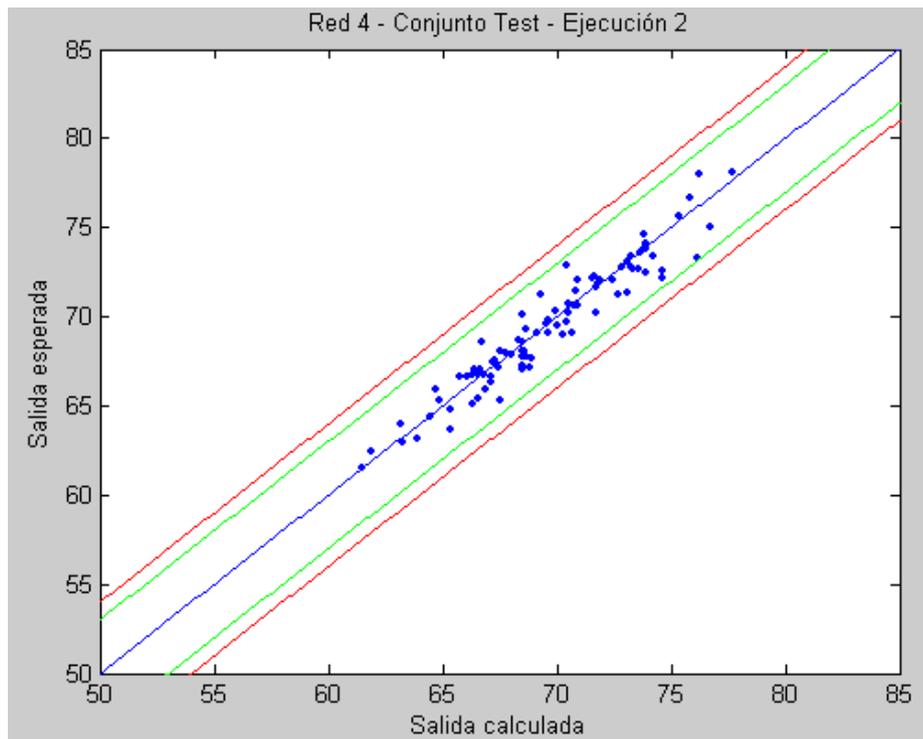


Figura 62. Gráfica comparativa del conjunto test, red 4, Ejecución 2

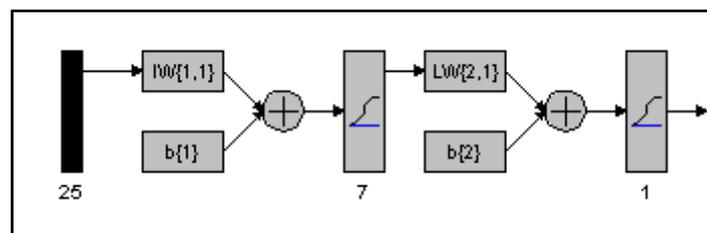


Figura 63. Estructura de la red neuronal propuesta

3.2 RESULTADOS DE LA RED NEURONAL PROPUESTA

Una vez elegida la estructura de la red neuronal que se va a utilizar, se construyen 5 grupos de datos diferentes para que los resultados puedan considerarse válidos, y no dependientes del azar. Los 5 grupos de datos se construyen aleatoriamente a partir de los 289 registros de entrada de los que disponemos. Así, se obtienen cinco grupos de datos con un conjunto de entrenamiento y un conjunto de test cada uno, formados éstos a su vez por registros diferentes. Cada conjunto de entrenamiento contiene 200 registros y el de test los otros 89 registros.

La estructura de la red neuronal fue ejecutada cinco veces con cada uno de los cinco conjuntos de datos obtenidos. Es decir, que se obtuvieron aleatoriamente cinco conjuntos de pesos iniciales diferentes, y la red se entrenó una vez para uno de estos conjuntos de pesos y para cada conjunto de entrenamiento. Esto da lugar a 25 pruebas y 25 resultados para el nivel de ruido. El comportamiento de la red se ha evaluado con el error respecto del resultado esperado, así como las medidas estadísticas de la media y la desviación típica del error.

Las Tablas 30 - 34 contienen los resultados para los cinco conjuntos de datos utilizados. Cada una se presenta con las 5 ejecuciones de la red, y en ella se muestran los parámetros de la red para cada ejecución, el número de épocas en el entrenamiento y los resultados. También se presentan los resultados relativos a la salida media de las 5 ejecuciones. Esto es de interés teniendo en cuenta la teoría de errores de las medidas físicas, debido al error que los aparatos de medida pueden cometer. La salida de la RN obtenida en cada ejecución, para cada ejemplo, se considera como una medida del L_{Aeq} y entonces se considera la media de los 5 valores como una medida más aproximada del valor real de L_{Aeq} .

Para analizar los resultados se tienen en cuenta los ejemplos que sobrepasan el 5% del error, que es una medida menor a lo que los expertos consideran válido en este ámbito. En el texto, se les llamará ejemplos erróneos. Asimismo se miden la media y la desviación típica del error. También se considera de interés indicar cuál es el error máximo que se ha cometido en el conjunto de test, observando que en sólo una ocasión ha sido superior a 9dB.

Como se deduce de las Tablas 30-34 , en todas las ejecuciones, el número de épocas ha sido muy bajo, especialmente en los conjuntos de muestras 2 y 5. Esto significa que el entrenamiento es muy rápido, lo que permitirá fácilmente la utilización de las RN entrenadas en tiempo prácticamente real.

En todos los conjuntos, la mayoría de las ejecuciones aprenden correctamente, y en el conjunto de test, como mucho 2 (3 en el conjunto 5) ejemplos superan el 5% del error respecto de la salida esperada. Solo en una ejecución (primera del conjunto 3) este número de ejemplos ha llegado a 5, debido, seguramente, a una mala configuración de pesos iniciales aleatorios.

Excepto en los conjuntos 1 y 5, siempre hay ejecuciones en que incluso el conjunto de test completo está por debajo del 5% en el error de salida. Solo un caso (1 ejemplo de entrenamiento o test en una ejecución, la 4 del conjunto 4) supera el error del 9 dB.

Tabla 30. Resultados de las cinco ejecuciones de la red con el conjunto de datos 1

<i>CONJUNTO 1</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	20	10	12	9	40	
Ejemplos erróneos (Entre/Test)	0/4	0/3	0/1	0/1	0/2	0/0
Error máximo conjunto test (dB)	6,36	4,61	3,68	3,07	5,18	2,69
Media del error (dB)	0,71	0,634	0,57	0,64	0,57	0,62
Desviación típica del error (dB)	0,72	0,59	0,56	0,56	0,65	0,51

Tabla 31. Resultados de las cinco ejecuciones de la red con el conjunto de datos 2

<i>CONJUNTO 2</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	15	14	13	12	11	
Ejemplos erróneos (Entre/Test)	0/3	0/4	0/3	0/0	0/1	0/0
Error máximo conjunto test (dB)	4,46	5,15	4,01	3,21	2,32	2,04
Media del error (dB)	0,78	0,64	0,70	0,62	0,60	0,64
Desviación típica del error (dB)	0,67	0,72	0,69	0,71	0,65	0,78

Tabla 32. Resultados de las cinco ejecuciones de la red con el conjunto de datos 3

<i>CONJUNTO 3</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	35	8	16	8	16	
Ejemplos erróneos (Entre/Test)	1/5	0/0	0/1	0/0	0/2	0/0
Error máximo conjunto test (dB)	5,27	3,20	4,08	2,72	3,92	3,10
Media del error (dB)	0,79	0,68	0,59	0,69	0,59	0,67
Desviación típica del error (dB)	0,81	0,57	0,56	0,55	0,59	0,50

Tabla 33. Resultados de las cinco ejecuciones de la red con el conjunto de datos 4

<i>CONJUNTO 4</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	27	16	20	14	27	
Ejemplos erróneos (Entre/Test)	0/2	0/0	0/2	0/2	0/2	0/1
Error máximo conjunto test (dB)	6,98	2,64	7,98	14,47	4,51	5,85
Media del error (dB)	0,67	0,57	0,58	0,5	0,59	0,61
Desviación típica del error (dB)	0,72	0,51	0,69	1,05	0,61	0,60

Tabla 34. Resultados de las cinco ejecuciones de la red con el conjunto de datos 5

<i>CONJUNTO 5</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	13	16	14	9	8	
Ejemplos erróneos (Entre/Test)	0/2	0/3	0/1	0/3	0/3	0/0
Error máximo conjunto test (dB)	4,49	4,47	3,98	5,08	4,58	3,57
Media del error (dB)	0,73	0,62	0,65	0,74	0,78	0,71
Desviación típica del error (dB)	0,66	0,65	0,59	0,71	0,70	0,63

Por otro lado, la media del número de ejemplos erróneos no llega a 2 (la media para las 25 ejecuciones es 1.92).

El error medio se mantiene siempre muy bajo (por debajo de 0.79 dB), lo que indica un aprendizaje global muy bueno.

La desviación típica se mantuvo constante entre 0.55 y 0.72 en la mayoría de los ejemplos, excepto en dos casos. Por un lado, en el caso del conjunto 4, ejecución 4, que es de 1.05, debido a un ejemplo en el que el error es de 14.47 dB. Y por otro, en el caso del conjunto 3, ejecución 1, que es de 0.81, puesto que el error máximo es también algo elevado (5.27 dB) y además el error medio asciende a 0.79, una cifra algo alta respecto al resto de ejecuciones. Este grupo de datos tiene además la particularidad de contener el único conjunto de entrenamiento con un error por encima del 5% del valor esperado. Esto confirma que el aprendizaje global obtenido por la RN es muy bueno.

Especial consideración tienen los resultados de salida medios (columna 6 de las tablas). El uso de estos valores se justifica claramente en la teoría de errores de las medidas físicas.

Como se puede observar, en todos los conjuntos de datos, el aprendizaje es correcto, y en los ejemplos de test, solo para el conjunto 4, un ejemplo tiene un error superior al 5%. El error medio es siempre menor que 0.71 y la desviación típica no supera el valor 0.83, lo que indica un aprendizaje global excelente.

El hecho de que el entrenamiento de las RN propuestas es muy rápido, y que pueden ser entrenadas en paralelo, apoya la utilización de valores de las salidas de varias RN, como salida calculada por el sistema que se propone.

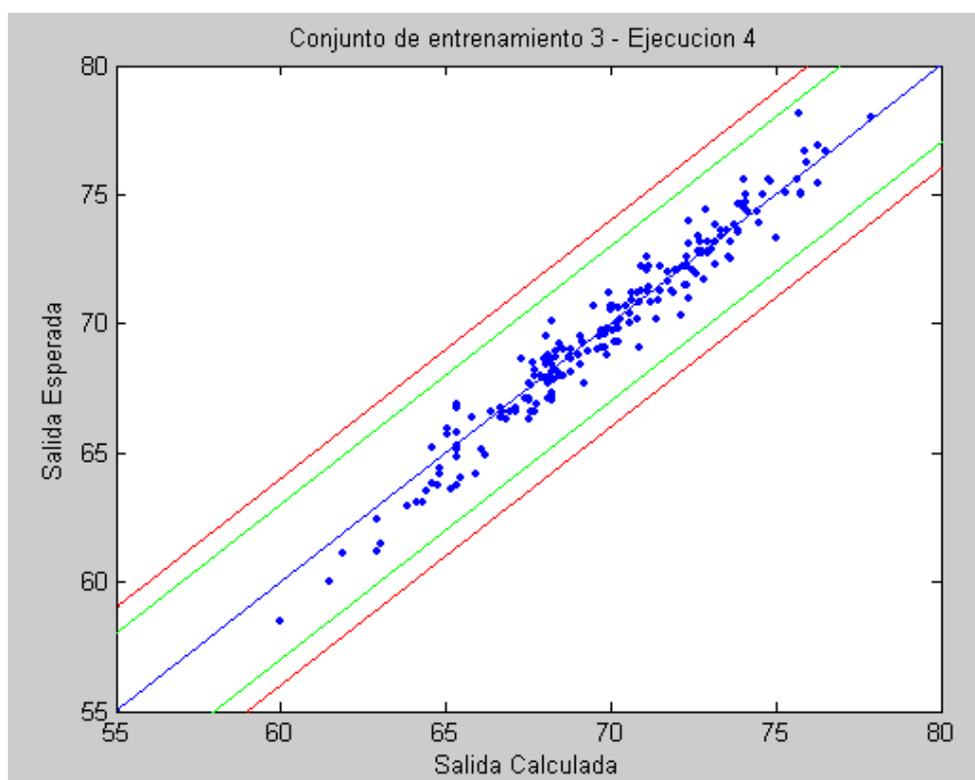


Figura 64. Gráfica comparativa del conjunto entrenamiento 3, ejecución 4

A la vista de estos resultados, en los que se obtiene una media del error de las 25 ejecuciones muy baja: 0.716 dB, y teniendo en cuenta que los expertos en la materia admiten un error superior al 5%, éstos se pueden considerar unos resultados excelentes, ya que implica reducir el error objetivo en un 92%. La media de las desviaciones de las 25 ejecuciones fue de 0.786, siendo un valor también muy bueno, dado que la media del error es baja.

En las Figuras 64 y 65 se muestran las gráficas comparativas del nivel de ruido medido con el obtenido, para los conjuntos de entrenamiento y de test. Se ilustran los resultados para el grupo de datos 3. La línea interior (más clara) indica $\pm 5\%$ respecto de 60dB (± 3 dB) y la exterior $\pm 5\%$ respecto de 80 dB (± 4 dB). Se observa que ningún dato se sale de estos límites, en ninguno de los dos casos.

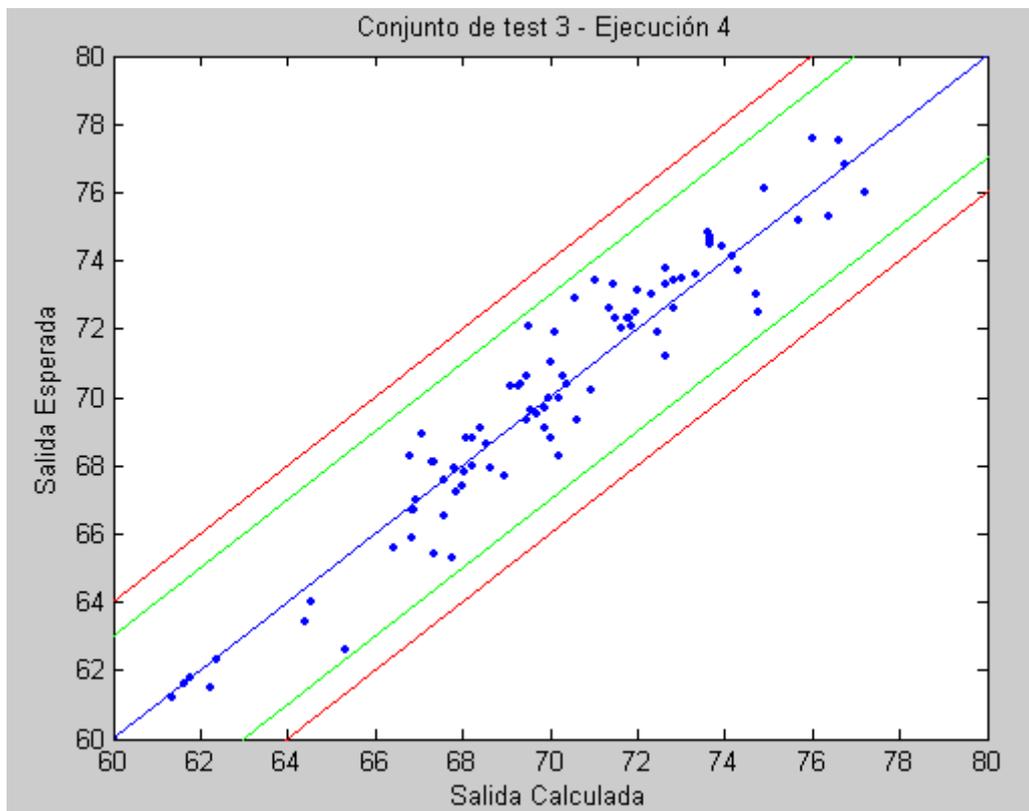


Figura 65. Gráfica comparativa del conjunto test 3, ejecución 4

A modo de ejemplo, en la Tabla 35 se muestran los resultados completos del conjunto 3, ejemplos de test.

Tabla 35. Resultados de la red del conjunto de test 3

	Datos	L _{Aeq} esperado	L _{Aeq} obtenido (dBA) Ejecución					L _{Aeq} media(dBA)	Error de L _{Aeq} media (dBA)
			1	2	3	4	5		
DATOS TEST	5	73	72,859	72,249	72,351	72,317	71,913	72,3378	0,6622
	7	72,3	71,674	72,18	71,225	71,813	72,29	71,8364	0,4636
	12	72,5	72,196	72,499	72,194	71,936	70,365	71,838	0,662

13	73,1	72,544	72,615	72,171	71,993	70,703	72,0052	1,0948
14	69,7	69,879	69,603	69,82	69,891	69,799	69,7984	0,0984
15	72,3	72,315	71,499	71,463	71,477	71,034	71,5576	0,7424
20	70	70,085	70,974	70,187	70,226	70,612	70,4168	0,4168
22	72,1	67,043	70,032	69,366	69,516	69,02	68,9954	3,1046
28	74,4	74,424	74,058	74,118	73,919	73,208	73,9454	0,4546
31	67,7	62,433	69,837	69,228	68,985	67,732	67,643	0,057
33	72,9	70,196	70,86	70,665	70,587	71,41	70,7436	2,1564
35	70,4	69,496	70,744	70,235	70,401	69,818	70,1388	0,2612
39	74,8	73,101	74,426	73,722	73,609	73,616	73,6948	1,1052
42	73,4	72,604	72,551	72,649	72,857	72,789	72,69	0,71
43	76,1	73,952	74,738	74,197	74,887	74,707	74,4962	1,6038
44	73,7	74,149	73,668	73,482	74,309	75,007	74,123	0,423
52	71,9	72,38	72,692	72,184	72,462	71,811	72,3058	0,4058
58	72,1	71,944	71,175	70,874	71,856	73,046	71,779	0,321
61	71,2	72,274	71,151	72,406	72,64	75,128	72,7198	1,5198
65	73,5	74,763	73,119	73,226	73,025	72,548	73,3362	0,1638
72	73,8	72,609	73,094	73,185	72,657	72,189	72,7468	1,0532
73	72,3	72,488	71,749	71,832	71,755	72,27	72,0188	0,2812
75	76	72,697	76,653	76,678	77,2	77,153	76,0762	0,0762
77	75,2	72,656	74,654	75,319	75,67	76,648	74,9894	0,2106
80	77,6	72,645	74,711	76,224	76,036	76,422	75,2076	2,3924
82	72,6	72,397	70,6	71,406	71,369	71,025	71,3594	1,2406
83	71,9	72,352	70,036	70,299	70,095	70,557	70,6678	1,2322
84	75,3	72,698	75,968	75,782	76,37	76,57	75,4776	0,1776
85	72	72,472	71,575	71,689	71,656	72,208	71,92	0,08
86	76,8	72,723	76,742	76,109	76,76	76,924	75,8516	0,9484
89	70,6	72,2	69,183	69,85	70,287	70,507	70,4054	0,1946
92	72,5	72,641	74,224	75,008	74,753	75,099	74,345	1,845
94	73	72,604	73,701	75,735	74,725	74,295	74,212	1,212
99	69,3	70,649	68,871	68,365	70,609	69,157	69,5302	0,2302
100	70	70,53	69,623	69,889	69,993	70,341	70,0752	0,0752
105	73,3	71,098	70,877	71,185	71,474	72,041	71,335	1,965
108	68,8	70,402	68,877	70,432	70,006	69,818	69,907	1,107
109	68	68,604	68,03	68,908	68,246	67,689	68,2954	0,2954
110	69,1	68,931	67,923	68,857	68,428	68,2	68,4678	0,6322
111	70,2	70,973	69,97	70,235	70,932	70,419	70,5058	0,3058
113	68,3	70,614	69,427	70,874	70,219	69,947	70,2162	1,9162
114	70,3	70,195	69,49	70,042	69,268	70,278	69,8546	0,4454
119	70,4	70,577	69,72	70,483	69,318	70,269	70,0734	0,3266
129	66,5	67,09	67,556	67,657	67,564	67,496	67,4726	0,9726
130	67,6	67,09	67,556	67,657	67,564	67,496	67,4726	0,1274
133	68,8	68,329	68,198	68,102	68,224	68,19	68,2086	0,5914
140	62,6	65,516	65,575	65,323	65,322	65,458	65,4388	2,8388
153	69,1	69,893	69,377	69,218	69,867	69,51	69,573	0,473
158	68,1	67,705	67,276	67,343	67,336	68,003	67,5326	0,5674
159	65,9	67,052	67,158	67,215	66,845	66,751	67,0042	1,1042
164	67,8	68,291	67,538	68,149	68,061	68,574	68,1226	0,3226

165	69,6	69,543	69,076	69,887	69,567	70,835	69,7816	0,1816
166	67,9	67,232	69,696	68,331	68,65	69,211	68,624	0,724
167	68,6	68,582	68,958	68,424	68,539	68,941	68,6888	0,0888
169	66,7	66,59	66,978	67,085	66,84	66,613	66,8212	0,1212
173	67,9	67,305	69,796	68,515	68,62	69,6	68,7672	0,8672
174	66,7	66,619	66,995	67,131	66,879	66,648	66,8544	0,1544
178	68,8	68,313	67,549	68,179	68,089	68,6	68,146	0,654
180	70,3	70,249	69,368	69,907	69,11	70,343	69,7954	0,5046
181	62,3	61,246	61,197	61,374	62,377	61,233	61,4854	0,8146
184	65,4	65,856	66,335	65,931	67,363	66,747	66,4464	1,0464
192	61,6	60,64	61,041	60,826	61,648	60,493	60,9296	0,6704
195	64	64,356	64,164	63,814	64,539	63,987	64,172	0,172
196	67,2	66,154	67,662	67,138	67,836	70,738	67,9056	0,7056
197	61,5	61,25	61,482	61,372	62,256	61,267	61,5254	0,0254
200	61,2	60,688	61,613	60,813	61,337	60,596	61,0094	0,1906
206	63,4	64,062	64,926	63,749	64,412	63,728	64,1754	0,7754
211	74,6	74,004	72,626	73,289	73,647	73,051	73,3234	1,2766
212	74,5	74,004	72,626	73,289	73,647	73,051	73,3234	1,1766
214	74,1	74,283	73,757	75,344	74,16	74,209	74,3506	0,2506
223	73,6	73,053	73,157	72,759	73,356	73,197	73,1044	0,4956
224	77,5	76,095	77,686	76,953	76,616	77,634	76,9968	0,5032
225	74,7	73,878	73,399	73,609	73,684	73,522	73,6184	1,0816
228	68,9	68,668	67,667	68,851	67,064	67,922	68,0344	0,8656
229	68,1	67,437	67,347	67,618	67,322	67,838	67,5124	0,5876
231	70,6	70,142	70,492	70,754	69,49	70,701	70,3158	0,2842
235	65,3	66,626	66,55	66,71	67,756	67,073	66,943	1,643
239	68,3	67,6	66,731	67,635	66,795	67,191	67,1904	1,1096
243	73,3	72,808	75,516	71,856	72,649	73,98	73,3618	0,0618
244	69,5	70,14	69,713	70,068	69,687	69,409	69,8034	0,3034
246	73,4	71,638	71,865	69,314	71,017	71,491	71,065	2,335
247	71	72,78	69,722	69,397	70,034	70,009	70,3884	0,6116
248	69,3	70,128	69,956	70,139	69,472	69,166	69,7722	0,4722
252	72,6	72,83	73,937	73,163	72,814	72,885	73,1258	0,5258
258	67	66,112	66,225	66,249	66,927	66,113	66,3252	0,6748
264	67,9	67,568	67,933	67,191	67,827	67,318	67,5674	0,3326
267	65,6	65,795	65,975	66,036	66,441	65,946	66,0386	0,4386
269	61,8	62,457	62,133	60,749	61,756	61,615	61,742	0,058
279	67,4	67,641	66,969	67,678	67,999	67,511	67,5596	0,1596
Máximo error		5,272	3,202	4,086	2,722	3,928	3,1046	
Media error		0,7936	0,6863	0,5988	0,6906	0,591	0,6721	
Media errores		0,6721						
Desviaciones		0,8186	0,5735	0,5684	0,5542	0,5934	0,5021	
Media desviaciones		0,6216						

El resto de los datos, para todos los conjuntos (1 a 5), en ejemplos de entrenamiento y test, se encuentran en los anexos. Igualmente se incluyen en el anexo los gráficos de comparación para entrenamiento y test, de las salidas medias (columna L_{Aeq} media) para todos los conjuntos seleccionados. Ver también [154].

3.4 ENTRENAMIENTO CON 9 NEURONAS OCULTAS

Como se ha indicado anteriormente, también se han realizado pruebas con una RN con 9 neuronas en la capa oculta, un solo conjunto de entrenamiento y cinco ejecuciones. Los resultados obtenidos fueron presentados en un congreso internacional [153], mostrados en la Tabla 36.

Tabla 36. Resultados de las cinco ejecuciones de la red neuronal con 9 neuronas en la capa intermedia

<i>CONJUNTO 1 (9 NEURONAS EN LA CAPA OCUILTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	14	14	12	12	10	
Ejemplos erróneos (Entre/Test)	0/2	0/4	0/3	0/0	0/0	0/0
Error máximo conjunto test (dB)	4,37	5,49	3,86	3,45	2,42	2,84
Media del error (dB)	0,76	0,62	0,68	0,66	0,62	0,66
Desviación típica del error	0,78	1,05	0,87	0,76	0,67	0,83

Como se observa en la Tabla 36, los resultados son similares a los que se obtuvieron con un menor número de neuronas en la capa intermedia.

Cabe destacar que el número de épocas es bajo, consiguiendo a su vez un bajo número de ejemplos erróneos, ninguno en el conjunto de entrenamiento. En general, el error máximo del conjunto de test es también bajo, aunque en una ocasión es algo superior al resto. Esto ocurre en la segunda ejecución, en la que también la desviación típica es mayor al resto de ejecuciones, aunque la media del error se mantiene baja.

Respecto a los resultados de salida media, se obtienen también valores bajos, destacando especialmente que no hay ningún ejemplo erróneo en los conjuntos de entrenamiento y test.

3.5 COMPARATIVA ENTRE LA RED NEURONAL Y LOS MODELOS PREDICTIVOS EXISTENTES

Como se vio en la introducción, el objetivo de esta red neuronal es mejorar los resultados de los modelos predictivos existentes. Estos modelos calculan el nivel de ruido a partir de las características del entorno con la ayuda de cálculos matemáticos. Dado que cada uno se ha desarrollado en un lugar diferente del mundo, y con las circunstancias particulares locales, no todos responden de forma adecuada universalmente. Además, dado el gran número de variables diferentes que pueden ser tomadas en consideración, se hace muy engorroso proceder con

cálculos matemáticos que envuelvan más de 6 ó 7 variables. Por esto, los modelos conocidos eliminan las variables que los expertos consideran menos influyentes dado el entorno concreto, lo que va en detrimento de la precisión de los modelos. Estos no son problemas para una red neuronal, ya que es capaz de manejar grandes cantidades de datos, y obtener unos resultados óptimos si se ha entrenado con un conjunto de datos suficientemente amplio y diverso.

Para poder comparar los resultados, se han implementado los diferentes modelos, y se han ejecutado con los datos usados en la RNA. La implementación está disponible en <http://wdb.ugr.es/local/nataliag>

Para comprobar que efectivamente la red neuronal es capaz de obtener un nivel de ruido más próximo al real, se compara, en una gráfica, cada uno de los modelos con el resultado que se obtiene con la red. Se han obtenido gráficas para cada grupo de datos, mostrando aquí únicamente la gráfica para el primer grupo de datos. Esta gráfica, que compara la salida medida del conjunto de test con la salida de la red neuronal y la de cada modelo, se muestra en las Figuras 66 y 67.

Según se observa, los modelos que más se aproximan a nivel de ruido real son el modelo lineal, el granadino I y el E. Gaja, aunque la red neuronal es la que en todo momento se ajusta más al nivel esperado. Uno de los puntos fuertes de la red se da cuando hay un ejemplo en el que el caudal de vehículos es cero (visto en los ejemplos de la calle mesones), momento en el que la red aproxima muy bien, mientras que los modelos predictivos dan un valor muy bajo. Esto se debe a que se basan primordialmente en el caudal de vehículos para el cálculo base, y, en el caso del caudal igual a cero, el nivel de ruido de salida es calculado sólo mediante las correcciones.

Como conclusión o resumen se puede indicar que, como es conocido, las características de las RNA hacen que éstas sean adecuadas para aplicaciones en las que no hay un patrón identificable que pueda ser programado, y en las cuales se dispone de un conjunto de datos de entrada clasificados o no. Esto hace a las redes neuronales la opción perfecta para abordar el problema de la predicción del ruido ambiental, ya que se dispone de un conjunto de datos de entrada y salida.

Se ha comprobado que, en efecto, la predicción realizada por el modelo de RN propuesto supera con creces la realizada por los modelos matemáticos que se usan habitualmente. Pero además, la predicción realizada es muy buena en el conjunto de datos utilizado, propuesto por expertos en ruido ambiental.

El número de variables de entrada es alto comparado con los modelos matemáticos, pero el tiempo de respuesta es suficientemente bajo para poder usar las RN en tiempo real.

Teniendo en cuenta lo anterior, hay que plantearse estudiar lo que pasa si el número de variables se reduce, y cuáles de ellas son las menos relevantes. Para ello será interesante hacer un Análisis de Componentes Principales y también un estudio y análisis de selección de características.

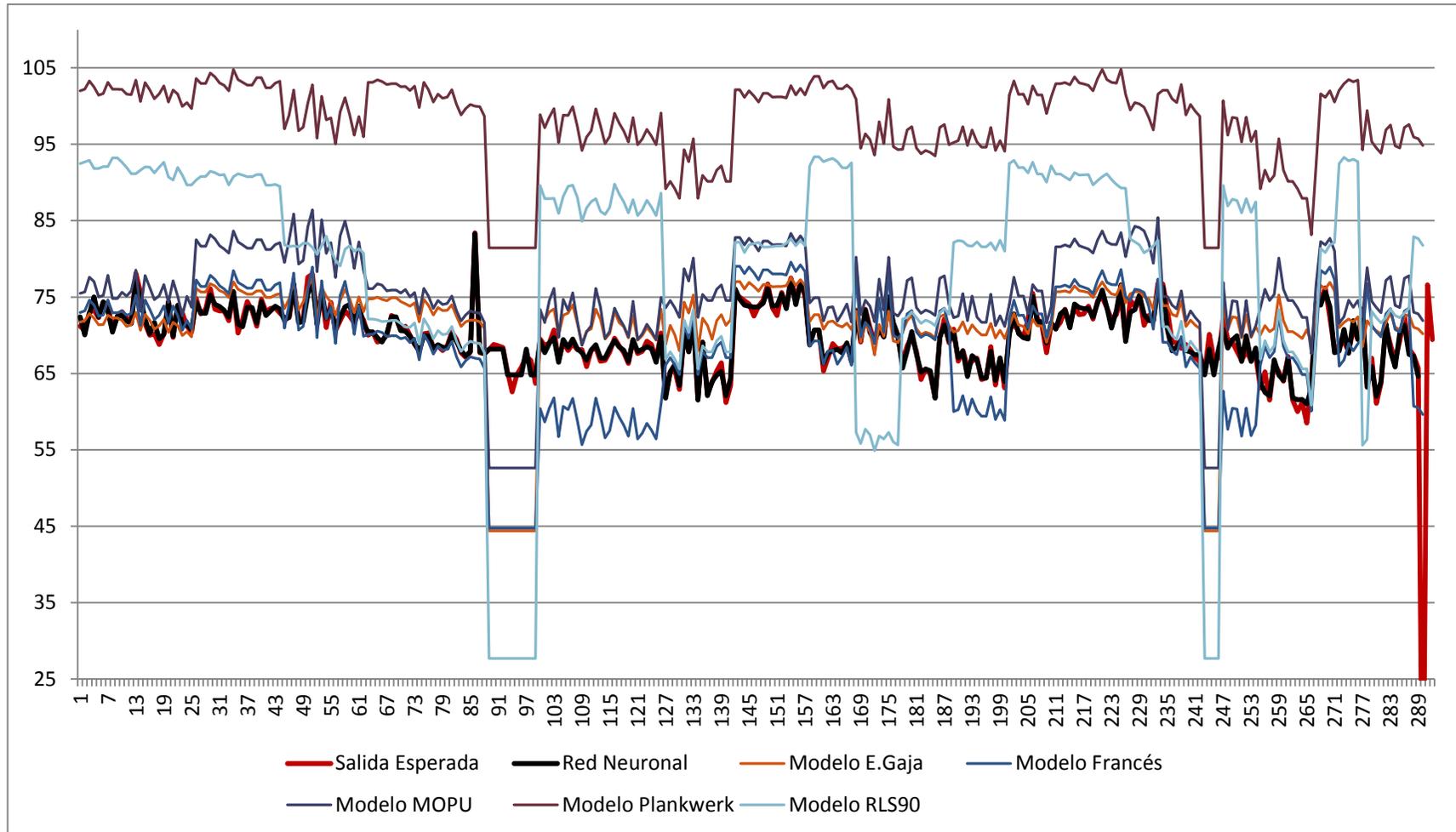


Figura 66. Gráfica comparativa de los resultados de los modelos clásicos y la red neuronal con el valor L_{Aeq} medido para el conjunto de datos 1

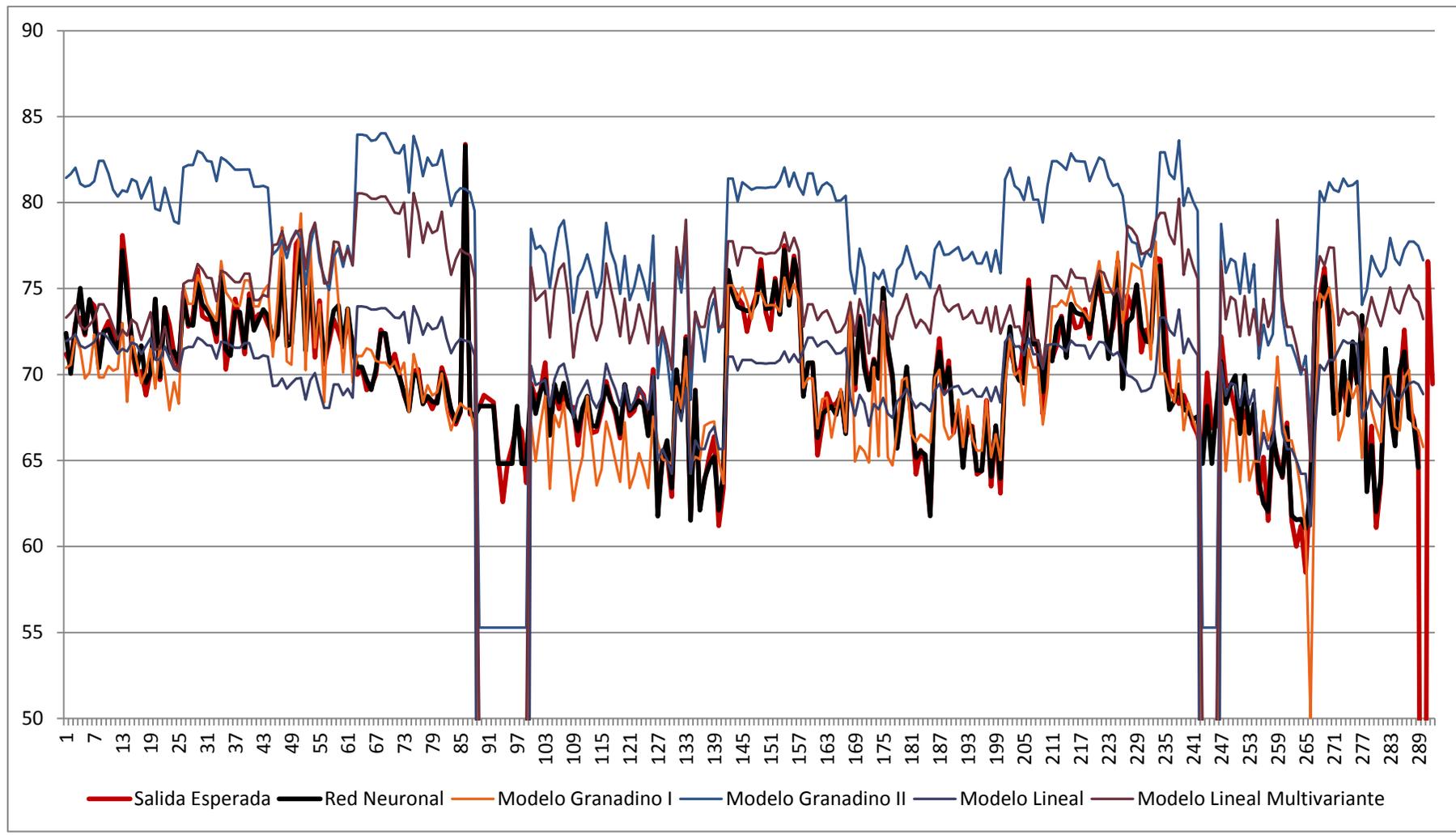


Figura 67. Gráfica comparativa de los resultados de los modelos clásicos y la red neuronal con el valor L_{Aeq} medido para el conjunto de datos 1

CAPÍTULO IV: SELECCIÓN Y EXTRACCIÓN DE CARACTERÍSTICAS

La red neuronal artificial entrenada para la predicción del nivel de presión sonora recibe como entrada registros de datos formados por 25 características. Como se ha indicado en el capítulo previo, las 25 variables se han obtenido haciendo la unión conjuntista de las variables usadas por los distintos métodos matemáticos descritos en el capítulo I, y adicionalmente se ha tenido en cuenta la opinión de algunos expertos para separar, en variables distintas, los valores descendentes o ascendentes en las vías públicas.

La forma de obtener el conjunto de las 25 variables permite reflexionar sobre si todas las variables consideradas son relevantes, o quizás alguna se pudiera eliminar, con el consiguiente ahorro de tiempo de cómputo.

Para comprobar, evaluar y analizar si es posible la reducción del número de variables que definen el paisaje sonoro en esta memoria, y seleccionar las variables no relevantes, se van a aplicar indistintamente métodos de reducción de características, tanto de extracción como de selección de características.

Como método de extracción de características se utilizará el Análisis de Componentes Principales (ACP), que permite reducir el número de entradas, obteniendo

combinaciones lineales de las características ya existentes, con lo que no necesariamente se identificarán en sí características concretas.

Este primer paso va a indicar, además del número de características, si alguna de ellas coincide con alguna de las 25 variables consideradas en la experimentación realizada, o todas son combinaciones lineales de algunas de estas 25 variables. En caso de que el ACP no extraiga variables del conjunto inicial propuesto, se usarán métodos de selección de características para ello. Se considerará el número de características extraídas como el mínimo número de variables directas que se pueden aceptar tras el proceso de selección.

Se evaluarán distintos métodos de selección de características, que seguramente proporcionarán distintos conjuntos de variables seleccionadas. Para obtener un subconjunto definitivo de las variables, se entrenarán, con el modelo de RNA considerada en el capítulo anterior, los diferentes conjuntos de variables obtenidos. Con los que hayan obtenido un nivel de aprendizaje suficiente, se hará un análisis de las variables incluidas, a fin de obtener un conjunto final (o quizás más de uno) con suficientes garantías de éxito.

En el capítulo III se comprobó, tras la batería de pruebas realizadas, que 7 neuronas en la capa oculta era una buena opción. Posteriormente se presentaron también, en un congreso internacional [153], experimentos con 9 neuronas en la capa oculta con resultados similares. Por ello, todos los experimentos de este capítulo se van a realizar con 7 y 9 neuronas en la capa intermedia.

Para todo ello, se van a utilizar diferentes métodos de los descritos en el capítulo I. Se pretende seleccionar, uno o más conjuntos de variables, con la etiqueta de “relevantes”, que se puedan considerar suficientes para entrenar una RNA para hacer una predicción del ruido ambiental urbano, suficientemente precisa.

4.1 EXTRACCIÓN DE CARACTERÍSTICAS: ANÁLISIS DE COMPONENTES PRINCIPALES

Como una primera aproximación a la reducción de características, se ha aplicado el Análisis de Componentes Principales[62]. Como se explica en el capítulo I, el Análisis de Componentes Principales (ACP) es un método estadístico que busca relaciones entre los datos para reducir el número de entradas (componentes del vector de entradas) en la medida de lo posible. MATLAB dispone de funciones para este preprocesamiento de las entradas. Concretamente la función *prepca* aplica el Análisis de Componentes Principales a los datos que recibe, devolviendo el conjunto de datos recibido reducido y una matriz de transformación para realizar el postprocesamiento en conjuntos de datos de las mismas características.

Tras aplicar el análisis al conjunto de 289 registros de 25 variables cada uno (una matriz de 25x289), se obtiene una nueva matriz de 11 filas y 289 columnas. Es decir, las 25 entradas han sido reducidas a 11. Estas entradas no se corresponden exactamente con ninguna de las primitivas, sino que son combinaciones lineales de las mismas. Esto indica que la reducción de entradas mediante una selección de características no es inmediata y

requiere un estudio más profundo que se abordará en las siguientes secciones de este capítulo.

De cualquier forma, es necesario evaluar el comportamiento del modelo de RNA considerando las 11 entradas obtenidas. Se hace un estudio similar al del capítulo III considerando los mismos conjuntos de entrenamiento y test, pero con sólo 11 entradas.

Aunque los resultados, como cabía esperar, no son tan espectaculares como los obtenidos en el capítulo III, se consideran de relevancia ya que el conjunto de datos se ha reducido bastante. Se muestran en las Tablas 37 - 41, los resultados de las 5 ejecuciones de la RN con los mismos 5 grupos de datos que el capítulo III.

Tabla 37. Resultados de las cinco ejecuciones de la red con el conjunto 1 (con ACP)

<i>CONJUNTO 1 (7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	13	14	14	14	15	
Ejemplos erróneos (Entre/Test)	0/5	1/3	0/3	2/4	0/5	0/3
Error máximo conjunto test (dB)	8,13	6,71	7,08	6,11	6,81	7,63
Media del error (dB)	0,88	0,91	0,82	0,99	0,83	0,89
Desviación típica del error (dB)	0,94	0,86	0,81	1,05	0,82	0,78

Tabla 38. Resultados de las cinco ejecuciones de la red con el conjunto 2 (con ACP)

<i>CONJUNTO 2 (7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	15	6	7	14	11	
Ejemplos erróneos (Entre/Test)	3/5	5/7	3/2	3/2	3/3	4/2
Error máximo conjunto test (dB)	8,78	5,91	7,00	4,64	6,83	8,37
Media del error (dB)	0,92	1,09	1	0,96	0,86	0,97
Desviación típica del error (dB)	1,01	1,11	1,00	0,96	0,92	0,86

Tabla 39. Resultados de las cinco ejecuciones de la red con el conjunto 3 (con ACP)

<i>CONJUNTO 3 (7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	11	22	7	17	21	
Ejemplos erróneos (Entre/Test)	4/1	2/3	4/1	4/1	5/0	3/0
Error máximo conjunto test (dB)	3,76	5,04	5,06	3,49	3,62	2,87
Media del error (dB)	1,14	0,88	0,96	0,98	1,06	0,86
Desviación típica del error (dB)	1,05	0,76	0,83	0,94	1,00	0,80

Tabla 40. Resultados de las cinco ejecuciones de la red con el conjunto 4 (con ACP)

<i>CONJUNTO 4(7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	12	15	11	14	12	
Ejemplos erróneos (Entre/Test)	1/3	1/5	1/2	0/2	1/3	1/2
Error máximo conjunto test (dB)	13,55	12,34	12,56	12,03	10,16	12,13
Media del error (dB)	0,85	0,81	0,83	0,83	0,78	0,82
Desviación típica del error (dB)	1,08	1,04	1,03	1,01	0,97	1,00

Tabla 41. Resultados de las cinco ejecuciones de la red con el conjunto 5 (con ACP)

<i>CONJUNTO 5(7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	6	4	7	7	6	
Ejemplos erróneos (Entre/Test)	3/1	5/7	6/8	2/8	4/8	3/6
Error máximo conjunto test (dB)	3,44	12,50	5,09	15,01	7,84	7,73
Media del error (dB)	1,02	1,17	1,35	1,16	1,13	1,17
Desviación típica del error (dB)	0,94	1,56	1,28	1,63	1,23	1,19

Observando las Tablas 37-41, se ve que se reduce bastante el tiempo de entrenamiento respecto del uso de datos sin preprocesamiento. De forma similar al entrenamiento con entradas sin preprocesamiento, los conjuntos 1, 3 y 4 son los que requieren más épocas en el entrenamiento, pero son también los que proporcionan mejores resultados.

De forma parecida, se obtiene una media del error baja en los conjuntos 1 y 4, coincidiendo con dos de los tres conjuntos con más épocas en el entrenamiento. Por otro lado, las cinco ejecuciones se comportan de manera diferente según los datos, lo que es consecuente con la obtención aleatoria de los conjuntos de entrenamiento y test.

El número total de ejemplos erróneos es más alto, para todas las ejecuciones, que cuando no hay preprocesamiento en las entradas. Para la salida media, los resultados siguen siendo buenos, ya que el número de errores se mantiene en 3 para los conjuntos 1, 3 y 4, y algo más alto en el 2 y el 5. Para los 5 conjuntos, solo hay un total de 24 ejemplos erróneos (menos del 2%).

El máximo error cometido es de 15 dB, y hay en total 17 ejemplos en los que el error supera 9 dB. Teniendo en cuenta que se han hecho pruebas con un total de 7225 registros de datos (25 conjuntos de 289 datos cada uno), 17 errores supone, a pesar de todo, un muy bajo porcentaje (0,23%). Mirando la media del error para todos los datos, se obtiene un valor de 0.92 dB, lo que es más que aceptable.

Es muy frecuente que sea el mismo ejemplo en el que la red se equivoca para todas sus ejecuciones, siendo esto visible en los conjuntos de datos 1 y 4. En el conjunto 4 se encuentran tres ejemplos en los que en todas las ejecuciones se da un error mayor al 5%. Lo mismo ocurre con el conjunto 1 en dos ocasiones. No son ejemplos cuya salida está comprendida en algún rango concreto, ni tampoco sus valores de entrada guardan alguna

relación de la que se puede intuir que la red no haya aprendido correctamente. En cambio, en ambos conjuntos, coincide el ejemplo de error en tres ocasiones. Es decir, en los conjuntos de datos 1 y 4 el error sube más de un 5% en los ejemplos 104, 126 y 186. En el resto de conjuntos, se observaron los errores para estos 3 registros de datos y también es alto. Incluso cuando forman parte del conjunto de entrenamiento, la red no aprende correctamente estos ejemplos. La desviación típica se mantiene alrededor de 1, aunque en el conjunto 1 baja un poco y en el 5 sube. La media de las desviaciones de las 25 ejecuciones fue de 1.03276, siendo un valor aceptable ya que la media del error es baja. El análisis de la salida media de las cinco ejecuciones, mejora los resultados respecto de las salidas de cada RN individual, lo que de nuevo justifica su utilización.

Como conclusión a este apartado, se considera interesante la introducción del Análisis de Componentes Principales en nuestro problema, ya que, a pesar de no dar unos resultados tan buenos como sin preprocesamiento de entradas, es positiva la obtención de una importante reducción en el conjunto de entradas.

Se muestran en las Figuras 68 y 69, las gráficas comparativas del nivel de ruido medido con el obtenido, para los conjuntos de entrenamiento y de test para el conjunto de datos 5. Al igual que en gráficas anteriores, la línea interior (más clara) indica $\pm 5\%$ respecto de 60dB ($\pm 3\text{dB}$) y la exterior $\pm 5\%$ respecto de 80 dB ($\pm 4\text{dB}$). Se ve que en el conjunto de entrenamiento, tres ejemplos sobrepasan los límites permitidos, mientras que en el conjunto de test lo hace uno.

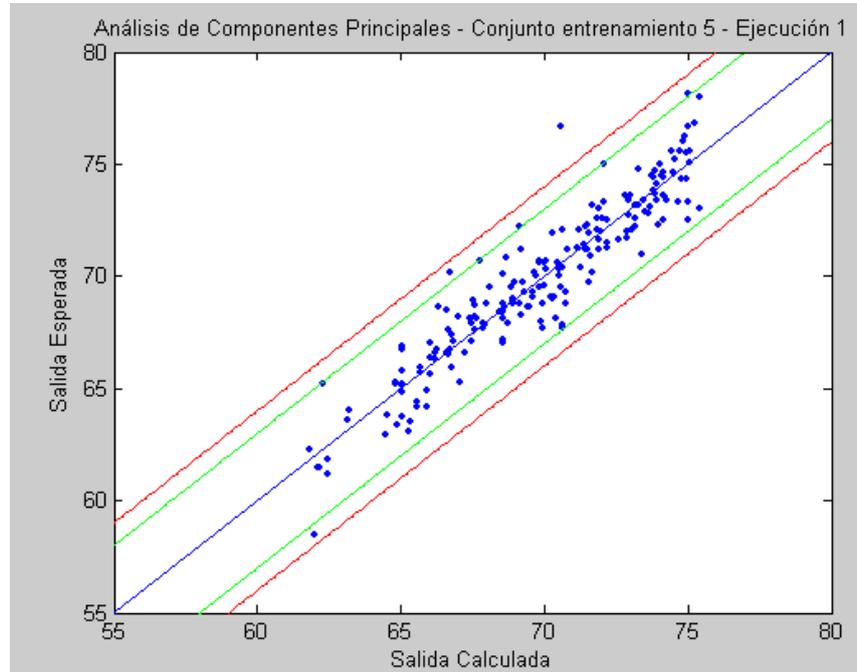


Figura 68. Gráfica comparativa resultados conjunto entrenamiento 5, Ejecución 1 (ACP)

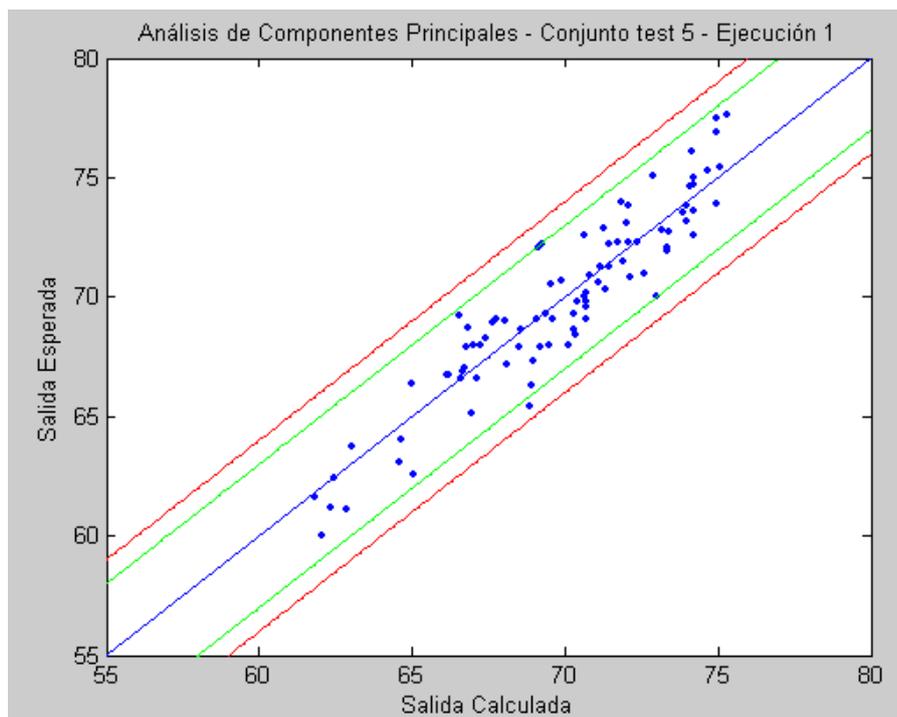


Figura 69. Gráfica comparativa resultados conjunto test 5, ejecución 1 (ACP)

En el anexo se incluyen las tablas de resultados para todos los conjuntos de datos, así como gráficas comparativas para las salidas medias de todos los conjuntos de datos.

El Análisis de Componentes Principales ha mostrado un comportamiento aceptable, reduciendo a 11 el número de características sin coincidir ninguna con alguna de las variables del estado inicial. Entonces se plantea realizar un proceso de Selección de Características y se asumirá, de acuerdo con los resultados, que el número mínimo de características a seleccionar debe de estar en torno a 11. En las siguientes secciones de este capítulo se usarán distintos métodos de selección de características, ya descritos en el capítulo I, y se planteará un rango del número de características entre 10 y 20.

4.2 SELECCIÓN DE CARACTERÍSTICAS

Se van a aplicar diferentes métodos de selección de características a los datos de ruido ambiental disponibles. Dadas las diferentes características de los métodos, se van a seleccionarse los más dispares posibles. Para ello, se buscarán:

1. Métodos que clasifiquen variables individualmente, premiando variables que tienen un poder predictivo individual considerable. Estos métodos devuelven un ranking de características, a partir de las cuales se podrá elegir el número de variables deseado.
2. Métodos que clasifiquen variables por subconjuntos: tipo filtro

3. Métodos que clasifiquen variables por subconjuntos: tipo wrapper

Con los diferentes métodos de selección de características, solo se va a considerar el conjunto 1 de las 5 selecciones aleatorias con las que se ha trabajado anteriormente.

En todos los métodos, excepto en el de mRMR (mínima redundancia, máxima relevancia) se ha considerado la implementación WEKA [66].

A. MÉTODOS QUE CLASIFICAN VARIABLES DE FORMA INDIVIDUAL

1. Relief

Una de las extensiones de Relief, Relief-F, ya comentada en el capítulo I, se ha ejecutado, con los datos del problema, con diferentes números de vecinos más cercanos (instancias de igual y diferente clase con las que se compara). La novedad que aporta Relief-F con respecto a Relief es que calcula la habilidad de los atributos para separar clases independientemente de qué clases sean los vecinos más cercanos. Para ello, no sólo calcula el vecino más cercano con una clase diferente (near-miss), sino que también encuentra vecinos del resto de clases diferentes y estima la media de sus aportaciones, usando finalmente este valor medio para el cálculo de los pesos de cada característica.

Tabla 42 pueden verse los resultados al ejecutar el método Relief-F con los datos de ruido ambiental con los que se está trabajando. Se han hecho pruebas para 10, 15, 20 y 50 vecinos más cercanos. Al ser éste un método que sigue una estrategia de evaluación de los atributos individual, devuelve una lista ordenada de características.

Tabla 42. Ranking de variables para el método Relief-F

Número de vecinos más cercanos	Ranking de variables
10	13,14,12,5,3,7,10,9,24,1,4,2,19,18,22,20,21,15,25,23,17,16,6,11,8
15	14,13,12,5,3,16,4,7,2,20,17,9,21,25,19,15,18,24,22,1,23,10,6,11,8
20	14,13,3,5,12,2,16,21,20,17,4,25,9,19,7,18,15,24,22,10,23,1,11,6,8
50	16,2,3,21,4,17,14,23,20,12,13,15,25,19,22,5,18,10,7,6,9,8,24,11,1

Se eligen los valores cuya posición promedio sea mejor para hacer un ranking. Los valores medios de las posiciones por variable se muestran en la Tabla 43.

Elegidas las mejores características, se harán pruebas con las primeras 10, 12 y 15 (se marcan en negrita las nuevas características incluidas en cada grupo):

10 variables: 2, 3, 4, 5, 12,13, 14, 16, 20, 21

12 variables: 2, 3, 4, 5, **7**, 12,13, 14, 16, **17**, 20, 21

15 variables: 2, 3, 4, 5, 7, 9, 12, 13, 14, 16, 17, 19, 20, 21, 25

Para comprobar la precisión con la que la red neuronal predecirá el valor del nivel de presión sonora solamente con el uso de estas 10, 12 y 15 variables, se prueba la red neuronal con las mismas. Para la selección de características, en lugar de hacer 5 ejecuciones con cada uno de los conjuntos de datos con los que se viene trabajando, se ha elegido el conjunto 1, realizándose todas las pruebas con el mismo. Los resultados se observan en las Tablas 44 - 49.

Tabla 43. Ordenación de las variables según el método Relief-F

Puesto	Variable	Media Posiciones
1	14	2,75
2	3	4
3	13	4
4	12	5,25
5	5	7
6	2	7,25
7	4	8,5
8	16	9
9	21	10,5
10	20	11
11	7	12
12	17	12
13	9	13,5
14	19	14
15	25	14,5
16	15	15,75
17	18	16
18	10	16,75
19	22	17
20	24	17
21	23	17,5
22	1	19,25
23	6	22,5
24	11	23,75
25	8	24,25

De las Tablas 44 y 45, cabe destacar que, para 9 neuronas en la capa oculta, se observa un importante descenso en el número de épocas, aunque los resultados son ligeramente peores que para 7 neuronas en la capa oculta. Se observa también que el número de ejemplos de test erróneos es similar en ambos casos (además, los errores

tienden a cometerse en los mismos ejemplos). Cabe destacar que el único caso en que no se comete ningún error superior al 5% es en el conjunto de entrenamiento en las salidas medias con 7 neuronas en la capa intermedia. El error cuadrático medio queda siempre por debajo de 1,5 dB.

10 variables → 2, 3, 4, 5, 12, 13, 14, 16, 20, 21

Tabla 44. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (Relief-F)

CONJUNTO 1 (7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	12	13	8	11	10	
Ejemplos erróneos (Entre/Test)	3/7	2/9	3/8	3/10	4/9	0/2
Error máximo conjunto test (dB)	5,71	6,20	5,70	5,88	4,54	5,24
Media del error (dB)	1,34	1,47	1,38	1,41	1,36	1,27
Desviación típica del error (dB)	0,76	0,80	0,77	0,78	0,76	0,75

Tabla 45. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (Relief-F)

CONJUNTO 1 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	7	6	11	8	
Ejemplos erróneos (Entre/Test)	3/7	3/9	6/8	4/8	4/9	1/3
Error máximo conjunto test (dB)	5,97	4,89	5,15	5,09	5,45	5,20
Media del error (dB)	1,39	1,36	1,39	1,39	1,34	1,30
Desviación típica del error (dB)	0,79	0,76	0,79	0,78	0,78	0,76

En general y como cabe esperar, los resultados son algo peores que con el conjunto de las 25 características, aunque el bajo número de errores en el test es positivo, sobre todo teniendo en cuenta la gran reducción de variables que supone.

En las Tablas 46 y 47 se muestran los resultados para 12 variables.

12 variables → 2, 3, 4, 5, 7, 12, 13, 14, 16, 17, 20, 21

Tabla 46. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (Relief-F)

CONJUNTO 1 (7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	10	12	13	22	9	
Ejemplos erróneos (Entre/Test)	2/6	1/6	2/8	1/6	2/7	0/2
Error máximo conjunto test (dB)	5,04	4,97	5,45	7,49	19,16	4,25
Media del error (dB)	1,09	1,02	1,12	1,37	2,26	0,99
Desviación típica del error (dB)	0,71	0,69	0,72	0,84	1,27	0,68

Tabla 47. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (Relief-F)

CONJUNTO 1 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	27	24	12	18	15	
Ejemplos erróneos (Entre/Test)	1/6	1/7	0/6	1/7	1/8	0/2
Error máximo conjunto test (dB)	5,50	4,66	4,82	11,94	5,86	4,79
Media del error (dB)	1,17	1,04	1,01	1,53	1,21	1,03
Desviación típica del error (dB)	0,73	0,67	0,68	0,96	0,77	0,69

En el caso de 12 variables, y teniendo en cuenta el número de errores de test, los resultados son solamente un poco mejores que para 10 variables. Existen algunas ejecuciones (en concreto, la 4ª ejecución con 9 neuronas en la capa oculta) que arrojan un error cuadrático medio alto, pero que esto no repercute negativamente en el correspondiente número de ejemplos erróneos en test.

Por otro lado, el número de épocas es mayor que para 10 variables, haciendo que se ajuste mejor el entrenamiento, siendo patente en el reducido número de ejemplos erróneos. También en test se reducen estos errores respecto a 10 variables, aunque menos drásticamente.

En general, los resultados son buenos, y se considera que han mejorado respecto a la prueba con 10 características.

En las Tablas 48 y 49 se pueden observar los resultados de aplicar la red neuronal a las 15 variables seleccionadas por Relief.

15 variables → 2, 3, 4, 5, 7, 9, 12, 13, 14, 16, 17, 19, 20, 21, 25

Tabla 48. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (Relief-F)

CONJUNTO 1 (7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	10	7	8	7	8	
Ejemplos erróneos (Entre/Test)	3/7	2/5	0/8	2/4	0/5	0/2
Error máximo conjunto test (dB)	4,07	2,96	4,35	5,77	4,64	4,24
Media del error (dB)	1,19	1,05	1,07	1,20	0,96	0,93
Desviación típica del error (dB)	0,70	0,65	0,68	0,72	0,66	0,66

Tabla 49. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (Relief-F)

CONJUNTO 1 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	6	9	7	11	
Ejemplos erróneos (Entre/Test)	2/7	0/9	4/9	1/7	1/7	0/2
Error máximo conjunto test (dB)	4,53	4,39	4,67	6,78	6,98	5,08
Media del error (dB)	1,07	1,14	1,26	1,25	1,33	1,03
Desviación típica del error (dB)	0,69	0,72	0,74	0,79	0,82	0,71

Ejecutando la red neuronal 5 veces con las 15 características elegidas se observa una ligera mejoría en el caso de la red con 7 neuronas en la capa oculta. No obstante, el número de ejemplos erróneos queda en cifras similares a las pruebas con 10 y 12 características.

El número de épocas es bajo en todos los casos, y a pesar de ello, la red se ajusta bien a los ejemplos de entrenamiento. El error medio en entrenamiento y test es bajo, pero ya que el número de errores en test permanece alto, no parece haber ganancia significativa con la inclusión de 3 características más respecto al caso anterior.

En resumen, una buena solución que ofrece Relief-F para el problema planteado es reducir las características a 15 o incluso a 12.

2. Evaluador de atributos según la ganancia de información

Para el uso de este método, ya que se evalúa la ganancia de información respecto a la clase, se han discretizado los valores de salida L_{Aeq} , previamente normalizados entre 0 y 1. Se han dividido en 5 clases, incluidas en la Tabla 50, obteniendo en cada clase un número de ejemplos similar.

Tabla 50. Clases al discretizar L_{Aeq}

Clase	Valores	Número de instancias
1	(-inf-0.3]	57
2	(0.3-0.4]	58
3	(0.4-0.5]	58
4	(0.5-0.6]	60
5	(0.6-inf)	56

La aplicación del método devuelve un ranking de variables, que se muestra en la Tabla 51, junto con el valor de la ganancia obtenido para cada una de ellas.

Escogiendo las 10, 12, y 15 mejores variables y probando con la red neuronal, queda patente que la evaluación de atributos de forma individual no implica unos resultados óptimos, ya que las variables pueden ser dependientes entre sí. Los resultados se muestran en las Tablas 52 - 57.

Tabla 51. Ranking de variables según el cálculo de la ganancia de información.

	#	Nombre variable	Ganancia
1	24	Anchura de la calzada	0.854
2	4	Tiempo de estabilización	0.788
3	25	Distancia al receptor	0.765
4	22	Anchura de la vía	0.751
5	16	Pendiente	0.73
6	23	Altura de la vía	0.59
7	8	Flujo ascendente de vehículos pesados	0.526
8	5	Tipo de flujo de tráfico	0.501
9	10	Flujo ascendente de motos	0.494
10	6	Flujo ascendente de vehículos ligeros	0.387
11	11	Flujo descendente de motos	0.291
12	21	Tipo de vía	0.277
13	17	Número de carriles ascendente	0.249
14	15	Velocidad de circulación promedio	0.235
15	13	Eventos relacionados con tráfico	0.2
16	18	Número de carriles descendente	0.193
17	9	Flujo descendente de vehículos pesados	0.189
18	19	Tipo de pavimento	0.156
19	7	Flujo descendente de vehículos ligeros	0.138
20	20	Estado del firme	0.129
21	2	Entorno comercial o de ocio	0.117
22	14	Eventos no relacionados con tráfico	0.112
23	1	Periodo del día	0
24	12	Número de vehículos con sirena	0
25	3	Presencia de obras	0

10 variables → 2, 4, 5, 6, 8, 10, 16, 23, 24, 25

Tabla 52. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (InfoGain)

CONJUNTO 1 (7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	15	9	15	13	7	
Ejemplos erróneos (Entre/Test)	'15/15'	'15/12'	'14/14'	'14/12'	'16/13'	4/7
Error máximo conjunto test (dB)	13,03	13,26	13,16	13,61	12,87	13,23
Media del error (dB)	2,77	2,84	2,70	2,68	2,98	2,69
Desviación típica del error (dB)	1,20	1,21	1,20	1,20	1,24	1,19

Tabla 53. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (InfoGain)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	15	15	15	15	10	
Ejemplos erróneos (Entre/Test)	'13/13'	'16/11'	'14/12'	'13/13'	'15/13'	5/6
Error máximo conjunto test (dB)	13,79	13,48	13,20	13,19	13,34	13,39
Media del error (dB)	2,72	2,61	2,71	2,62	2,63	2,53
Desviación típica del error (dB)	1,20	1,17	1,19	1,16	1,20	1,16

12 variables→2, 4, 5, 6, 8, 10, 11,16, 21, 23, 24, 25

Tabla 54.Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (InfoGain)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	14	9	8	15	9	
Ejemplos erróneos (Entre/Test)	'14/10'	'13/13'	'14/14'	'14/12'	'16/9'	'3/6'
Error máximo conjunto test (dB)	13,44	13,48	13,16	13,59	13,49	13,40
Media del error (dB)	2,71	2,46	2,65	2,65	2,65	2,46
Desviación típica del error (dB)	1,16	1,16	1,18	1,17	1,17	1,15

Tabla 55. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (InfoGain)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	15	10	11	13	9	
Ejemplos erróneos (Entre/Test)	'12/13'	'11/16'	'14/13'	'14/16'	'11/18'	'2/6'
Error máximo conjunto test (dB)	13,76	13,48	13,16	13,59	13,49	13,37
Media del error (dB)	2,54	2,62	2,62	2,63	2,62	2,46
Desviación típica del error (dB)	1,19	1,16	1,18	1,17	1,17	1,16

15 variables→2, 4, 5, 6, 8, 10, 11, 13, 15, 16,17, 21, 23, 24, 25

Tabla 56. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (InfoGain)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	8	7	8	7	6	
Ejemplos erróneos (Entre/Test)	'14/12'	'11/10'	'10/8'	'12/7'	'11/9'	'2/6'
Error máximo conjunto test (dB)	11,91	14,26	13,28	13,21	13,38	13,11
Media del error (dB)	2,32	2,32	2,13	2,14	2,38	2,06
Desviación típica del error (dB)	1,10	1,16	1,11	1,12	1,15	1,11

Tabla 57. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (InfoGain)

CONJUNTO 1 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	9	6	7	5	6	
Ejemplos erróneos (Entre/Test)	'10/10'	'11/8'	'10/9'	'10/9'	'9/8'	'2/5'
Error máximo conjunto test (dB)	11,91	13,60	13,28	13,21	13,38	13,39
Media del error (dB)	2,27	2,22	2,21	2,22	2,11	2,04
Desviación típica del error (dB)	1,10	1,21	1,11	1,12	1,15	1,11

En las Tablas 52 - 57, se observa que el número de ejemplos erróneos es mayor al obtenido para las pruebas con las variables que devuelve el método Relief.

De la misma manera, tanto la media del error como la desviación típica del mismo es alta. Aunque el error máximo del conjunto de test se mantiene constante, este valor es normalmente alto (en torno a 13 dB).

En cuanto al número de épocas de entrenamiento, viene disminuyendo conforme se añaden características al conjunto de entrada. No obstante, este valor no tiene repercusión positiva en los resultados, que son en general menos buenos que para los obtenidos para otros métodos.

En resumen, no parece que la evaluación de atributos según la ganancia de información aporte un conjunto de características con los que la red aproxime el objetivo de forma óptima.

B. MÉTODOS QUE CLASIFICAN SUBCONJUNTOS DE VARIABLES: TIPO FILTRO

1. Selección de subconjuntos basado en la correlación

Se ha implementado este método para los datos disponibles, junto con diferentes algoritmos de búsqueda disponibles en Weka [66]. Los resultados se muestran en la Tabla 58. El número de variables elegidas usando cada algoritmo de búsqueda es similar, moviéndose en torno a 9-11 variables. Además en muchos casos coincide un gran número de ellas.

Probando la red neuronal con las 9, 10 y 11 variables que devuelven los métodos de búsqueda indicados en la Tabla 58, los resultados que ésta obtiene son los indicados en las páginas siguientes. Se ha ejecutado la red neuronal 5 veces, para el mismo conjunto de entrenamiento y test, para 7 y 9 neuronas en la capa intermedia. Se pueden observar estos resultados en las Tablas 59 - 64.

Tabla 58. Resultados para el método de selección de características basado en la correlación

<i>Método de búsqueda</i>		Variables elegidas: número total
<i>Primero el mejor hacia adelante</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Primero el mejor hacia atrás</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Primero el mejor bidireccional</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Búsqueda Exhaustiva</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Búsqueda genética</i> Probabilidad de cruce: 0.6 Probabilidad de mutación: 0.033	Tamaño población: 20 Número de generaciones: 20	5,8,10,11,12,13,14,16,17,19 : 10
	Tamaño población: 100 Número de generaciones: 20	4,5,8,9,10,12,13,14,16,17,19 : 11
	Tamaño población: 1000 Número de generaciones: 35	5,8,10,11,12,13,14,16,17,19 : 10
	Tamaño población: 1000 Número de generaciones: 150	5,8,10,11,12,13,14,16,17 : 9
	Tamaño población: 1000 Número de generaciones: 250	5,8,10,11,12,13,14,16,17 : 9
<i>Greedy StepWise</i>		8,17,10,13,5,12,11,19,16,4,14,9,6,22,20,2 3,25,21,15,18,24,2,7,3,1
<i>Selección lineal hacia adelante</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Selección lineal hacia adelante (flotante)</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Búsqueda al azar</i>	Nº de iteraciones: 8.388.608 (25.0% del espacio de búsqueda)	5,8,9,10,12,13,14,16,17,19 : 10
<i>ScatterSearchV1 Greedy Combination</i>		5,8,10,11,12,13,14,16,17 : 9
<i>ScatterSearchV1 Reduced Greedy Combination</i>		5,8,10,11,12,13,14,16,17 : 9
<i>Subset Size Forward Selection</i>		5,8,10,11,12,13,14,16,17 : 9

9 variables→5,8,10,11,12,13,14,16,17

Tabla 59. Resultados de las 5 ejecuciones de la red con el conjunto 1, 9 variables (correlación)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	5	10	11	8	12	
Ejemplos erróneos (Entre/Test)	'11/10'	'10/11'	'12/9'	'11/12'	'10/9'	'2/5'
Error máximo conjunto test (dB)	4,68	6,28	6,27	8,78	8,04	6,39
Media del error (dB)	1,85	1,86	1,93	2,07	1,98	1,80
Desviación típica del error (dB)	0,87	0,88	0,91	0,97	0,93	0,88

Tabla 60. Resultados de las 5 ejecuciones de la red con el conjunto 1, 9 variables (correlación)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	9	7	9	6	6	
Ejemplos erróneos (Entre/Test)	'10/15'	'10/11'	'11/14'	'11/10'	'12/11'	'2/4'
Error máximo conjunto test (dB)	6,28	6,27	9,32	8,78	8,04	6,62
Media del error (dB)	1,95	2,18	1,77	1,96	1,99	1,73
Desviación típica del error (dB)	0,88	0,91	0,99	0,97	0,93	0,89

10 variables→5,8,10,11,12,13,14,16,17,19

Tabla 61. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (correlación)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	15	15	14	15	15	
Ejemplos erróneos (Entre/Test)	'5/8'	'3/9'	'3/9'	'4/7'	'6/6'	'1/2'
Error máximo conjunto test (dB)	4,70	6,14	5,15	6,82	8,57	3,69
Media del error (dB)	1,33	1,42	1,29	1,44	1,53	1,06
Desviación típica del error (dB)	0,78	0,84	0,76	0,82	0,85	0,67

Tabla 62. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (correlación)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	11	13	8	15	15	6
Ejemplos erróneos (Entre/Test)	'5/8'	'4/7'	'4/7'	'3/6'	'3/9'	'1/2'
Error máximo conjunto test (dB)	4,70	8,29	6,14	6,11	5,15	3,49
Media del error (dB)	1,30	1,42	1,26	1,13	1,39	1,07
Desviación típica del error (dB)	0,78	0,86	0,84	0,84	0,76	0,68

11 variables→4,5,8,9,10,12,13,14,16,17,19

Tabla 63. Resultados de las 5 ejecuciones de la red con el conjunto 1, 11 variables (correlación)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	11	13	12	11	18	
Ejemplos erróneos (Entre/Test)	'7/10'	'10/10'	'8/7'	'5/9'	'5/8'	'0/2'
Error máximo conjunto test (dB)	5,61	4,52	4,04	6,19	4,02	4,37
Media del error (dB)	1,51	1,47	1,37	1,62	1,27	1,23
Desviación típica del error (dB)	0,81	0,80	0,74	0,86	0,74	0,72

Tabla 64. Resultados de las 5 ejecuciones de la red con el conjunto 1, 11 variables (correlación)

<i>CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	8	12	18	11	18	
Ejemplos erróneos (Entre/Test)	'4/7'	'5/6'	'4/10'	'4/7'	'4/5'	'1/2'
Error máximo conjunto test (dB)	5,99	4,52	4,04	9,73	4,02	3,56
Media del error (dB)	1,25	1,16	1,31	1,32	1,29	1,06
Desviación típica del error (dB)	0,84	0,80	0,74	1,00	0,74	0,68

De las Tablas 59 - 64 se extrae que el número de ejemplos erróneos disminuye considerablemente al añadir una característica más al conjunto de 9 variables. Al pasar a 11 características, en cambio, éste se mantiene similar.

Por otro lado, el error máximo del conjunto de test se mueve en valores similares en los 3 casos (9,10 y 11 variables) aunque la media del error mejora considerablemente para 10 y 11 características. La desviación típica es igualmente algo más alta en el primero de los casos.

Esto indica que la inclusión de la variable 19 (tipo de pavimento) ha provocado una mejoría en los resultados, no obteniendo una mejoría de igual medida al incluir la variable 4 (tiempo de estabilización).

En general, los resultados obtenidos con la clasificación de características basándose en la correlación son buenos y proporcionan información concreta y útil acerca de qué variables pueden ayudar a aproximar mejor el nivel de presión sonora.

2. Evaluación de la consistencia de los subconjuntos

Este método sólo es capaz de obtener un subconjunto de características cuando la variable objetivo es nominal, es decir, ha sido necesario dividir en clases el valor de L_{Aeq} . Para poder trabajar con este método, se ha dividido la variable objetivo en 5 clases, una vez normalizado su valor, mostradas en la Tabla 50.

Se ha probado utilizando diferentes métodos de búsqueda, mostrando los resultados en la Tabla 65.

Las variables elegidas por la mayoría de los métodos de búsqueda combinados con el algoritmo de clasificación son:

9 variables -> 5,6,8,9,10,11,13,14,24

10 variables - >5,6,8,9,10,11,13,14,17,24

10 variables ->2,5,6,8,9,10,11,13,14,24

11 variables -> 5,7,8,9,10,11,13,14,18,23,25

14 variables ->3,5,6,8,9,10,11,12,13,14,15,16,19,20

Tabla 65. Resultados para el método de selección de características según la consistencia

<i>Método de búsqueda</i>		<i>Variables elegidas</i>
<i>Best First Forward</i>		5,6,8,9,10,11,13,14,24 : 9
<i>Best First Backward</i>		5,6,8,9,10,11,13,14,24 : 9
<i>Best First BiDirectional</i>		5,6,8,9,10,11,13,14,24 : 9
<i>Búsqueda genética</i> Probabilidad de cruce: 0.6 Probabilidad de mutación: 0.033	Tamaño población: 20 Número de generaciones: 20	3,5,6,8,9,10,11,12,13,14,15,16,19,20 : 14
	Tamaño población: 100 Número de generaciones: 20	2,5,6,8,9,10,11,13,14,24 : 10
	Tamaño población: 1000 Número de generaciones: 35	5,6,8,9,10,11,13,14,17,24 : 10
	Tamaño población: 1000 Número de generaciones: 150	5,6,8,9,10,11,13,14,24 : 9
	Tamaño población: 1000 Número de generaciones: 250	5,6,8,9,10,11,13,14,24 : 9
<i>Greedy StepWise</i>		5,6,8,9,10,11,13,14,24 : 9
<i>Linear Forward Selection</i>		5,6,8,9,10,11,13,14,24 : 9
<i>Linear Forward Selection (flotante)</i>		5,6,8,9,10,11,13,14,24 : 9
Búsqueda al azar	Número de iteraciones: 8388608 (25.0% del espacio de búsqueda)	1,2,5,6,7,8,9,10,11,13,14,16,18,20,23,25 : 16
<i>ScatterSearchV1 Greedy Combination</i>		5,6,8,9,10,11,13,14,24 : 9
<i>ScatterSearchV1 Reduced Greedy Combination</i>		5,7,8,9,10,11,13,14,18,23,25 : 11
<i>Subset Size Forward Selection</i>		5,6,8,9,10,11,13,14,24 : 9

Los resultados de probar la red neuronal con estas variables se muestran en las Tablas 66 - 75.

9 variables→5,6,8,9,10,11,13,14,24

Tabla 66. Resultados de las 5 ejecuciones de la red con el conjunto 1, 9 variables (consistencia)

<i>CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	16	11	15	14	8	
Ejemplos erróneos (Entre/Test)	'5/11'	'13/12'	'7/8'	'11/11'	'20/12'	'1/4'
Error máximo conjunto test (dB)	10,16	10,96	9,53	12,66	10,66	10,56
Media del error (dB)	1,87	2,31	1,97	2,15	2,51	1,84
Desviación típica del error (dB)	0,98	1,08	1,02	1,09	1,11	1,00

Tabla 67. Resultados de las 5 ejecuciones de la red con el conjunto 1, 9 variables (consistencia)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	15	15	15	13	15	
Ejemplos erróneos (Entre/Test)	'9/12'	'5/12'	'7/14'	'6/9'	'7/13'	'1/5'
Error máximo conjunto test (dB)	10,16	10,96	12,66	10,68	10,66	9,22
Media del error (dB)	1,94	2,07	2,26	1,93	1,93	1,70
Desviación típica del error (dB)	0,98	1,08	1,09	1,15	1,11	0,95

10 variables→5,6,8,9,10,11,13,14,17,24

Tabla 68. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (consistencia)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	11	15	11	10	12	
Ejemplos erróneos (Entre/Test)	'5/10'	'4/10'	'10/14'	'10/9'	'10/16'	'2/3'
Error máximo conjunto test (dB)	10,72	9,92	11,08	10,41	10,10	10,41
Media del error (dB)	1,71	1,74	2,08	2,00	2,17	1,63
Desviación típica del error (dB)	0,98	0,99	1,06	1,02	1,05	0,94

Tabla 69. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (consistencia)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	9	15	17	14	11	
Ejemplos erróneos (Entre/Test)	'8/10'	'10/9'	'3/9'	'8/10'	'8/9'	'2/3'
Error máximo conjunto test (dB)	10,72	9,92	11,06	11,08	10,41	9,54
Media del error (dB)	1,73	1,85	1,37	1,60	1,87	1,44
Desviación típica del error (dB)	0,98	0,99	1,16	1,06	1,02	0,90

10 variables→2,5,6,8,9,10,11,13,14,24

Tabla 70. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (consistencia)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	15	8	13	14	13	
Ejemplos erróneos (Entre/Test)	'4/10'	'5/7'	'6/8'	'6/11'	'6/6'	'3/3'
Error máximo conjunto test (dB)	12,27	10,78	9,83	11,38	11,48	11,49
Media del error (dB)	1,91	1,64	1,67	1,82	1,68	1,55
Desviación típica del error (dB)	1,03	0,95	0,96	0,99	0,98	0,96

Tabla 71. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (consistencia)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	14	10	11	9	8	
Ejemplos erróneos (Entre/Test)	'6/8'	'5/12'	'5/8'	'6/10'	'6/10'	'2/3'
Error máximo conjunto test (dB)	14,12	12,27	10,78	10,54	11,38	10,99
Media del error (dB)	1,71	2,00	1,67	1,76	2,10	1,54
Desviación típica del error (dB)	1,15	1,03	0,95	0,96	0,99	0,95

11 variables→5,7,8,9,10,11,13,14,18,23,25

Tabla 72. Resultados de las 5 ejecuciones de la red con el conjunto 1, 11 variables (consistencia)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	13	12	13	16	13	
Ejemplos erróneos (Entre/Test)	'6/12'	'7/6'	'5/7'	'18/14'	'5/10'	'2/3'
Error máximo conjunto test (dB)	13,98	12,98	13,93	9,49	14,69	13,49
Media del error (dB)	2,40	1,99	1,86	2,29	2,00	1,80
Desviación típica del error (dB)	1,13	1,02	1,05	1,06	1,11	1,04

Tabla 73. Resultados de las 5 ejecuciones de la red con el conjunto 1, 11 variables (consistencia)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	13	13	9	10	11	
Ejemplos erróneos (Entre/Test)	'2/6'	'4/8'	'2/7'	'5/6'	'3/7'	'2/2'
Error máximo conjunto test (dB)	13,98	12,98	13,42	9,49	14,69	13,82
Media del error (dB)	1,96	1,76	1,90	1,89	1,74	1,65
Desviación típica del error (dB)	1,13	1,02	1,17	1,06	1,11	1,01

14 variables→3,5,6,8,9,10,11,12,13,14,15,16,19,20

Tabla 74. Resultados de las 5 ejecuciones de la red con el conjunto 1, 14 variables (consistencia)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	13	17	14	18	14	
Ejemplos erróneos (Entre/Test)	'4/5'	'1/10'	'1/8'	'1/7'	'1/5'	'1/2'
Error máximo conjunto test (dB)	4,11	9,80	4,40	3,80	4,23	3,62
Media del error (dB)	1,10	1,44	0,98	0,93	0,89	0,80
Desviación típica del error (dB)	0,71	0,92	0,69	0,66	0,64	0,60

Tabla 75. Resultados de las 5 ejecuciones de la red con el conjunto 1, 14 variables (consistencia)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	14	14	10	11	10	
Ejemplos erróneos (Entre/Test)	'1/4'	'1/12'	'1/5'	'2/6'	'1/4'	'1/2'
Error máximo conjunto test (dB)	14,04	9,80	4,40	3,80	4,23	5,49
Media del error (dB)	0,97	1,36	1,12	0,97	0,89	0,82
Desviación típica del error (dB)	1,24	0,92	0,69	0,66	0,64	0,64

En las Tablas 66 y 67 se observa que el conjunto de 9 variables no ofrece unos resultados muy buenos. El error máximo del conjunto de test es alto (en torno a 10dB) en todas las ejecuciones y la media del mismo pasa de 2dB en varios casos. Comparando con el primero de los conjuntos de 10 variables (incluyendo la variable 17- número de carriles ascendentes, Tabla 68 - Tabla 69), se observa que estos datos se mantienen para el caso de 7 neuronas en la capa intermedia, mejorando ligeramente para 9 neuronas.

Al observar las Tablas 70 - 73, se puede observar como este rango de valores se mantiene en el segundo conjunto de 10 variables (incluyendo la característica 2-entorno comercial o de ocio) y con el de 11. Incluso en la media de las salidas, se obtienen 2 y 3 ejemplos erróneos y el error máximo del conjunto de test sólo baja de 10 dB en un par de ocasiones. También la media del error se acerca a 2dB.

Cuando la red neuronal recibe como entrada 14 características diferentes, sí se obtienen unos resultados mejores a los anteriores (Tablas 74 - 75). El error máximo del conjunto de test se mantiene en torno a 4dB en la mayoría de los casos, reduciéndose también el número de ejemplos erróneos. En la media, este valor queda en 1 y 2 para entrenamiento y test, respectivamente. Además, la media del error baja de 1dB, y la desviación típica del error es también baja, llegando a 0,60dB para 7 neuronas en la media de las salidas.

Resumiendo, el método de evaluación de subconjuntos de variables por consistencia proporciona los mejores resultados obtenidos hasta ahora, especialmente al considerar 14 características. La adición de 7 variables al conjunto de 11, y la eliminación de 4 más de éste, ha permitido mejorar considerablemente los resultados. Las variables añadidas son: presencia de obras, flujo ascendente de vehículos ligeros, número de vehículos con sirena, velocidad de circulación promedio, pendiente, tipo de pavimento y estado del firme. Mientras que las variables eliminadas del conjunto de 11 son: altura de la vía, distancia al receptor, número de carriles descendente, flujo descendente de vehículos ligeros.

3. Método de mínima redundancia - máxima relevancia

Este método fue desarrollado en el año 2003 por Hanchuan Peng, Fuhui Long y Chris Ding y ya ha sido comentado en el capítulo I. Sus dos variantes, MID y MIQ, han sido evaluadas con los datos disponibles.

Al ejecutar MID con los datos de nivel de ruido, y pidiendo que devuelva las características más relevantes, se obtuvieron los siguientes resultados:

10 Características → 19,21,16,17,18,14,20,23,22,23

Características más relevantes (MaxRel)	19,21,16,17,18,14,20,23,22,23
Mejor conjunto de características (mRMR)	19,21,16,17,18,14,20,23,22,23

12 Características → 19,21,16,17,18,14,20,22,23,24,6,8

Características más relevantes (MaxRel)	19,21,16,17,18,14,20,13,22,23,24,15
Mejor conjunto de características (mRMR)	19,21,16,17,18,14,20,22,23,24,6,8

15 Características → 19,21,16,17,18,14,20,13,22,23,24,6,8,4,5,7

Características más relevantes (MaxRel)	19,21,16,17,18,14,20,13,22,23,24, 15, 1,6, 8
Mejor conjunto de características (mRMR)	19,21,16,17,18,14,20,13,22,23,24,6,8,4,5,7

Se evaluarán las 10, 12 y 15 mejores características en ambos sentidos (relevancia máxima y redundancia mínima) con la red neuronal. Los resultados se muestran en las Tablas 76 - 81.

10 características → 19,21,16,17,18,14,20,22,23,24

Tabla 76. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (mRMR - MID)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	12	12	12	11	13	
Ejemplos erróneos (Entre/Test)	'13/14'	'13/13'	'14/14'	'13/14'	'13/14'	'6/6'
Error máximo conjunto test (dB)	10,45	10,50	11,03	10,57	10,45	10,75
Media del error (dB)	2,60	2,57	2,69	2,57	2,58	2,60
Desviación típica del error (dB)	1,15	1,14	1,17	1,14	1,14	1,15

Tabla 77. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (mRMR - MID)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	10	8	7	6	10	
Ejemplos erróneos (Entre/Test)	'14/14'	'14/13'	'13/14'	'13/13'	'14/13'	'5/6'
Error máximo conjunto test (dB)	10,45	11,56	11,03	10,57	10,45	10,46
Media del error (dB)	2,64	2,55	2,59	2,55	2,55	2,55
Desviación típica del error (dB)	1,15	1,18	1,17	1,14	1,14	1,14

12 Características → 19,21,16,17,18,14,20,22,23,24,6,8

Tabla 78. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (mRMR - MID)

<i>CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	15	11	19	15	10	
Ejemplos erróneos (Entre/Test)	'9/13'	'8/10'	'6/10'	'8/11'	'12/11'	'4/6'
Error máximo conjunto test (dB)	10,57	10,89	10,76	11,53	13,42	11,46
Media del error (dB)	2,31	2,20	2,13	2,17	2,62	2,18
Desviación típica del error (dB)	1,13	1,10	1,07	1,10	1,21	1,12

Tabla 79. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (mRMR - MID)

<i>CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	13	12	6	17	9	
Ejemplos erróneos (Entre/Test)	'10/12'	'7/12'	'7/11'	'8/13'	'7/13'	'4/5'
Error máximo conjunto test (dB)	10,89	10,76	17,44	13,42	11,78	11,24
Media del error (dB)	2,19	2,38	2,24	2,87	2,48	2,03
Desviación típica del error (dB)	1,10	1,07	1,45	1,21	1,22	1,09

15 Características → 19,21,16,17,18,14,20,13,22,23,24,6,8,4,5,7

Tabla 80. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (mRMR - MID)

<i>CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	13	12	10	11	12	
Ejemplos erróneos (Entre/Test)	'4/7'	'3/6'	'5/9'	'3/11'	'2/6'	'1/3'
Error máximo conjunto test (dB)	14,15	14,41	13,06	13,60	15,33	14,54
Media del error (dB)	1,91	1,95	1,95	1,96	1,98	1,78
Desviación típica del error (dB)	1,07	1,09	1,05	1,10	1,12	1,07

Tabla 81. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (mRMR - MID)

<i>CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	12	8	17	12	17	
Ejemplos erróneos (Entre/Test)	'3/7'	'5/8'	'3/9'	'4/5'	'4/9'	'1/2'
Error máximo conjunto test (dB)	14,41	13,18	13,06	13,60	15,33	14,81
Media del error (dB)	1,86	2,07	1,87	1,82	2,15	1,79
Desviación típica del error (dB)	1,09	1,05	1,05	1,10	1,12	1,08

En las Tablas 76 - 81, se puede observar cómo, respecto al error máximo que la red comete en el conjunto de test, no siempre al añadir variables a la red neuronal, y de

esta manera información, se mejoran considerablemente los resultados. Es con 15 características cuando la red neuronal comete errores máximos más altos. Se puede intuir entonces los valores de qué variables son los que más confunden a la red. En este caso, se han añadido las variables 4, 5, y 7 (tiempo de estabilización, tipo de flujo de tráfico, flujo descendente de vehículos ligeros), siendo las mismas sospechosas de conducir a la red a cometer errores altos, según el criterio de este método.

Por otro lado, sí disminuye considerablemente el número de ejemplos erróneos conforme se añaden variables a la red, cayendo desde 5-6 ejemplos en 10 variables a 1-3 en 15 variables. Igual ocurre con la media del error y la desviación típica del mismo, aunque caen más levemente desde 2,6dB a 1,79dB.

Al ejecutar MIQ con los datos de nivel de ruido, y pidiendo que devuelva las características más relevantes, se obtuvieron los siguientes resultados:

10 Características

Características más relevantes (MaxRel)	19,21,16,17,18,14,20,13,22,23
Mejor conjunto de características (mRMR)	19,21,16,17,18,14,20,13,22,23

12 Características

Características más relevantes (MaxRel)	19,21,16,17,18,14,20,13,22,23, 24,15
Mejor conjunto de características (mRMR)	19,21,16,17,18,14,20,13,22,23, 24,15

15 Características

Características más relevantes (MaxRel)	19,21,16,17,18,14,20,13,22,23,24,15, 1,6, 8
Mejor conjunto de características (mRMR)	19,21,16,17,18,14,20,13,22,23,24,15, 1,6, 8

Se evaluarán las 10, 12, y 15 mejores características en ambos sentidos (relevancia máxima y redundancia mínima) con la red neuronal.

Según se observa en las Tablas 82 - 87, para MIQ, también la mejoría es leve al pasar de 10 a 12 variables. Los errores máximos se mantienen en torno a 10-11dB en ambos casos y el número de ejemplos erróneos es, en todas las ejecuciones, de 12-14. La media del error supera 2,5 dB en todos los casos.

10 variables→19,21,16,17,18,14,20,13,22,23**Tabla 82. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (mRMR - MIQ)**

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	12	9	8	9	
Ejemplos erróneos (Entre/Test)	'13/13'	'14/14'	'14/14'	'16/12'	'16/14'	'5/6'
Error máximo conjunto test (dB)	10,42	10,89	10,47	11,53	11,44	11,06
Media del error (dB)	2,55	2,62	2,65	2,66	2,79	2,60
Desviación típica del error (dB)	1,14	1,16	1,16	1,14	1,20	1,15

Tabla 83. Resultados de las 5 ejecuciones de la red con el conjunto 1, 10 variables (mRMR - MIQ)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	14	9	12	13	8	
Ejemplos erróneos (Entre/Test)	'14/13'	'13/14'	'13/14'	'13/14'	'13/14'	'5/6'
Error máximo conjunto test (dB)	10,55	10,89	12,08	11,53	11,44	10,53
Media del error (dB)	2,55	2,59	2,55	2,58	2,63	2,57
Desviación típica del error (dB)	1,14	1,16	1,20	1,14	1,20	1,14

12 variables→19,21,16,17,18,14,20,13,22,23,24,15**Tabla 84. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (mRMR - MIQ)**

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	8	13	12	7	
Ejemplos erróneos (Entre/Test)	'14/13'	'13/13'	'15/13'	'12/14'	'13/13'	'5/6'
Error máximo conjunto test (dB)	10,45	10,44	10,49	10,47	10,45	10,58
Media del error (dB)	2,55	2,56	2,57	2,60	2,56	2,56
Desviación típica del error (dB)	1,14	1,14	1,14	1,15	1,14	1,14

Tabla 85. Resultados de las 5 ejecuciones de la red con el conjunto 1, 12 variables (mRMR - MIQ)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	11	9	10	6	
Ejemplos erróneos (Entre/Test)	'14/13'	'13/13'	'13/13'	'13/13'	'13/13'	'5/6'
Error máximo conjunto test (dB)	10,63	10,44	10,49	10,47	10,45	10,45
Media del error (dB)	2,57	2,55	2,57	2,56	2,55	2,56
Desviación típica del error (dB)	1,15	1,14	1,14	1,15	1,14	1,14

15 variables→19,21,16,17,18,14,20,13,22,23,24,15,1,6,8

Tabla 86. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (mRMR - MIQ)

CONJUNTO 1 (7NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	8	11	12	8	12	
Ejemplos erróneos (Entre/Test)	'8/13'	'8/10'	'7/9'	'7/11'	'11/16'	'3/5'
Error máximo conjunto test (dB)	11,82	9,12	10,68	11,93	10,60	11,39
Media del error (dB)	2,38	2,02	2,12	2,31	2,41	2,12
Desviación típica del error (dB)	1,16	1,04	1,09	1,12	1,12	1,10

Tabla 87. Resultados de las 5 ejecuciones de la red con el conjunto 1, 15 variables (mRMR - MIQ)

	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	9	12	11	7	8	
Ejemplos erróneos (Entre/Test)	'7/10'	'8/10'	'7/11'	'8/11'	'7/11'	'4/5'
Error máximo conjunto test (dB)	12,16	9,12	13,43	11,93	10,60	10,43
Media del error (dB)	2,10	1,93	2,17	2,14	2,04	2,00
Desviación típica del error (dB)	1,16	1,04	1,20	1,12	1,12	1,07

En cambio, al pasar a 15 variables (se añaden 1, 6 y 8, período del día, flujo ascendente de vehículos ligeros, flujo ascendente de vehículos pesados), sí se mejora algo más el resultado, bajando mucho el número de ejemplos erróneos (en las ejecuciones, en la media no cae tan drásticamente). También la media del error baja hasta quedarse en torno a 2dB, aunque el error máximo del conjunto de test se mantiene en valores altos.

En general, con el método de evaluación de características mRMR no se obtienen unos resultados demasiado prometedores, ya que anteriormente se han observado comportamientos mejores con otros conjuntos de variables.

C. MÉTODOS QUE CLASIFICAN SUBCONJUNTOS DE VARIABLES: TIPO WRAPPER

1. Tipo Wrapper con conjunto de entrenamiento

Con Weka [66], usando de clasificador el **perceptrón multicapa (MLP en inglés)**, el subconjunto de características que devuelve es el siguiente: 1, 2, 3, 5, 6, 8, 10, 11, 12, 13, 14, 15, 17, 21, 23, 24, 25 (17 variables). Los resultados con la red neuronal han sido los mostrados en las Tablas 88 - 89.

Tabla 88. Resultados de las 5 ejecuciones de la red con el conjunto 1, 17 variables (wrapper - train)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	10	8	7	5	9	
Ejemplos erróneos (Entre/Test)	'3/6'	'2/5'	'3/4'	'5/6'	'2/6'	'0/3'
Error máximo conjunto test (dB)	6,01	4,41	4,89	5,13	7,15	4,83
Media del error (dB)	1,26	1,04	1,03	1,18	1,22	0,95
Desviación típica del error (dB)	0,76	0,67	0,68	0,72	0,79	0,67

Tabla 89. Resultados de las 5 ejecuciones de la red con el conjunto 1, 17 variables (wrapper - train)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	5	9	6	7	7	
Ejemplos erróneos (Entre/Test)	'4/6'	'2/6'	'2/6'	'1/4'	'2/4'	'0/3'
Error máximo conjunto test (dB)	4,41	4,89	4,65	5,13	7,15	4,65
Media del error (dB)	1,39	1,38	1,38	0,98	0,99	0,89
Desviación típica del error (dB)	0,67	0,68	0,68	0,72	0,79	0,66

Como era de esperar, al obtener este método un número de variables más alto, los resultados son mejores que para métodos anteriores. En cuanto al número de ejemplos erróneos, llega a ser nulo en entrenamiento para las salidas medias y de solamente 3 para test. También los errores máximos del conjunto de test son bajos, manteniéndose alrededor de 4-5 dB en casi todos los casos, valores similares a Relief. Por otro lado, la media del error es muy baja, cayendo por debajo de 1 en varias ocasiones, siendo su desviación típica también baja, entre 0,67 y 0,79.

Este método ofrece un conjunto de características bastante significativo, ya que obtiene resultados muy buenos aunque el número de variables es alto.

2. Tipo Wrapper con validación cruzada

Con el clasificador PART y el método de validación cruzada para 10 iteraciones, se obtienen los siguientes resultados, que ordenados de manera ascendente según el número de subconjuntos en los que es elegida cada variable, son los mostrados en la Tabla 90.

Para 10 iteraciones, se eligen las 7, 11 y 19 variables que en más iteraciones fueron seleccionadas. Los resultados de predecir el nivel de presión sonora con estas variables mediante la red neuronal son los mostrados en las Tablas 91 - 96. Aunque el punto de partida es en general de 10, 12 y 15 en este capítulo, como en este caso el método devuelve un conjunto de 7 variables, también se analiza.

Tabla 90. Resultados del método Wrapper con Validación Cruzada (10 iteraciones)

Posición	# variable	% subconjuntos
1	4	100
2	13	100
3	14	100
4	2	50
5	5	50
6	1	30
7	19	30
8	3	20
9	15	20
10	18	20
11	24	20
12	6	10
13	9	10
14	12	10
15	16	10
16	17	10
17	20	10
18	22	10
19	23	10
20	7	0
21	8	0
22	10	0
23	11	0
24	21	0
25	25	0

7 variables → 1,2,4,5,13,14,19

Tabla 91. Resultados de las 5 ejecuciones de la red con el conjunto 1, 7 variables (wrapper - cross10)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	12	19	16	10	11	
Ejemplos erróneos (Entre/Test)	'11/8'	'8/13'	'11/10'	'8/13'	'8/10'	'3/4'
Error máximo conjunto test (dB)	10,93	8,36	9,94	10,95	10,23	9,88
Media del error (dB)	2,11	2,30	2,05	2,16	2,10	1,96
Desviación típica del error (dB)	1,07	1,05	1,02	1,05	1,04	1,01

Tabla 92. Resultados de las 5 ejecuciones de la red con el conjunto 1, 7 variables (wrapper - cross10)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	10	17	7	17	13	
Ejemplos erróneos (Entre/Test)	'9/11'	'8/13'	'9/10'	'6/13'	'10/14'	'3/4'
Error máximo conjunto test (dB)	10,93	8,36	9,94	10,23	9,93	10,51
Media del error (dB)	2,11	2,16	2,12	2,32	2,23	1,99
Desviación típica del error (dB)	1,07	1,05	1,02	1,04	1,20	1,03

Para 7 variables, el número de ejemplos erróneos en la salida media es bajo, comparado con resultados anteriores con número de características similares. En cambio, el error máximo del conjunto de test, la media del error y la desviación típica aún devuelven valores algo altos, moviéndose el error máximo en torno a 10dB y la media del error alrededor de 2dB.

Cuando la red recibe 4 características más, estos valores no bajan considerablemente. Baján levemente el número de ejemplos erróneos y la media del error, pero el error máximo del conjunto de test sube 1dB. Por esto, se considera que el introducir estas 4 nuevas variables no ha aportado grandes mejoras.

11 variables→1,2,4,5,13,14,19,3,15,18,24

Tabla 93. Resultados de las 5 ejecuciones de la red con el conjunto 1,11 variables(wrapper – cross10)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	6	14	7	7	9	
Ejemplos erróneos (Entre/Test)	'9/10'	'9/10'	'8/12'	'5/12'	'3/13'	'1/3'
Error máximo conjunto test (dB)	10,27	9,04	11,17	12,10	11,27	11,06
Media del error (dB)	2,07	2,04	2,23	2,11	2,13	1,89
Desviación típica del error (dB)	1,00	0,99	1,07	1,07	1,06	1,00

Tabla 94. Resultados de las 5 ejecuciones de la red con el conjunto 1,11 variables(wrapper – cross10)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	8	8	8	6	5	
Ejemplos erróneos (Entre/Test)	'8/10'	'3/10'	'3/11'	'7/11'	'8/11'	'2/4'
Error máximo conjunto test (dB)	9,04	11,17	10,82	11,27	12,72	11,23
Media del error (dB)	2,08	1,94	2,09	2,12	2,10	1,86
Desviación típica del error (dB)	0,99	1,07	1,04	1,06	1,06	1,01

19 variables→1,2,4,5,13,14,19,3,15,18,24,6,9,12,16,17,20,22,23

Tabla 95. Resultados de las 5 ejecuciones de la red con el conjunto 1,19 variables(wrapper – cross10)

CONJUNTO 1(7 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	11	13	11	7	
Ejemplos erróneos (Entre/Test)	'1/7'	'2/6'	'3/6'	'1/6'	'3/6'	'0/2'
Error máximo conjunto test (dB)	3,54	3,41	4,21	5,09	4,00	4,05
Media del error (dB)	1,10	1,05	1,14	1,12	1,10	0,98
Desviación típica del error (dB)	0,68	0,67	0,69	0,70	0,69	0,66

Tabla 96. Resultados de las 5 ejecuciones de la red con el conjunto 1,19 variables(wrapper – cross10)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	9	9	6	8	9	
Ejemplos erróneos (Entre/Test)	'1/7'	'2/8'	'3/7'	'1/9'	'2/8'	'1/2'
Error máximo conjunto test (dB)	5,36	3,41	4,21	5,09	4,00	4,43
Media del error (dB)	1,09	1,17	1,20	1,09	1,21	0,98
Desviación típica del error (dB)	0,81	0,67	0,69	0,70	0,69	0,66

Al devolver este método el % de subconjuntos en los que cada variable ha sido incluida, se pasa de 11 variables a 19 drásticamente, puesto que las 8 variables en juego devuelven el mismo resultado. En el caso en que la red recibe 19 características de entrada, los resultados mejoran notablemente, como es de esperar por ser un alto número de variables. El error máximo del conjunto de test baja en alguna ejecución por debajo de 3dB, la media del error está en torno a 1dB y la desviación típica es también baja (entre 0,66 y 0,81dB).

En resumen, los resultados con 7 variables son bastante buenos, teniendo en cuenta el reducido número de características usado. En cambio, la adición de 4 variables no arroja resultados tan atractivos. El uso de 19 características implica un comportamiento muy bueno de la red.

Para 25 iteraciones (dado que hay 25 variables, esto sería equivalente al método leave-one-out), se obtuvieron unos resultados que, ordenados de manera ascendente según el número de subconjuntos en los que es elegida cada variable, son los que se observan en la Tabla 97.

Para 25 iteraciones, se eligen las 10, 13 y 15 variables que en más iteraciones fueron seleccionadas. Los resultados de predecir el nivel de presión sonora con estas variables mediante la red neuronal son los mostrados en las Tablas 98 -103.

Observando las Tablas 98 -103, se ve cómo la red se comporta de forma muy similar para los 3 casos (10, 12 y 15 variables). Los errores máximos se mantienen en casi todas las ejecuciones en torno a 5 dB, la media del error en torno a 1dB y la desviación típica se mueve en rangos similares (entre 0,7 dB y 0,9 dB). Este es un claro ejemplo de cómo el añadir más variables no necesariamente implica una mejoría en el comportamiento de la red.

El comportamiento de la red en este método es relativamente bueno, pero hasta ahora se han obtenido resultados mejores con otros métodos.

Tabla 97. Resultados del método Wrapper con Validación Cruzada (25 iteraciones)

Posición	# variable	% subconjuntos
1	4	100
2	13	100
3	14	100
4	5	44
5	10	36
6	1	32
7	21	32
8	12	28
9	3	24
10	19	24
11	2	20
12	16	20
13	18	20
14	17	16
15	22	16
16	6	12
17	23	12
18	9	8
19	20	8
20	24	8
21	11	4
22	7	0
23	8	0
24	15	0
25	25	0

10 variables→4,13,14,5,10,1,21,12,3,19

Tabla 98. Resultados de las 5 ejecuciones de la red con el conjunto 1,10 variables(wrapper - cross25)

CONJUNTO 1(7NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	22	13	12	13	22	
Ejemplos erróneos (Entre/Test)	'2/8'	'2/9'	'3/7'	'3/8'	'3/7'	'1/2'
Error máximo conjunto test (dB)	4,72	5,15	5,31	5,27	5,81	5,17
Media del error (dB)	1,15	1,23	1,13	1,18	1,15	1,06
Desviación típica del error (dB)	0,70	0,74	0,70	0,71	0,72	0,68

Tabla 99. Resultados de las 5 ejecuciones de la red con el conjunto 1,10 variables(wrapper - cross25)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	19	12	9	17	10	
Ejemplos erróneos (Entre/Test)	'1/8'	'1/7'	'2/9'	'2/10'	'1/9'	'1/3'
Error máximo conjunto test (dB)	4,72	5,31	5,22	5,06	5,81	5,25
Media del error (dB)	1,17	1,13	1,15	1,39	1,18	1,15
Desviación típica del error (dB)	0,70	0,70	0,73	0,73	0,72	0,73

13 variables→4,13,14,5,10,1,21,12,3,19,2,16,18

Tabla 100. Resultados de las 5 ejecuciones de la red con el conjunto 1,13 variables(wrapper - cross25)

CONJUNTO 1(7NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	10	9	6	8	6	
Ejemplos erróneos (Entre/Test)	'4/6'	'3/9'	'3/7'	'5/7'	'2/9'	'0/2'
Error máximo conjunto test (dB)	4,48	5,23	5,60	5,38	5,23	5,07
Media del error (dB)	1,22	1,33	1,26	1,29	1,35	1,14
Desviación típica del error (dB)	0,71	0,75	0,73	0,76	0,78	0,70

Tabla 101. Resultados de las 5 ejecuciones de la red con el conjunto 1,13 variables(wrapper - cross25)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	6	6	10	6	7	
Ejemplos erróneos (Entre/Test)	'4/9'	'1/7'	'1/5'	'5/7'	'2/7'	'0/2'
Error máximo conjunto test (dB)	7,71	5,23	5,60	5,38	5,23	5,10
Media del error (dB)	1,24	1,19	1,11	1,32	1,26	1,12
Desviación típica del error (dB)	0,82	0,75	0,73	0,76	0,78	0,70

15 variables→4,13,14,5,10,1,21,12,3,19,2,16,18,17,22

Tabla 102. Resultados de las 5 ejecuciones de la red con el conjunto 1,15 variables(wrapper - cross25)

CONJUNTO 1(7NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	9	8	22	16	17	
Ejemplos erróneos (Entre/Test)	'2/10'	'1/6'	'5/7'	'3/7'	'2/8'	'0/2'
Error máximo conjunto test (dB)	5,37	4,93	4,67	6,22	5,06	5,01
Media del error (dB)	1,27	1,15	1,22	1,28	1,12	1,09
Desviación típica del error (dB)	0,75	0,71	0,72	0,78	0,69	0,69

Tabla 103. Resultados de las 5 ejecuciones de la red con el conjunto 1,15 variables(wrapper - cross25)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	8	5	7	6	9	
Ejemplos erróneos (Entre/Test)	'1/8'	'2/7'	'1/11'	'2/10'	'1/7'	'1/2'
Error máximo conjunto test (dB)	5,37	4,67	5,06	9,34	8,89	5,18
Media del error (dB)	1,11	1,10	1,19	1,23	1,19	1,05
Desviación típica del error (dB)	0,75	0,72	0,69	0,87	0,91	0,69

4.3 RESUMEN DE RESULTADOS

En las Tablas 104 y 105 se resumen los resultados de cada uno de los métodos de selección de características evaluados en las secciones anteriores. A partir de estas tablas se deducirán una serie de variables del entorno descriptivas del paisaje sonoro que, como se ha venido haciendo hasta ahora, se probarán en la red neuronal con objeto de tomar una decisión final acerca de qué características reducen el conjunto inicial de forma aceptable.

A la vista de las Tablas 104 y 105, se han extraído un grupo de características diferentes a las elegidas por un único método. Este grupo de variables se caracteriza por haber dado resultados positivos en uno o varias de las pruebas que se acaban de realizar, así como en la observación de un significado característico en el paisaje sonoro. Cuando ha sido necesario se ha recurrido a la opinión de expertos en ruido ambiental, para considerar la inclusión (o no) de alguna variable sobre la que se tenía duda. Estas características son las relacionadas en la Tabla 106. Se forman así cuatro grupos de características:

- El grupo de variables A, compuesto por 10 características, será el grupo con el que inicialmente se probará la red neuronal.
- El grupo de variables B, compuesto por dos variables, que, unido a las del grupo A, será testeado con la red neuronal en segundo lugar.
- El grupo de variables C, que contiene tres características y que junto con los grupos A y B constituirá el conjunto de características con el que se probará la red neuronal en último lugar.
- El grupo de variables D, que representará a las variables descartadas, entendidas como las menos relevantes.

Para analizar el comportamiento de estos conjuntos de variables de entrada que se han seleccionado, se consideran de nuevo los cinco conjuntos de entrenamiento utilizados en el capítulo III. Esto permite una comparación adecuada.

Los resultados que devuelve la red recibiendo como entrada los datos de las características elegidas se muestran en las Tablas 107 - 121. Estos resultados son los correspondientes a la ejecución de la red neuronal con 9 neuronas en la capa oculta. En el anexo puede encontrarse los resultados con 7 neuronas.

Tabla 104. Resumen resultados de los métodos de Selección de Características evaluados

Método	Puesto																									N° variables							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	HN	7	10	11	12	13	15	19
RELIEF	14	3	13	12	5	2	4	16	21	20	7	17	9	19	25	15	18	10	22	24	23	1	6	11	8	7		2		2		2	
N° ejemplos erróneos test																										9		3		2		2	
Media error																										7		1,27		0,99		0,93	
Ganancia información	24	4	25	2	16	23	8	5	10	6	11	21	17	15	13	18	9	19	7	20	2	14	1	12	3	7	7	7		6		6	
N° ejemplos erróneos test																										9		6		6		5	
Media error																										7		2,69		2,46		2,06	
mRMR - MID	19	21	16	17	18	14	20	22	23	24	6	8	4	5	7											7		6		6		2	
N° ejemplos erróneos test																										9		6		5		3	
Media error																										7		2,60		2,18		1,78	
mRMR - MIQ	19	21	16	17	18	14	20	13	22	23	24	15	1	6	8											7		6		6		5	
N° ejemplos erróneos test																										9		6		6		5	
Media error																										7		2,60				2,12	
Validación Cruzada 10 iter	4	13	14	2	5	1	19	3	15	18	24	6	9	12	16	17	20	22	23							7	4		3				2
N° ejemplos erróneos test																										9	4		4				2
Media error																										7	1,96		1,89				0,98
Validación Cruzada 25 iter	4	13	14	5	10	1	21	12	3	19	2	16	18	17	22											7		2			2	2	
N° ejemplos erróneos test																										9		3			2	2	
Media error																										7		1,06			1,14	1,09	
																										9		1,15			1,12	1,05	

Ranking de variables

Tabla 105. Resumen resultados de los métodos de Selección de Características evaluados

Método	Subconjunto de Variables																	Nº variables						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	HN	7	9	10	11	14	17
Correlación - 9 variables	5	8	10	11	12	13	14	16	17										7	5				
Nº errores > 5% test																			9	4				
Media error																			7	1,80				
																			9	1,73				
10 variables	5	8	10	11	12	13	14	16	17	19									7		2			
Nº errores > 5% test																			9		2			
Media error																			7		1,06			
																			9		1,07			
11 variables	4	5	8	10	11	12	13	14	16	17	19								7			2		
Nº errores > 5% test																			9			2		
Media error																			7			1,23		
																			9			1,06		
Consistencia - 9 variables	5	6	8	9	10	11	13	14	24										7	4				
Nº errores > 5% test																			9	5				
Media error																			7	1,84				
																			9	1,70				
10 variables	5	6	8	9	10	11	13	14	17	24									7		3			
Nº errores > 5% test																			9		3			
Media error																			7		1,63			
																			9		1,44			
10 variables	2	5	6	8	9	10	11	13	14	24									7		3			
Nº errores > 5% test																			9		3			
Media error																			7		1,55			
																			9		1,54			
11 variables	5	7	8	9	10	11	13	14	18	23	25								7			3		
Nº errores > 5% test																			9			2		
Media error																			7			1,80		
																			9			1,65		
14 variables	3	5	6	8	9	10	11	12	13	14	15	16	19	20									2	
Nº errores > 5% test																			9				2	
Media error																			7				0,80	
																			9				0,82	
Wrapper Red Neuronal	1	2	3	5	6	7	10	11	12	13	14	15	17	21	23	24	25	7					3	
Nº errores > 5% test																			9					3
Media error																			7					0,95
																			9					0,89

Tabla 106. Agrupación de variables según los resultados del capítulo IV

A	3	PRESENCIA OBRAS CONSTRUCCIÓN
A	4	TIEMPO DE ESTABILIZACIÓN
A	5	TIPO DE FLUJO
A	8	CAUDAL DE VEHÍCULOS PESADOS ASCENDENTE
A	10	CAUDAL DE MOTOCICLETAS ASCENDENTE
A	12	APARICIÓN DE VEHÍCULOS CON SIRENA
A	13	APARICIÓN ESA RELACIONADOS CON EL TRÁFICO
A	14	APARICIÓN ESA NO RELACIONADOS CON EL TRÁFICO
A	16	PENDIENTE (%)
A	17	CARRILES PENDIENTE ASCENDENTE
B	6	CAUDAL DE VEHÍCULOS LIGEROS ASCENDENTE
B	19	TIPO DE PAVIMENTO
C	2	ENTORNO COMERCIAL/OCIO
C	11	CAUDAL DE MOTOCICLETAS DESCENDENTE
C	20	ESTADO DEL FIRME
D	1	PERÍODO DEL DÍA
D	7	CAUDAL DE VEHÍCULOS LIGEROS DESCENDENTE
D	9	CAUDAL DE VEHÍCULOS PESADOS DESCENDENTE
D	15	VELOCIDAD DE CIRCULACIÓN PROMEDIO
D	18	CARRILES PENDIENTE DESCENDENTE
D	21	TIPO DE VÍA
D	22	ANCHURA DE LA VÍA
D	23	ALTURA DE LA VÍA
D	24	ANCHURA DE LA CALZADA
D	25	DISTANCIA AL RECEPTOR

10 variables→3, 4, 5, 8, 10, 12, 13, 14, 16, 17

Tabla 107. Resultados de las 5 ejecuciones de la red con el conjunto 1,10 variables (selección propia)

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	12	10	7	18	15	
Ejemplos erróneos (Entre/Test)	'4/8'	'5/8'	'6/10'	'6/10'	'6/9'	'2/2'
Error máximo conjunto test (dB)	9,47	3,65	6,15	5,05	3,76	5,26
Media del error (dB)	1,33	1,30	1,56	1,54	1,35	1,30
Desviación típica del error (dB)	0,93	0,71	0,80	0,76	0,73	0,78

Tabla 108. Resultados de las 5 ejecuciones de la red con el conjunto 2,10 variables (selección propia)

CONJUNTO 2 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	13	9	9	15	18	
Ejemplos erróneos (Entre/Test)	'5/7'	'4/8'	'5/10'	'4/12'	'4/8'	'2/2'
Error máximo conjunto test (dB)	6,37	5,25	5,75	4,52	4,39	5,17
Media del error (dB)	1,34	1,45	1,51	1,55	1,36	1,26
Desviación típica del error (dB)	0,79	0,83	0,82	0,72	0,74	0,76

Tabla 109. Resultados de las 5 ejecuciones de la red con el conjunto 3,10 variables (selección propia)

CONJUNTO 3 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	11	18	11	18	14	
Ejemplos erróneos (Entre/Test)	'4/9'	'4/9'	'3/4'	'5/9'	'5/7'	'1/2'
Error máximo conjunto test (dB)	5,81	6,06	6,14	5,18	5,06	5,38
Media del error (dB)	1,30	1,46	1,06	1,44	1,19	1,15
Desviación típica del error (dB)	0,77	0,81	0,80	0,77	0,77	0,72

Tabla 110. Resultados de las 5 ejecuciones de la red con el conjunto 4,10 variables (selección propia)

CONJUNTO 4 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	8	22	19	15	16	
Ejemplos erróneos (Entre/Test)	'7/6'	'5/7'	'6/10'	'6/9'	'6/8'	'2/1'
Error máximo conjunto test (dB)	3,48	4,21	3,74	3,55	3,88	3,55
Media del error (dB)	1,25	1,23	1,31	1,24	1,25	1,17
Desviación típica del error (dB)	0,76	0,77	0,71	0,76	0,73	0,71

Tabla 111. Resultados de las 5 ejecuciones de la red con el conjunto 5,10 variables (selección propia)

CONJUNTO 5 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	22	14	15	13	7	
Ejemplos erróneos (Entre/Test)	'4/11'	'3/10'	'2/12'	'4/10'	'5/14'	'0/3'
Error máximo conjunto test (dB)	5,38	5,58	6,35	11,41	6,32	5,90
Media del error (dB)	1,43	1,41	1,39	1,35	1,55	1,34
Desviación típica del error (dB)	0,78	0,79	0,82	1,00	0,81	0,79

12 variables→3, 4, 5, 8, 10, 12, 13, 14, 16, 17, 6, 19**Tabla 112. Resultados de las 5 ejecuciones de la red con el conjunto 1,12 variables (selección propia)**

CONJUNTO 1(9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	9	9	8	9	
Ejemplos erróneos (Entre/Test)	'5/8'	'5/6'	'4/9'	'5/8'	'7/4'	'2/0'
Error máximo conjunto test (dB)	3,54	8,49	4,20	6,11	5,98	3,54
Media del error (dB)	1,14	1,17	1,27	1,22	1,09	0,99
Desviación típica del error (dB)	0,66	0,84	0,75	0,84	0,76	0,67

Tabla 113. Resultados de las 5 ejecuciones de la red con el conjunto 2,12 variables (selección propia)

<i>CONJUNTO 2 (9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	5	7	9	11	7	
Ejemplos erróneos (Entre/Test)	'5/11'	'5/11'	'5/9'	'5/7'	'4/10'	'2/2'
Error máximo conjunto test (dB)	5,40	5,20	5,45	5,49	4,69	5,05
Media del error (dB)	1,37	1,33	1,35	1,25	1,29	1,13
Desviación típica del error (dB)	0,77	0,83	0,76	0,74	0,78	0,73

Tabla 114. Resultados de las 5 ejecuciones de la red con el conjunto 3,12 variables (selección propia)

<i>CONJUNTO 3 (9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	9	16	9	11	18	
Ejemplos erróneos (Entre/Test)	'5/7'	'4/8'	'4/10'	'5/7'	'5/5'	'1/2'
Error máximo conjunto test (dB)	6,24	5,64	5,33	5,01	4,90	5,20
Media del error (dB)	1,14	1,20	1,36	1,30	1,13	1,12
Desviación típica del error (dB)	0,82	0,74	0,82	0,71	0,77	0,72

Tabla 115. Resultados de las 5 ejecuciones de la red con el conjunto 4,12 variables (selección propia)

<i>CONJUNTO 4 (9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	12	8	8	9	8	
Ejemplos erróneos (Entre/Test)	'6/9'	'6/6'	'5/8'	'6/6'	'7/7'	'2/1'
Error máximo conjunto test (dB)	3,58	3,96	3,76	3,45	4,02	3,40
Media del error (dB)	1,23	1,22	1,04	1,23	1,19	1,05
Desviación típica del error (dB)	0,72	0,75	0,70	0,72	0,75	0,69

Tabla 116. Resultados de las 5 ejecuciones de la red con el conjunto 5,12 variables (selección propia)

<i>CONJUNTO 5 (9 NEURONAS EN LA CAPA OCULTA)</i>	<i>Ejecuc.1</i>	<i>Ejecuc.2</i>	<i>Ejecuc.3</i>	<i>Ejecuc.4</i>	<i>Ejecuc.5</i>	<i>Media</i>
Épocas	9	17	7	6	7	
Ejemplos erróneos (Entre/Test)	'5/10'	'3/12'	'3/11'	'7/9'	'3/8'	'0/2'
Error máximo conjunto test (dB)	5,03	6,59	9,82	6,95	5,11	5,32
Media del error (dB)	1,34	1,53	1,45	1,33	1,24	1,19
Desviación típica del error (dB)	0,75	0,85	0,91	0,84	0,77	0,74

15 variables → 3, 4, 5, 8, 10, 12, 13, 14, 16, 17, 6, 19, 2, 11, 20

Tabla 117. Resultados de las 5 ejecuciones de la red con el conjunto 1,15 variables (selección propia)

CONJUNTO 1 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	6	9	6	11	7	
Ejemplos erróneos (Entre/Test)	'3/5'	'3/4'	'3/5'	'4/3'	'3/7'	'2/0'
Error máximo conjunto test (dB)	3,47	13,55	4,48	4,93	9,01	3,65
Media del error (dB)	1,00	0,93	1,03	0,88	1,14	0,84
Desviación típica del error (dB)	0,63	1,00	0,67	0,69	0,79	0,61

Tabla 118. Resultados de las 5 ejecuciones de la red con el conjunto 2,15 variables (selección propia)

CONJUNTO 2 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	6	9	5	7	
Ejemplos erróneos (Entre/Test)	'4/4'	'5/6'	'4/4'	'3/5'	'3/2'	'1/1'
Error máximo conjunto test (dB)	4,82	4,15	4,65	5,14	5,60	4,90
Media del error (dB)	1,12	1,10	1,09	1,09	1,10	0,95
Desviación típica del error (dB)	0,73	0,66	0,71	0,69	0,73	0,66

Tabla 119. Resultados de las 5 ejecuciones de la red con el conjunto 3,15 variables (selección propia)

CONJUNTO 2 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	6	7	8	7	11	0
Ejemplos erróneos (Entre/Test)	'3/5'	'3/5'	'4/3'	'3/5'	'4/3'	'1/2'
Error máximo conjunto test (dB)	4,09	6,32	5,26	5,86	5,10	5,11
Media del error (dB)	1,14	1,10	1,04	1,09	1,04	0,92
Desviación típica del error (dB)	0,71	0,73	0,71	0,72	0,67	0,64

Tabla 120. Resultados de las 5 ejecuciones de la red con el conjunto 4,15 variables (selección propia)

CONJUNTO 4 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	6	8	6	5	6	
Ejemplos erróneos (Entre/Test)	'3/6'	'4/6'	'4/5'	'3/5'	'4/6'	'2/0'
Error máximo conjunto test (dB)	3,41	3,94	3,76	3,70	3,47	3,46
Media del error (dB)	1,15	1,21	1,20	1,11	1,20	0,96
Desviación típica del error (dB)	0,66	0,67	0,66	0,67	0,66	0,64

Observando la evolución del comportamiento de la red neuronal conforme se añaden características, se detecta que mejora mucho en cada paso.

Tabla 121. Resultados de las 5 ejecuciones de la red con el conjunto 5,15 variables (selección propia)

CONJUNTO 5 (9 NEURONAS EN LA CAPA OCULTA)	Ejecuc.1	Ejecuc.2	Ejecuc.3	Ejecuc.4	Ejecuc.5	Media
Épocas	7	6	6	4	7	
Ejemplos erróneos (Entre/Test)	'3/5'	'2/8'	'3/5'	'2/7'	'2/9'	'0/3'
Error máximo conjunto test (dB)	4,71	5,25	5,26	5,69	5,25	5,26
Media del error (dB)	1,30	1,24	1,24	1,17	1,28	0,94
Desviación típica del error (dB)	0,71	0,75	0,72	0,75	0,74	0,66

El número de ejemplos erróneos baja, aunque en el caso de las salidas medias, se mantiene en torno a 0-2 para el entrenamiento y 0-3 para test. En general, el conjunto de test se predice peor en todas las ejecuciones que implican al conjunto 5, seguramente porque algún ejemplo más complicado de predecir (o cuyas características no se han aprendido al entrenar) se ha incluido en el conjunto de test.

Para las salidas medias, el error máximo en el conjunto de test cae, en el conjunto 1 casi 2dB al pasar de 10 a 12 características, aunque se mantiene en cifras similares al anterior en 15 características. En el resto de conjuntos se mantiene cercano a 5db excepto en el conjunto 4, que se mueve siempre alrededor de 3,5 dB. En los casos en los que se baja de 4dB se puede considerar como positivo, ya que con los métodos expuestos anteriormente en pocas ocasiones se ha bajado de 4dB.

La media del error baja en menor medida, pero considerando que se trata de un valor cuadrático medio, ésta bajada es remarcable, puesto que va de 1,34dB a 0,84dB. También la desviación típica del error se reduce, llegando a 0,61dB para 15 variables con el conjunto 1. Sólo en el método de evaluación de la consistencia de las características se obtienen valores similares con 14 variables.

Los resultados obtenidos con las variables elegidas entre los resultados obtenidos han sido bastante buenos, pudiendo elegir indistintamente como conjunto final de reducción de características el obtenido para el método de evaluación de la consistencia de variables de 14 características (3,5,6,8,9,10,11,12,13,14,15,16,19,20) o bien el conjunto de 15 características evaluado en este apartado (2,3,4,5,6,8,10,11,12,13,14,16,17,19,20).

Los resultados tan sorprendentes del capítulo III son también obtenidos, según se ha visto durante todo este capítulo, tras el proceso de Selección de Características., a lo que podría ayudar, sin duda, el que los datos de los que se dispone sean demasiado homogéneos. En el siguiente capítulo se tratará de analizar, con un nuevo conjunto de datos, que no han sido tomados específicamente para este proceso, si efectivamente las RNA son capaces de predecir el ruido urbano, al margen de un posible exceso de homogeneidad en los datos iniciales.

Tabla 122. Tabla de variables con variables elegidas en negrita

	Nombre variable
1	Periodo del día
2	Entorno comercial o de ocio
3	Presencia de obras
4	Tiempo de estabilización
5	Tipo de flujo de trafico
6	Flujo ascendente de vehículos ligeros
7	Flujo descendente de vehículos ligeros
8	Flujo ascendente de vehículos pesados
9	Flujo descendente de vehículos pesados
10	Flujo ascendente de motos
11	Flujo descendente de motos
12	Número de vehículos con sirena
13	Eventos relacionados con tráfico
14	Eventos no relacionados con tráfico
15	Velocidad de circulación promedio
16	Pendiente
17	Número de carriles ascendente
18	Número de carriles descendente
19	Tipo de pavimento
20	Estado del firme
21	Tipo de vía
22	Anchura de la vía
23	Altura de la vía
24	Anchura de la calzada
25	Distancia al receptor

CAPÍTULO V: VALIDACIÓN DEL MODELO CON NUEVOS CONJUNTOS DE DATOS

Los resultados obtenidos en el capítulo anterior hacen necesario, o al menos conveniente, una nueva validación del modelo, ya que es posible que los muy buenos resultados, obtenidos con pocas variables, se deban en parte a que los datos utilizados en los entrenamientos de las RNA sean demasiados homogéneos en algunas de las variables.

Por ejemplo, las variables “número de vehículos con sirena” o “presencia de obras” tenían, en la gran mayoría de los casos (98% y 90%, respectivamente) el mismo valor, con lo que eliminar estas variables en la selección de características, es posible que no haya tenido consecuencias para el almacenamiento distribuido de la información en los pesos. Entonces, para la RNA no ha tenido gran importancia la eliminación de estas variables.

De igual forma ha podido pasar con otras variables, donde la variabilidad de los datos disponibles es escasa. Es el caso del caudal de motocicletas tanto ascendente como descendente, que se mantiene igual o por debajo de 1 en el 70% de los casos.

Por ello, es necesario, o al menos muy conveniente, poner de manifiesto el comportamiento del modelo de la RNA, utilizando datos donde la variabilidad de los valores sea clara y más abundante.

Los mismos expertos⁴ que proporcionaron los datos usados en los capítulos anteriores, proporcionaron también nuevos datos, que podían ser usados y cuya variabilidad estaba mucho más clara. Estos nuevos datos se tomaron con otros objetivos, y

⁴ Se le agradece al Dr. Ruiz Padillo de la UGR y a su grupo de investigación.

tuvieron que sufrir un proceso de adaptación para que pudieran ser tratados conjuntamente con los datos iniciales.

Este preprocesamiento inicial consistió en dar formato, normalizar y dar dimensiones similares a las de los datos iniciales con los que se desarrolló el modelo propuesto.

Entre los tratamientos de los datos que fue necesario realizar cabe destacar el proceso de normalización, dado que los límites superiores superaban aquellos definidos para los datos anteriores. A modo de ejemplo, el caudal descendente de vehículos ligeros excedía el límite superior de la normalización en los datos anteriores, siendo antes 80 vehículos por hora y en los datos nuevos se disponía de datos de hasta 160 vehículos a la hora.

Por otro lado, también en las variables en las que se establecen categorías (como son el estado del firme, el tipo de flujo, el tipo de pavimento y el tipo de vía), aparecen ejemplos en los nuevos datos en categorías en las que antes no se disponía de ningún dato. Ya que dichas categorías estaban contempladas, esto no ha supuesto un problema.

Asimismo, existían variables en los datos nuevos que no aparecían anteriormente. Estas variables han sido ignoradas para estas pruebas, aunque serán de utilidad para futuros experimentos. Ejemplos de estas variables son: presencia de vegetación, presencia de fuentes de agua o caudal de vehículos especiales, que se ha añadido a vehículos pesados.

De la misma forma, variables como el tiempo de estabilización no son proporcionadas en este nuevo conjunto de datos. En este caso, se ha tomado un valor constante previamente establecido, 25 minutos, ya que era el tiempo medio de los datos iniciales.

Una primera aproximación a la evaluación de los nuevos datos, una vez procesados, fue ejecutar con los mismos la red neuronal ya entrenada con los datos anteriores. Estas pruebas arrojaron unos resultados poco convincentes y fuera de la línea de los resultados obtenidos en el capítulo III. Esto llevó a deducir que los datos nuevos eran en exceso diferentes de los anteriores y ratificaba la posibilidad de que los datos iniciales podían ser demasiado homogéneos. Con estos resultados en la mano, se consideró que la mejor opción para sacar partido de los nuevos datos sería unirlos a los datos iniciales y tratar todo como un único conjunto.

Una vez realizado el preprocesamiento descrito, los experimentos llevados a cabo son de utilidad para verificar la bondad del modelo, es decir, si las RNA puede ser una metodología válida para la predicción del ruido ambiental urbano.

A su vez, la unión de los datos iniciales y los datos posteriores aportarán información sobre las variables tras un proceso de selección de características similar al desarrollado en el capítulo IV. Estos resultados son comparados con los obtenidos anteriormente y con los modelos matemáticos “clásicos” que se usan actualmente.

5.1 RED NEURONAL ARTIFICIAL PARA LA TOTALIDAD DE LOS DATOS DISPONIBLES

El número de ejemplos con los que inicialmente se contaba era de 289. El conjunto de datos proporcionado posteriormente incluía 367 instancias, teniendo así, un total de 656 ejemplos. Los datos se normalizaron todos juntos.

De la misma forma que con el conjunto inicial de datos, este nuevo conjunto se divide en conjunto de entrenamiento y test. El nuevo conjunto de entrenamiento estará compuesto por 450 ejemplos, y el de test, por los restantes 206. Esta división de las instancias en dos conjuntos se hace, aleatoriamente, de cinco formas diferentes, con lo que se dispondrá de nuevo de cinco conjuntos de datos diferentes.

Las pruebas con estos cinco conjuntos de datos se han realizado con distinto número de neuronas en la capa oculta, concretamente con 7, 9, 12 y 15. Para simplificar y agilizar la presentación, sólo se muestran los resultados de la media de las cinco ejecuciones realizadas con cada conjunto de entrenamiento, ya que la media indica unos resultados mejores y más realistas. Esto se seguirá a lo largo de todo el capítulo.

Se muestran en las Tablas 123 - 126 los resultados para la salida media después de las cinco ejecuciones para cada uno de los cinco conjuntos. Los resultados completos se incluyen en los anexos.

Tabla 123. Resultados de las salidas medias de la red, para los cinco conjuntos, 7HN

7 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'18/15'	'26/13'	'24/18'	'15/17'	'22/16'
Error máximo conjunto test (dB)	12,58	11,83	10,76	9,46	7
Media del error (dB)	2,92	2,94	3,33	2,98	3
Desviación típica del error (dB)	1,27	1,23	1,34	1,26	1,23

Tabla 124. Resultados de las salidas medias de la red, para los cinco conjuntos, 9HN

9 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'12/10'	'16/8'	'19/18'	'15/21'	'14/13'
Error máximo conjunto test (dB)	13,84	13,24	12,95	10,32	7,70
Media del error (dB)	2,48	2,38	3,12	3	2,46
Desviación típica del error (dB)	1,16	1,11	1,33	1,28	1,15

El número de ejemplos erróneos es más alto con 7 neuronas ocultas. Para 9,12 y 15 es similar con los conjuntos 1 y 2, y en los otros conjuntos se reduce para 12 y 15 neuronas.

El error máximo en el conjunto de test es similar en los cuatro casos, aunque ligeramente menor para 7 neuronas ocultas. En el conjunto 5 se obtiene claramente

siempre un error máximo bastante más bajo que con los otros conjuntos, por la coincidencia de uno o más ejemplos "difíciles de aprender" en los ejemplos de entrenamiento, y que están como ejemplos de test en los otros conjuntos considerados.

Tabla 125. Resultados de las salidas medias de la red, para los cinco conjuntos, 12HN

12 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'10/12'	'13/9'	'12/11'	'10/17'	'11/14'
Error máximo conjunto test (dB)	13,82	14,01	11,15	9,81	7,96
Media del error (dB)	2,40	2,29	2,54	2,51	2,33
Desviación típica del error (dB)	1,16	1,11	1,20	1,19	1,14

Tabla 126. Resultados de las salidas medias de la red, para los cinco conjuntos, 15HN

15 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'9/12'	'12/7'	'10/11'	'11/16'	'9/16'
Error máximo conjunto test (dB)	13,62	14,02	15,08	12,07	8,80
Media del error (dB)	2,18	2,15	2,48	2,62	2,27
Desviación típica del error (dB)	1,11	1,08	1,23	1,25	1,15

La media del error es también menor al aumentar el número de neuronas en la capa oculta, bajando también, aunque más ligeramente, la desviación típica del error. En general, es con el conjunto 2 con el que la red se comporta mejor, sobre todo en los ejemplos erróneos y la media del error. Globalmente, el mejor resultado se obtiene para el conjunto 2 con 15 neuronas, donde el número de errores es menor al 3,5% tanto en entrenamiento (2,7%) como en test (3,4%).

Todos los resultados con 12 y 15 neuronas dan un número de errores más que aceptable, por debajo del 3% en los ejemplos de entrenamiento y para los ejemplos de test, por debajo del 8%, aunque baja al 6% en la mayoría de los casos (7 de 10).

Los resultados obtenidos con el conjunto de datos inicial (presentados en el capítulo III y en las Tablas 30 - 34), son mejores que estos nuevos, pero estos siguen siendo suficientemente buenos. Esto ocurre, incluso con el aumento del número de neuronas en la capa oculta, debido al incremento del número de ejemplos a "aprender". Es de destacar que el conjunto de datos ahora es mayor, y de carácter más heterogéneo, con lo que la nueva red neuronal entrenada estaría preparada para un número más amplio de situaciones.

En las Figuras 70 y 71 se comparan los resultados de la red neuronal con los obtenidos con los modelos matemáticos con este nuevo conjunto de datos. Aquí también puede verse como la red neuronal se sigue comportando bastante mejor que los modelos clásicos.

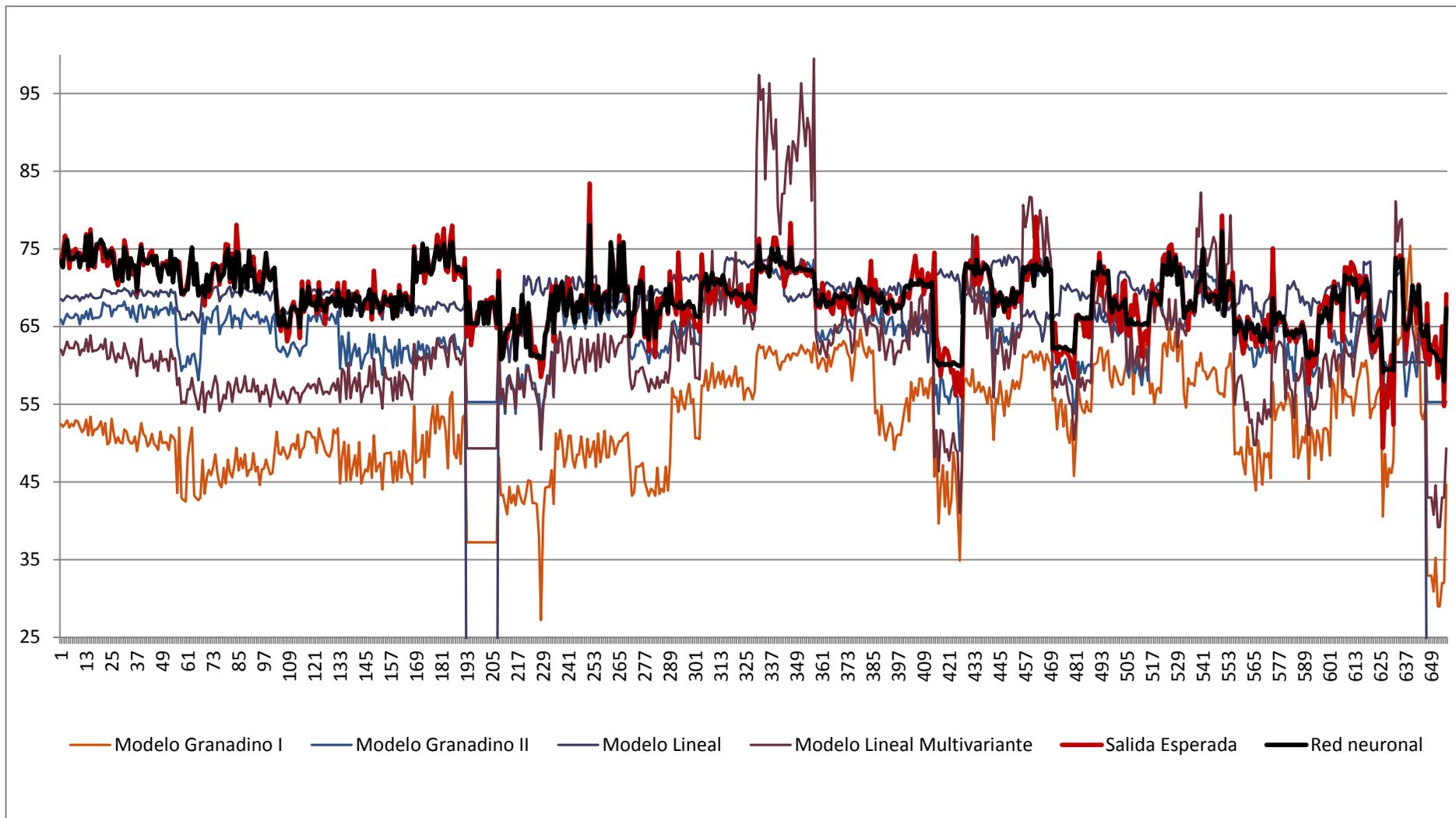


Figura 70. Gráfica comparativa de los resultados de los modelos clásicos y la red neuronal con el valor L_{Aeq} medido para el conjunto de datos 1 (a)

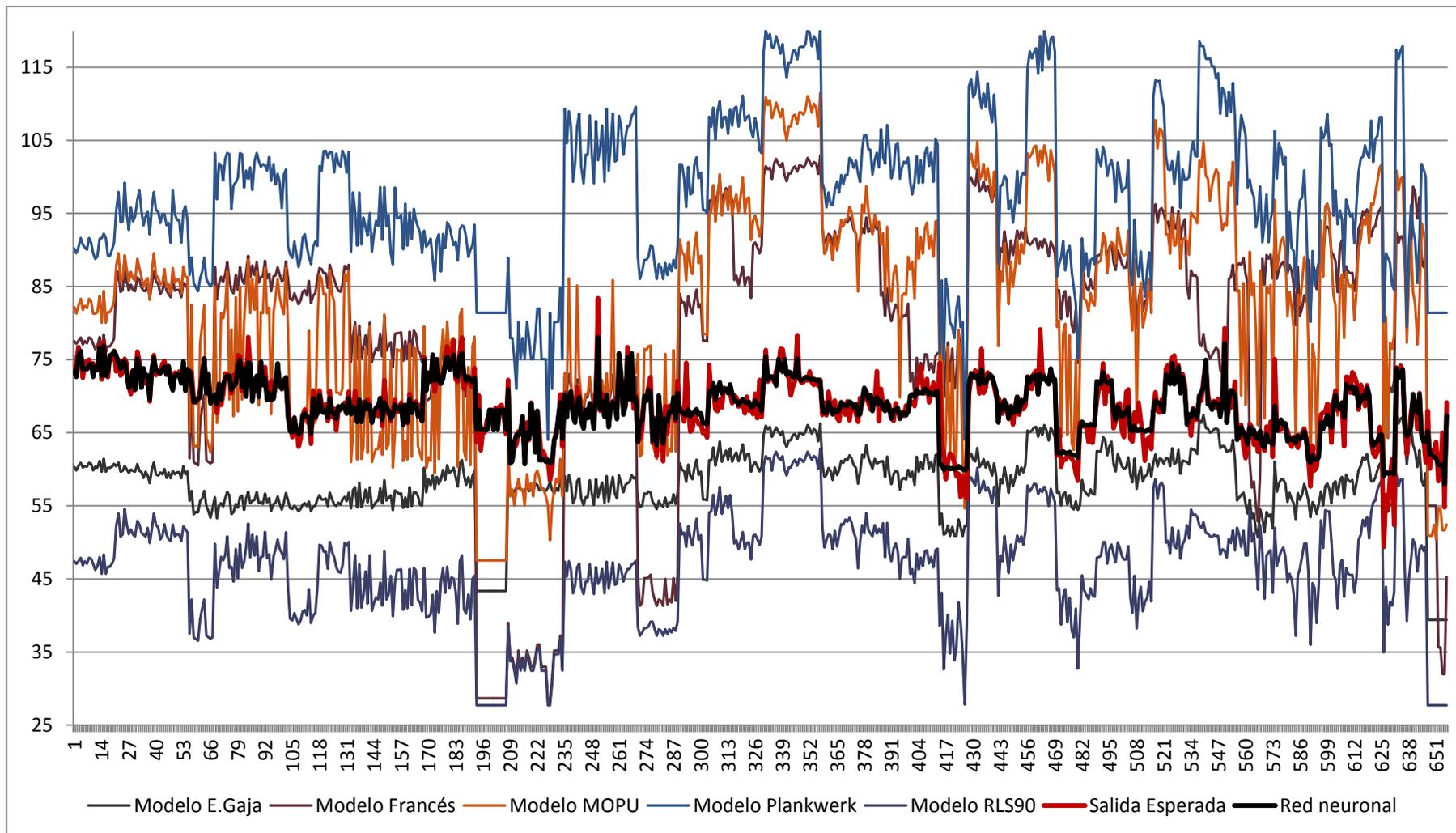


Figura 71. Gráfica comparativa de los resultados de los modelos clásicos y la red neuronal con el valor L_{Aeq} medido para el conjunto de datos 1 (b)

5.2 SELECCIÓN DE CARACTERÍSTICAS CON EL NUEVO CONJUNTO DE DATOS

En esta sección se pretende hacer un estudio de selección de características similar al capítulo IV, pero sobre los resultados obtenidos en la sección previa. Esto va a llevarse a cabo mediante la aplicación de varios de los métodos de selección de características mencionados, concretamente se han elegido aquellos cuyas variables elegidas ofrecieron un mejor resultado con la red neuronal. Al igual que en el capítulo IV, para estos métodos se mostrarán los resultados para el conjunto 1.

MÉTODO 1: RELIEF-F

Se ejecutó este método de la misma forma que en el capítulo IV, obteniendo un ranking de variables según los vecinos más cercanos (Tabla 127). Cada variable fue entonces ordenada según la media de sus posiciones en el ranking, y posteriormente comparada con este mismo resultado en el caso de los datos de partida, obteniendo la Tabla 128.

Para poder comparar mejor los resultados de este método, que arrojó los mejores resultados con los datos iniciales, se ha construido la Tabla 129, en la que se observan diferencias notables en cuanto a las variables elegidas. Sólo coinciden, en el mejor de los casos, un 66% de las variables.

Debido a que los resultados son tan diferentes del mismo método aplicado a los datos iniciales y con objeto de hacer una comparación más profunda, se ejecuta la red neuronal con las variables elegidas por Relief-F para este caso. Los resultados se observan en las Tablas 130-132. Como viene ocurriendo en capítulos anteriores, la red neuronal es evaluada para 7 y 9 neuronas en la capa oculta. Además, se evalúa también con 12 y 15 neuronas ya que el número de ejemplos es ahora mucho mayor.

Tabla 127. Resultados de Relief-F en los datos iniciales y nuevos según el número de vecinos

Número de vecinos más cercanos	Ranking de variables
10	12,2,21,13,3,15,23,25,17,19,16,14,24,20,1,7,18,22,5,11,4,6,10,9,8
15	12,2,15,21,3,1,23,25,19,16,20,22,24,17,18,14,7,13,6,11,4,10,5,9,8
20	2,3,21,23,15,1,12,25,16,19,22,14,24,20,17,18,7,13,6,11,4,5,10,9,8
50	3,2,23,16,19,21,15,12,20,18,25,24,14,1,7,17,4,22,6,11,13,10,9,5,8

Evaluando el comportamiento de la red neuronal con las 10 variables elegidas por el método Relief-F, los resultados (mostrados en la Tabla 130) son parecidos independientemente del número de neuronas en la capa oculta. Aunque con 12 y 15 neuronas baja el número de ejemplos erróneos, el error máximo del conjunto de test, la media del error y la desviación típica se mantienen en valores muy similares en todos los casos. No obstante, con 12 neuronas en la capa oculta se obtienen unos resultados ligeramente mejores.

Tabla 128. Ranking de variables según Relief-F

Relief-F para los datos iniciales y nuevos			Relief-F para los datos iniciales		
Puesto	Variable	Media Posiciones	Puesto	Variable	Media Posiciones
1	2	1,75	1	14	2,75
2	3	3,25	2	3	4
3	21	4	3	13	4
4	12	4,25	4	12	5,25
5	15	5,25	5	5	7
6	23	5,25	6	2	7,25
7	16	8,5	7	4	8,5
8	19	8,5	8	16	9
9	25	8,75	9	21	10,5
10	1	10,25	10	20	11
11	20	12	11	7	12
12	24	12,75	12	17	12
13	14	13,25	13	9	13,5
14	17	13,5	14	19	14
15	18	14,5	15	25	14,5
16	22	14,75	16	15	15,75
17	13	15,25	17	18	16
18	7	16,25	18	10	16,75
19	6	19,75	19	22	17
20	4	20	20	24	17
21	11	20	21	23	17,5
22	5	22	22	1	19,25
23	10	22,5	23	6	22,5
24	9	23,75	24	11	23,75
25	8	25	25	8	24,25

Observando la Tabla 131, se puede comprobar que, evaluando la red neuronal con las 12 variables elegidas por Relief-F, se obtienen similares conclusiones que en el caso anterior. Son similares los resultados independientemente del número de neuronas de la capa oculta, aunque el error máximo en el conjunto de test es menor si se usan 7 neuronas en la capa intermedia; mientras que con 15 neuronas se obtiene la mínima media del error y el menor número de ejemplos que sobrepasan un error del 5%.

Tabla 129. Comparativa de la aplicación de Relief-F según los datos

Número de variables	Datos iniciales y nuevos	Datos iniciales	Variabes coincidentes
10	1, 2, 3, 12, 15, 16, 19, 21, 23, 25	2, 3, 4, 5, 12, 13, 14, 16, 20, 21	2, 3, 12, 16, 21 (5/10)
12	1, 2, 3, 12, 15, 16, 19, 20, 21, 23, 24, 25	2, 3, 4, 5, 7, 12, 13, 14, 16, 17, 20, 21	2, 3, 12, 16, 20, 21 (6/12)
15	1, 2, 3, 12, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25	2, 3, 4, 5, 7, 9, 12, 13, 14, 16, 17, 19, 20, 21, 25	2, 3, 12, 14, 16, 17, 19, 20, 21, 25 (10/15)

Tabla 130. Resultados de las salidas medias de la red neuronal,10 variables (Relief)

<i>CONJUNTO 1 (10 VARIABLES)</i>	<i>Media 7HN</i>	<i>Media 9HN</i>	<i>Media 12HN</i>	<i>Media 15HN</i>
Ejemplos erróneos (Entre/Test)	'54/24'	'58/25'	'36/21'	'32/22'
Error máximo conjunto test (dB)	14,18	12,64	14,52	14,51
Media del error (dB)	4,95	5,16	4,33	4,43
Desviación típica del error (dB)	1,49	1,51	1,45	1,44

Tabla 131. Resultados de las salidas medias de la red neuronal,12 variables (Relief)

<i>CONJUNTO 1 12 VARIABLES</i>	<i>Media 7HN</i>	<i>Media 9HN</i>	<i>Media 12HN</i>	<i>Media 15HN</i>
Ejemplos erróneos (Entre/Test)	'64/26'	'63/31'	'31/17'	'33/19'
Error máximo conjunto test (dB)	12,57	12,54	14,38	14,45
Media del error (dB)	5,69	5,66	4,02	3,94
Desviación típica del error (dB)	1,63	1,62	1,38	1,37

Tabla 132. Resultados de las salidas medias de la red neuronal,15 variables (Relief)

<i>CONJUNTO 1 15 VARIABLES</i>	<i>Media 7HN</i>	<i>Media 9HN</i>	<i>Media 12HN</i>	<i>Media 15HN</i>
Ejemplos erróneos (Entre/Test)	'35/19'	'38/21'	'26/15'	'28/17'
Error máximo conjunto test (dB)	12,49	13,19	14,26	14,15
Media del error (dB)	4,62	4,72	3,60	3,56
Desviación típica del error (dB)	1,50	1,50	1,33	1,32

Para el caso de 15 variables (Tabla 133), se obtienen unos resultados ligeramente diferentes. En este caso, sí se puede concluir que, con 12 y 15 neuronas en la capa oculta, la red se comporta considerablemente mejor. Aunque el error máximo del conjunto de test es ligeramente mayor, el número de ejemplos erróneos es más bajo, así como la media del error y la desviación típica.

Comparando con los resultados obtenidos con las variables elegidas por Relief-F con los datos iniciales (Tablas 44 - 49), éstos últimos son mejores. Como se concluye en el apartado anterior, al ser los datos de naturaleza más heterogénea y las variables elegidas tan diferentes, los resultados no son comparables, siendo más adecuado contrastar con los resultados obtenidos para la totalidad de las variables (Tablas 123 -126). En este caso, el número ejemplos erróneos es más bajo (entre 10 y 15 ejemplos en el conjunto de test, siendo en este caso entre 15 y 25) y también la media del error (alrededor de 2 y en torno a 4). En cambio, el error máximo del conjunto de test es similar (en ambos casos entre 12 y 14dB), por lo tanto, se puede considerar como positiva la reducción de variables de 25 a 15, que es el caso en el que se obtienen mejores resultados.

MÉTODO 2: SELECCIÓN DE SUBCONJUNTOS DE CARACTERÍSTICAS BASADAS EN CORRELACIÓN

De la misma forma que en el capítulo IV, se han seleccionado características según la correlación entre las mismas. Este método se ha elegido por los buenos resultados arrojados anteriormente. En la Tabla 133 se muestran los resultados de aplicar este método con diferentes técnicas de búsqueda. Ya que, según la técnica de búsqueda, el método de correlación devuelve un número de variables diferentes, se elige el número de variable que coincide en mayor número de casos. Estas variables seleccionadas son comparadas por las elegidas en el capítulo IV para los datos iniciales y coinciden una gran parte de ellas (Tabla 134). Es decir, que en este caso, es posible que se haya llegado a un conjunto de variables que defina bien a ambos conjuntos de datos.

Se resumen en la Tabla 135 los resultados de la aplicación de la red neuronal al nuevo conjunto de datos con las 9 variables elegidas con 7, 9, 10 y 12 variables en la capa oculta.

Observando los resultados de la Tabla 135, son mejores los obtenidos con 12 neuronas en la capa intermedia para todas las medidas comparables. El número de ejemplos erróneos es menor (32 en conjunto de entrenamiento y 20 en conjunto de test), la media del error es menor a 4dB y el error máximo del conjunto de test está alrededor de 10dB. En el resto de los casos siempre se superan estos valores.

Comparando con los resultados de la red neuronal ejecutada con las 25 variables y con el conjunto de datos inicial y nuevo (Tablas 123 - 126), se puede observar cómo, en el caso de 12 neuronas en la capa oculta incluso se mejora el error máximo del conjunto de test (10,28dB frente a 12,58 en el mejor de los casos -7 neuronas-), aunque el resto de medidas dan unos resultados peores.

El número de ejemplos erróneos en test es algo mayor, también superando en al menos 1dB la media del error. Aunque el comportamiento con la reducción de variables es ligeramente peor, el hecho de que se obtenga una buena medida en el error máximo del conjunto de test, hace pensar que ésta pueda ser una aceptable reducción de variables.

Ya que en este caso sí coinciden las variables elegidas con el mismo método de selección de características para solamente los datos iniciales, es posible hacer una comparación y sacar alguna conclusión sobre la naturaleza de ambos conjuntos de datos.

Observando la Tabla 135 y las Tablas 59 - 64, se puede concluir que, en ambos casos, la reducción de variables ha sido equilibrada, ya que se mantienen unos resultados positivos en general, no mermando la calidad de las predicciones de la red, lo que permite considerar este subconjunto de variables como bastante prometedor de cara a considerar un subconjunto definitivo.

Tabla 133. Variables seleccionadas según el método de búsqueda para los datos nuevos y antiguos

Método de búsqueda		Variables elegidas	
		Datos antiguos	Datos nuevos y antiguos
<i>Primero el mejor hacia adelante</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,12,16,20 : 5
<i>Primero el mejor hacia atrás</i>		5,8,10,11,12,13,14,16,17 : 9	8,9,10,11,12,13,16,17,19 : 9
<i>Primero el mejor bidireccional</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,12,16,20 : 5
<i>Búsqueda Exhaustiva</i>		5,8,10,11,12,13,14,16,17 : 9	8,9,10,11,12,13,16,17,19 : 9
<i>Búsqueda genética</i>	Tamaño población: 20 Número de generaciones: 20	5,8,10,11,12,13,14,16,17,19 : 10	8,10,12,16,20 : 5
Probabilidad de cruce: 0.6	Tamaño población: 100 Número de generaciones: 20	4,5,8,9,10,12,13,14,16,17,19 : 11	8,9,10,12,13,16,17,19 : 8
Probabilidad de mutación: 0.033	Tamaño población: 1000 Número de generaciones: 35	5,8,10,11,12,13,14,16,17,19 : 10	8,9,10,11,12,13,16,17,19 : 9
	Tamaño población: 1000 Número de generaciones: 150	5,8,10,11,12,13,14,16,17 : 9	8,9,10,11,12,13,16,17,19 : 9
	Tamaño población: 1000 Número de generaciones: 250	5,8,10,11,12,13,14,16,17 : 9	8,9,10,11,12,13,16,17,19 : 9
<i>Greedy StepWise</i>		8,17,10,13,5,12,11,19,16,4,14,9,6,22,20,23,25,21,15,18,24,2,7,3,1	6,8,12,16,20 : 5
<i>Selección lineal hacia adelante</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,12,16,20 : 5
<i>Selección lineal hacia adelante (flotante)</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,12,16,20 : 5
<i>Búsqueda al azar</i>	Nº de iteraciones: 8.388.608 (25.0% del espacio búsqueda)	5,8,9,10,12,13,14,16,17,19 : 10	8,9,10,11,12,13,16,17,19 : 9
<i>ScatterSearch Greedy Combination</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,12,16,20 : 5
<i>ScatterSearch Reduced Greedy</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,9,10,12,13,16,17,19 : 9
<i>Subset Size Forward Selection</i>		5,8,10,11,12,13,14,16,17 : 9	6,8,12,16,20 : 5

Tabla 134. Comparativa de la aplicación del método de correlación según los datos

Datos iniciales y nuevos	Datos iniciales	Variables coincidentes
8,9,10,11,12,13,16,17,19 (9)	5,8,10,11,12,13,14,16,17 (9) 5,8,10,11,12,13,14,16,17,19 (10) 4,5,8,9,10,12,13,14,16,17,19 (11)	8,9,10,12,13,16,17,19 (7) 8,9,10,12,13,16,17,19 (8) 8,9,10,12,13,16,17,19 (8)

Tabla 135. Resultados de las salidas medias de la red neuronal,9 variables (Correlación)

CONJUNTO 1 - 9 VARIABLES	Media 7HN	Media 9HN	Media 12HN	Media 15HN
Ejemplos erróneos (Entre/Test)	'43/26'	'46/29'	'32/20'	'42/21'
Error máximo conjunto test (dB)	11,55	12,15	10,28	11,94
Media del error (dB)	5,03	5,24	3,95	4,47
Desviación típica del error (dB)	1,55	1,59	1,39	1,50

MÉTODO 3: WRAPPER CON VALIDACIÓN CRUZADA

Se ejecuta el método wrapper con validación cruzada para 10 y 25 iteraciones.

a. Validación cruzada de 10 iteraciones

Tabla 136. Datos iniciales y nuevos

Posición	# variable	% subconjuntos
1	13	100
2	23	90
3	2	80
4	3	80
5	16	80
6	14	80
7	8	70
8	10	70
9	12	70
10	17	60
11	20	60
12	1	50
13	7	50
14	9	50
15	18	50
16	22	50
17	4	40
18	21	40
19	6	30
20	19	30
21	24	30
22	25	30
23	11	20
24	15	20
25	5	0

Tabla 137. Datos iniciales

Posición	# variable	% subconjuntos
1	4	100
2	13	100
3	14	100
4	2	50
5	5	50
6	1	30
7	19	30
8	3	20
9	15	20
10	18	20
11	24	20
12	6	10
13	9	10
14	12	10
15	16	10
16	17	10
17	20	10
18	22	10
19	23	10
20	7	0
21	8	0
22	10	0
23	11	0
24	21	0
25	25	0

Se seleccionan los conjuntos de variables según la Tabla 136, eligiendo las 11 y 16 mejores variables. En la Tabla 138 se muestra como en el caso de 11 variables, no coinciden los resultados al comparar con la ejecución de este método con los datos iniciales. Para 16 variables y comparando con las 19 elegidas anteriormente, coinciden en mayor medida.

Tabla 138. Comparativa de variables elegidas según el conjunto de datos

Datos iniciales y nuevos	Datos iniciales	Variables coincidentes
2, 3, 8, 10, 12, 13, 14, 16, 17, 20, 23 (11)	1, 2, 3, 4, 5, 13, 14, 15, 18, 19, 24 (11)	2, 3, 13, 14 (4)
1, 2, 3, 7, 8, 9, 10, 12, 13, 14, 16, 17, 18, 20, 22, 23 (16)	1, 2, 3, 4, 5, 6, 9, 12, 13, 14, 15, 16, 17, 18, 19, 18, 20, 22, 23, 24 (19)	1, 2, 3, 9, 12, 13, 14, 16, 17, 18, 20, 22, 23 (13)

En las Tablas 139 y 140 se muestran los resultados para la ejecución de la red neuronal con los datos iniciales y nuevos sólo para las características elegidas por este método.

Tabla 139. Resultados de las salidas medias de la red neuronal, 11 variables (Wrapper-Cross 10)

CONJUNTO 1 - 11 VARIABLES	Media 7HN	Media 9HN	Media 12HN	Media 15HN
Ejemplos erróneos (Entre/Test)	'43/21'	'40/19'	'32/21'	'32/16'
Error máximo conjunto test (dB)	13,42	11,93	11,81	11,27
Media del error (dB)	4,89	4,26	4,25	3,79
Desviación típica del error (dB)	1,57	1,47	1,47	1,38

Tabla 140. Resultados de las salidas medias de la red neuronal, 16 variables (Wrapper-Cross 10)

CONJUNTO 1- 16 VARIABLES	Media 7HN	Media 9HN	Media 12HN	Media 15HN
Ejemplos erróneos (Entre/Test)	'34/24'	'31/20'	'25/23'	'23/17'
Error máximo conjunto test (dB)	13,50	12,42	11,28	12,95
Media del error (dB)	4,53	4,00	3,82	3,60
Desviación típica del error (dB)	1,51	1,43	1,44	1,39

La reducción de 16 variables no tiene un comportamiento suficientemente bueno (Tabla 140). En el caso de 15 neuronas en la capa oculta (para ambos casos la mejor solución), mejora, respecto al caso de 11 variables, sólo en número de ejemplos erróneos (de 32/16 a 23/17). Por lo tanto, la solución que puede ser la más aceptable con este método es con el subconjunto de 11 variables.

Comparando con la ejecución de la red neuronal para las 25 variables (Tablas 123 - 126), se mejora el error máximo del conjunto de test, siendo éste de alrededor de 11dB en el mejor de los casos, mientras que anteriormente se obtenía 12,58dB como mínimo. El resto de valores son peores al reducir el número de variables.

b. Validación cruzada de 25 iteraciones

La validación cruzada con 25 iteraciones devuelve los resultados indicados en la Tabla 142, de donde se seleccionan las primeras 12 y 16 características elegidas por el

método. Se pueden comparar los resultados de las variables elegidas con los distintos conjuntos de datos en la Tabla 141.

Tabla 141. Comparativa de variables elegidas según el conjunto de datos

Datos iniciales y nuevos	Datos iniciales	Variables coincidentes
1, 2, 7, 8, 9, 10, 13, 14, 16, 18, 20, 23	1,2, 3, 4,5,13,14,15,18, 19, 24	1,2,13,14,18
1, 2, 3, 7, 8, 9, 10, 12, 13, 14, 16, 18, 20, 22, 23, 25	1,2,3,4,5,6,9,12,13,14,15,16,17,18,19,18,20,22,23,24	1,2,3,9,12,13,14,16,18,20,22,2

Tabla 142. Datos iniciales y nuevos

Posición	# variable	% subconjuntos
1	14	96
2	10	92
3	13	92
4	2	84
5	16	80
6	1	76
7	8	72
8	20	72
9	7	68
10	9	68
11	23	68
12	18	68
13	3	56
14	12	56
15	22	56
16	25	56
17	4	48
18	15	44
19	24	44
20	17	32
21	19	28
22	6	24
23	21	12
24	11	12
25	5	4

Tabla 143. Datos iniciales

Posición	# variable	% subconjuntos
1	4	100
2	13	100
3	14	100
4	5	44
5	10	36
6	1	32
7	21	32
8	12	28
9	3	24
10	19	24
11	2	20
12	16	20
13	18	20
14	17	16
15	22	16
16	6	12
17	23	12
18	9	8
19	20	8
20	24	8
21	11	4
22	7	0
23	8	0
24	15	0
25	25	0

En las Tablas 144 y 145 se muestran los datos devueltos por la red neuronal con las 12 y 16 variables elegidas por este método, para 7, 9, 12 y 15 neuronas en la capa intermedia.

Tabla 144. Resultados de las salidas medias de la red neuronal,12 variables (Wrapper-Cross 25)

<i>CONJUNTO 1- 12 VARIABLES</i>	<i>Media 7HN</i>	<i>Media 9HN</i>	<i>Media 12HN</i>	<i>Media 15HN</i>
Ejemplos erróneos (Entre/Test)	'21/19'	'18/19'	'22/22'	'21/16'
Error máximo conjunto test (dB)	14,41	14,72	14,98	14,09
Media del error (dB)	3,91	4,14	3,74	3,72
Desviación típica del error (dB)	1,47	1,50	1,46	1,42

Tabla 145. Resultados de las salidas medias de la red neuronal,16 variables (Wrapper-Cross 25)

<i>CONJUNTO 1</i>	<i>Media 7HN</i>	<i>Media 9HN</i>	<i>Media 12HN</i>	<i>Media 15HN</i>
Ejemplos erróneos (Entre/Test)	'20/19'	'12/18'	'24/17'	'18/20'
Error máximo conjunto test (dB)	11,32	13,64	12,28	13,50
Media del error (dB)	3,73	2,85	3,58	3,30
Desviación típica del error (dB)	1,41	1,22	1,39	1,32

Observando las Tablas 144 - 145, se puede deducir, como en el caso anterior, que tampoco la ampliación de 12 a 16 variables mejora demasiado los resultados. Aunque baja la media del error en todos los casos y el error máximo del conjunto de test es también más bajo, la diferencia no es tan grande como podría esperarse al incorporar 4 variables más. No obstante, en el caso de 9 neuronas en la capa oculta, sí se obtienen unos buenos resultados para las 16 variables, ya que los ejemplos erróneos bajan a 12/18 (entrenamiento/test), y la media del error baja de 3 dB.

Al comparar con la ejecución de la red neuronal para las 25 variables (Tablas 123 - 126), se observa cómo, en la reducción a 12 variables, los valores son peores, sobre todo en la media del error (alrededor de 4dB y nunca mayor a 3dB) y en los ejemplos erróneos(en torno a 20 y en torno a 10). En cambio, si se elige el mejor caso de las 16 variables (para 9 neuronas en la capa oculta), estos valores son similares: la media del error es de 2,85dB, y los ejemplos erróneos son 12/18. Por lo tanto, se podría considerar como válida esta reducción a 16 variables.

5.3 RESUMEN DE RESULTADOS

Teniendo en cuenta todos los subconjuntos de variables elegidos por los diferentes métodos de selección de características de la sección anterior, una de las mejores reducciones es la obtenida con validación cruzada (25 iteraciones) con 16 variables (Tablas 144 - 145), que da mejores resultados en media del error y en el número de ejemplos erróneos. Es interesante indicar que con la reducción a 9 variables obtenida por el método de correlación (Tabla 135) se obtiene el menor error máximo del conjunto de test, siendo además la mayor reducción de variables.

No obstante, y tal como se hizo en el capítulo IV, en esta sección se van a analizar los conjuntos de variables obtenidos en todos los procesos de selección de características que se ha llevado a cabo en la sección anterior. La Tabla 146 incluye un resumen de los resultados obtenidos con cada uno de los métodos de selección de características evaluados. Esta tabla servirá de base para deducir un conjunto de variables del entorno descriptivas del paisaje sonoro, esta vez con el nuevo conjunto de datos. Cuando ha sido necesario se ha recurrido a la opinión de expertos en ruido ambiental, para considerar la inclusión (o no) de alguna variable sobre la que se tenía duda. Con los conjuntos de variables seleccionados, se entrena con la red neuronal para poder obtener decisiones concluyentes sobre las características que reducen el conjunto inicial de forma suficiente.

Se han agrupado las características en cuatro conjuntos, de la misma forma que en el capítulo IV (Tabla 147). En un primer conjunto (A) se han seleccionado 10 características más destacadas del resto, por haber sido seleccionadas en todos o casi todos los métodos de selección de características. El segundo grupo (B), reúne además dos características más que se consideran también relevantes en alguno de los métodos; mientras que en el tercer grupo (C), se completan los anteriores con tres características más, que, si bien no son elegidos por todos los métodos, si son consideradas por algunos como importantes. Así, se dispone de tres conjuntos de 10, 12 y 15 variables, con los que se volverá a probar la red neuronal. El grupo D incluye variables que inicialmente no son consideradas suficientemente relevantes como para tenerlas en cuenta.

Comparando las variables elegidas tras los procesos de Selección de Características para los dos conjuntos de datos (Tablas 106 y 147), los dos grupos son algo diferentes. Se podría destacar que en el grupo A de ambos casos coinciden 8 de las 10 variables, lo cual indica que estas variables son relevantes e influyentes en el nivel de presión sonora producido.

En las Tablas 148 - 150, se resumen los resultados de la red neuronal para las 10, 12 y 15 variables elegidas. En ellas se indica la salida media para cada uno de los conjuntos, para 12 neuronas en la capa oculta. Los resultados para 7, 9 y 15 neuronas se encuentran en el anexo. Puede observarse cómo hay conjuntos que destacan sobre el resto y que, aunque los valores evaluados no son tan buenos como con todas las variables, compensa el haber podido reducir de 25 a 10, 12 ó 15 características.

La media del error baja en pocas ocasiones de los 3dB, mientras que con las 25 variables esto ocurre en la mayoría de los casos. También los ejemplos erróneos son más numerosos. La desviación típica se mueve entre 1,3 y 1,5 dB, mientras que con todas las variables, estaba entre 1,1 dB y, como máximo, 1,34 dB. En estos aspectos, los resultados, como era previsible, no son tan buenos como los que devuelve la red neuronal con 25 variables.

Por otro lado, la red neuronal se comportaba mejor con todas las variables para los conjuntos 2 y 5. En este caso se obtienen unos resultados similares. En el conjunto 5 se baja mucho en el error máximo del conjunto de test, y en el resto de valores, las pruebas con los conjuntos 1 y 2 son las que mejores resultados ofrecen.

Tabla 146. Resumen de resultados de Selección de Características con nuevo conjunto de datos

Método	Puesto (ranking variables) / Subconjunto	Nº variables						
		HN	9	10	11	12	15	16
RELIEF	2,3,21,12,15,23,16,19,25,1,20,24,14,17,18,22,13,7,6,4,1 1,5,10,9,8	7		24		26	19	
Ejemplos erróneos en test		9		25		31	21	
		12		21		17	15	
		15		22		19	17	
Media error		7		4,95		5,69	4,62	
		9		5,16		5,66	4,72	
		12		4,33		4,02	3,60	
		15		4,43		3,94	3,56	
CORRELACIÓN	8,10,11,12,13,14,16,17,19	7	26					
Ejemplos erróneos en test		9	29					
		12	20					
		15	21					
Media error		7	5,03					
		9	5,24					
		12	3,95					
		15	4,47					
Validación Cruzada 10 iter	13,23,2,3,16,14,8,10,12,17,20,1,7,9,18,22	7			21			24
Ejemplos erróneos en test		9			19			20
		12			21			23
		15			16			17
		7			4,89			4,53
Media error		9			4,26			4
		12			4,25			3,82
		15			3,79			3,6
		7				19		19
Validación Cruzada 25 iter	14,10,13,2,16,1,8,20,7,9,23,18,3,12,22,25	7				19		19
Ejemplos erróneos en test		9				19		18
		12				22		17
		15				16		20
Media error		7				3,91		3,73
		9				4,14		2,85
		12				3,74		3,58
		15				3,72		3,3

Respecto a la cantidad de variables consideradas, se mejora ligeramente conforme se cuenta con un mayor número de ellas. No obstante, ya con 10 características se obtienen unos resultados bastante aceptables.

En general, se puede considerar interesante la reducción de 25 a 10 variables ya que es positivo que se obtengan resultados aceptables con un número de características tan limitado.

En definitiva, globalmente, parece más adecuado considerar la reducción a 15 variables y la red neuronal con 12 neuronas ocultas, y en concreto, el conjunto de entrenamiento 2 es con el que se obtiene un mejor comportamiento de la red.

Parece factible utilizar redes neuronales para predecir el nivel de ruido ambiental urbano, usando como entradas las 25 variables consideradas inicialmente, y parece factible también reducir el número de variables de entrada, sin una pérdida grande de efectividad.

Tabla 147. Resumen de variables elegidas tras la selección de características

A	2	ENTORNO COMERCIAL/OCIO
A	3	PRESENCIA OBRAS CONSTRUCCIÓN
A	8	CAUDAL DE VEHÍCULOS PESADOS ASCENDENTE
A	10	CAUDAL DE MOTOCICLETAS ASCENDENTE
A	12	APARICIÓN DE VEHÍCULOS CON SIRENA
A	13	APARICIÓN ESA RELACIONADOS CON EL TRÁFICO
A	14	APARICIÓN ESA NO RELACIONADOS CON EL TRÁFICO
A	16	PENDIENTE (%)
A	17	CARRILES PENDIENTE ASCENDENTE
A	20	ESTADO DEL FIRME
B	7	CAUDAL DE VEHÍCULOS LIGEROS DESCENDENTE
B	23	ALTURA DE LA VÍA
C	1	PERÍODO DEL DÍA
C	18	CARRILES PENDIENTE DESCENDENTE
C	19	TIPO DE PAVIMENTO
D	9	CAUDAL DE VEHÍCULOS PESADOS DESCENDENTE
D	4	TIEMPO DE ESTABILIZACIÓN
D	5	TIPO DE FLUJO
D	6	CAUDAL DE VEHÍCULOS LIGEROS ASCENDENTE
D	11	CAUDAL DE MOTOCICLETAS DESCENDENTE
D	15	VELOCIDAD DE CIRCULACIÓN PROMEDIO
D	21	TIPO DE VÍA
D	22	ANCHURA DE LA VÍA
D	24	ANCHURA DE LA CALZADA
D	25	DISTANCIA AL RECEPTOR

Tabla 148. Resultados de las salidas medias de la red, 12HN. 10 variables (selección propia)

12 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'29/17'	'28/10'	'31/19'	'23/25'	'30/14'
Error máximo conjunto test (dB)	13,75	13,49	12,44	13,08	7,59
Media del error (dB)	3,96	3,50	3,70	4,12	3,77
Desviación típica del error (dB)	1,42	1,31	1,38	1,50	1,37

Tabla 149. Resultados de las salidas medias de la red, 12HN. 12 variables (selección propia)

12 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'26/18'	'30/12'	'31/18'	'23/23'	'42/20'
Error máximo conjunto test (dB)	12,07	13,19	13,55	12,25	7,82
Media del error (dB)	3,38	3,21	4,09	3,75	4,38
Desviación típica del error (dB)	1,32	1,26	1,52	1,40	1,47

Tabla 150. Resultados de las salidas medias de la red, 12HN. 15 variables (selección propia)

12 NEURONAS	CONJUNTO 1	CONJUNTO 2	CONJUNTO 3	CONJUNTO 4	CONJUNTO 5
Ejemplos erróneos (Entre/Test)	'24/16'	'23/11'	'24/16'	'27/27'	'24/18'
Error máximo conjunto test (dB)	12,16	11,70	12,03	13,81	8,94
Media del error (dB)	3,16	3,06	3,55	4,38	3,21
Desviación típica del error (dB)	1,27	1,24	1,37	1,53	1,30

CAPÍTULO VI: CONCLUSIONES Y TRABAJO FUTURO

Para finalizar la memoria se analiza el trabajo realizado y se resume mediante las conclusiones explicitadas sobre cada uno de los objetivos descritos en la introducción. A lo largo del trabajo realizado se han encontrado algunos aspectos o temas, que porque no incidían directamente con los objetivos planteados, se presentan como líneas de trabajo futuras que quedan abiertas y con las que este trabajo puede verse ampliado. Finalmente se incluyen las publicaciones relacionadas con el contenido de esta memoria.

6.1 CONCLUSIONES

Como primera aportación de esta memoria, se ha revisado la literatura actual sobre el uso de técnicas de Soft Computing en el Análisis del Ruido Ambiental Urbano, detallando las aplicaciones más interesantes para esta memoria, todo ello incluido y descrito ampliamente en el capítulo II. Este estudio y revisión de la literatura ha permitido, no sólo conocer el estado del arte en el momento inicial de la realización de este trabajo, sino ir comprobando cómo durante su desarrollo, éste trabajo ha servido de base para otros, y se ha utilizado como referencia en otros trabajos y publicaciones.

La mayor parte de los trabajos encontrados describen métodos de predicción de flujo de tráfico o del grado de molestia producida por el ruido, abundando las redes neuronales y los modelos difusos como técnicas de predicción en estos casos. Los algoritmos genéticos son utilizados en su mayoría para predecir el ruido urbano, mientras que los modelos ocultos de Markov se usan para clasificar el tipo de ruido.

En general, no hay muchos artículos prediciendo el ruido urbano con redes neuronales, y la mayoría de ellos usan modelos similares a las de este trabajo. No obstante, todos ellos utilizan un conjunto de datos de entrada menos completo, en número y naturaleza de las variables, y todos los que obtienen resultados más parecidos lo hacen en fechas posteriores y citando, en la mayoría de las ocasiones, las publicaciones de los resultados explicitados en esta memoria. Además, las redes neuronales propuestas en los diferentes artículos revisados se entrenan durante un número muy elevado de épocas, aunque incluyen un número de ejemplos similar al de este trabajo. Con la red neuronal desarrollada en esta tesis el entrenamiento es siempre mucho más rápido. No es necesario usar un conjunto de validación por lo que se dispone de más ejemplos para el aprendizaje, lo que permite obtener mejores resultados en menos tiempo.

La segunda aportación es el desarrollo de una red neuronal artificial para predecir el nivel de ruido urbano en la ciudad de Granada, que se presenta en el capítulo III. Se han seleccionado 25 variables de entrada, y como salida el nivel equivalente de ruido. Se han tomado 289 muestras para el conjunto inicial de datos, usando 200 ejemplos para el entrenamiento y dejando 89 para el conjunto de test. La estructura de red es del tipo perceptrón multicapa, y el algoritmo de entrenamiento ha sido el Backpropagation, con la variante de Levenberg-Marquardt, usando la implementación de MATLAB. Se ha considerado una sola capa oculta y, tras algunos experimentos, se seleccionaron 7 y 9 neuronas en la capa oculta.

Se seleccionaron 5 conjuntos diferentes para el entrenamiento y la red se entrenó 5 veces con cada uno. Con 7 neuronas ocultas en la red neuronal se obtienen unos resultados excepcionales, en todas las 25 ejecuciones, llegando a predecir el nivel de presión sonora con gran precisión. La media del error obtenido siempre ha estado por debajo de 0,8 dB, y se ha obtenido un número de ejemplos erróneos en test (error superior al 5%) siempre por debajo de 6.

Además, se ha considerado usar como salida definitiva de la red neuronal, la media de las 5 salidas correspondientes a un mismo conjunto de entrenamiento, lo que mejora aún más los resultados, reduciendo el número de casos con error, en los 5 conjuntos de entrenamiento seleccionados, hasta reducirlos a 0 en casi todos los conjuntos de entrenamiento y test. El entrenamiento es rápido, lo que podría permitir el uso de la red en tiempo real.

Los resultados de la red neuronal para 9 neuronas en la capa intermedia, son similares. La comparación con los modelos matemáticos clásicos utilizados en diversos países y/o ciudades, permite observar directamente cómo la red neuronal ofrece un comportamiento muy superior al de estos modelos.

Como tercera aportación vinculada al segundo objetivo, en el capítulo IV se estudia la posibilidad de reducir el número de variables de entrada, detectando si hay algunas poco relevantes, porque su eliminación no redundaría en pérdida de efectividad. Para tratar de reducir este número de variables, se han utilizado diferentes técnicas de selección y extracción de características.

Como técnica de extracción de características se usa el Análisis de Componentes Principales, que permite reducir las características de 25 a 11, que son en todos los casos combinaciones lineales de las primeras 25. Es decir, no se obtiene ninguna característica concreta, sino sólo una reducción de los datos. La red neuronal entrenada con estos nuevos datos obtiene unos suficientemente buenos resultados, aunque no tan buenos como los anteriores, pero siguen siendo mejores que los aportados por los métodos matemáticos clásicos.

Para la selección de características concretas, se ha utilizado un amplio número de métodos, obteniendo los mejores resultados con el método RELIEF y con el método de evaluación de subconjuntos de variables por consistencia. En el caso de RELIEF, la media del error se mueve en torno a 1 dB y el número de ejemplos erróneos nunca supera 9. Por otro lado, con el método de consistencia se obtienen, para 14 variables, unos resultados muy similares a los obtenidos con las 25 variables, con una media del error por debajo de 1dB en la mayoría de los casos y un bajo número de ejemplos erróneos, sobre todo para las salidas medias.

Como resumen de los resultados obtenidos, se seleccionan tres conjuntos con 10, 12 y 15 variables y se evalúa el resultado de la red neuronal con ellos. Los mejores resultados se obtienen con 15 variables, obteniendo también una media del error por debajo de 1dB en la mayoría de los casos, y llegando a no obtener ningún ejemplo erróneo en el conjunto de test para las salidas medias con 12 y 15 características. Los resultados están prácticamente al mismo nivel que los obtenidos con las 25 variables, y superan ampliamente a los métodos clásicos con los 3 conjuntos de variables seleccionados.

Estos resultados tan sorprendentes han llevado a considerar una posible excesiva, y no natural, homogeneidad de los datos de entrenamiento, que en algunas variables se manifestaba en que en casi todos los ejemplos considerados, el valor de la variable es el mismo o prácticamente el mismo. Surge la necesidad de buscar nuevos conjuntos de datos para verificar o no los resultados obtenidos.

Nuevos datos, también de la ciudad de Granada, y con mayor heterogeneidad en los valores de las variables, fueron aportados, y tras una necesaria adaptación de los datos, porque las variables de entrada no eran totalmente coincidentes, se usan los nuevos datos como test para la red neuronal desarrollada. Los resultados no tan excelentes, confirman la posibilidad de que los datos iniciales fueran demasiado homogéneos.

Para confirmar la validez del modelo, y como última aportación, se ha realizado de nuevo el estudio de los capítulos III y IV, pero con la totalidad de los datos disponibles, un total de 656 ejemplos. Se utilizan 450 para el entrenamiento y 206 para el test y se consideran también 12 y 15 neuronas en la capa oculta.

Los resultados son buenos, aunque no tan espectaculares como los obtenidos previamente. Con 15 neuronas en la capa oculta, la media del error se mueve en torno a 2-3dB, y el número de ejemplos erróneos en test está en cifras similares a 15 (8% del total). No obstante, tras comparar estos resultados con los obtenidos por los modelos matemáticos de predicción, los obtenidos por la red neuronal son amplia y claramente mejores que los obtenidos por éstos últimos.

También para este nuevo conjunto de datos se lleva a cabo un proceso de selección de características, en el que se evalúan los métodos que previamente han dado mejor resultado. Éstos devuelven diversos conjuntos de variables y como viene siendo habitual, al añadir características se obtienen mejores resultados. Se selecciona, finalmente, un conjunto con 15 variables, algo diferentes al obtenido con los datos iniciales y se evalúa la red neuronal con las mismas, obteniendo unos resultados mejores que los demás métodos de selección de características: una media del error en torno a 4dB y un número de ejemplos erróneos en test siempre menor a 20 (10% del conjunto de test).

Como conclusión final, ha quedado patente que las redes neuronales son sistemas válidos para identificar, y así poder predecir, el modelo del ruido ambiental urbano, mejorando ampliamente los modelos matemáticos clásicos que actualmente se utilizan en diversas ciudades y países.

6.2 TRABAJOS FUTUROS

Sin duda, el aspecto más importante que queda por hacer, es ampliar el modelo de la red neuronal para que pueda ser válido en diferentes países y ciudades. Para ello será necesario disponer de datos de mediciones de ruido ambiental en distintas ciudades de varios países, para determinar si una única red neuronal puede “aprender” el comportamiento del ruido ambiental urbano de forma global, o es más conveniente considerar una red neuronal para cada ciudad y/o país.

La predicción de la estructura espectral del ruido urbano es un aspecto que aporta mucha información en el análisis del ruido, y la utilización de redes neuronales para ello parece interesante.

En el tema de selección de características, sería interesante desarrollar un método de selección de características propio en el que el método de evaluación fuera la propia red neuronal entrenada.

Un aspecto interesante que se pretende abordar también, es la aplicación de técnicas de Soft Computing al análisis y predicción de molestias por el ruido, aprovechando los datos recogidos de las encuestas que se han realizado en la ciudad de Granada.

Usar RNA para la clasificación del ruido ambiental podría aportar también resultados interesantes.

6.3 PUBLICACIONES RELACIONADAS

Genaro, N.; Torija, A.; Ramos-Ridao, A.; Requena, I.; Ruiz, D.P.; Zamorano, M. "*A neural network based model for urban noise prediction.*" *The Journal of the Acoustical Society of America*, Vol. 128 pp.1738-1746.2010, ISSN: 1520-8524.

Torija, A.; Genaro, N.; Ruiz, D.P.; Ramos-Ridao, A.; Zamorano, M.; Requena, I. "*Priorization of acoustic variables: Environmental decision support for the physical characterization of urban sound environments*", *Building and Environment*, Vol. 45, Issue 6, pp. 1477-1489, 2010, ISSN: 0360-1323.

Genaro, N.; Torija, A.; Ramos-Ridao, A.; Requena, I.; Ruiz, D.P.; Zamorano, M. "*Modeling Environmental Noise Using Artificial Neural Networks,*" *Intelligent Systems Design and Applications*, International Conference on Intelligent Systems Design and Applications, pp. 215-219, 2009. ISBN: 978-0-7695-3872-3.

Genaro, N.; Requena, I. "*A neural network for urban noise prediction*". *New Trends on Intelligent Systems and Soft Computing*. Vol. I. 2009, pp. 69-82, ISBN: 978-84-338-4979-3.

BIBLIOGRAFÍA

- [1] Comisión Europea, «<http://ec.europa.eu/environment/noise/greenpap.htm>,» Noviembre 1996. [En línea]. [Último acceso: 23 06 2014].
- [2] Bruel & Kjaer, «<http://www.bksv.com/doc/br1630.pdf>,» 2000. [En línea]. [Último acceso: 23 06 2014].
- [3] Comisión Europea, «Directiva 2002/49/CE del parlamento europeo y del consejo de 25 de junio de 2002 sobre evaluación y gestión del ruido ambiental,» n^o Diario Oficial n^o 189, pp. 0012-0026, 2002.
- [4] «Ley 37/2003, de 17 de noviembre, del ruido,» B.O.E. de 18 Noviembre de 2003., 2003.
- [5] Junta de Andalucía, «Reglamento de Protección contra la Contaminación Acústica en Andalucía,» 2012. [En línea]. [Último acceso: 23 06 2014].
- [6] Ayuntamiento de Granada, «Ordenanza municipal de protección del ambiente acústico de Granada,» 2007.
- [7] A. Torija, D. Ruiz, O. Herrera y S. Serrano, «Estudio de la Relación entre LAeq y los niveles percentiles para la descripción del ruido ambiental,» *Revista de Acústica*, vol. 37, n^o 3-4, pp. 106-112, 2006.
- [8] A. Torija, D. Ruiz y A. Ramos-Ridao, A method for prediction of the stabilization time in traffic noise measurements, Madrid: Congreso ICA (International Congress on Acoustics), 2007.
- [9] A. Torija, D. Ruiz y A. Ramos-Ridao, «Modelo Lineal Multivariante de predicción de descriptores de ruido en la ciudad de Granada. Uso del L50 para la descripción del ruido de tráfico rodado,» de *Tecnoacústica Gandía*, 2006.
- [10] A. Torija, D. Ruiz y A. Ramos-Ridao, «Obtaining of a factor to describe the anomalous sound events in traffic noise measurements,» de *Congreso ICA (International Congress on Acoustics)*, Madrid , 2007.
- [11] Bruel & Kjaer, «Sound & Vibration Measurement,» 2000. [En línea]. Available: <http://www.bksves.com/tbdoc/3571/Libro%20Medioambiental.pdf>.
- [12] «Recomendación de la Comisión Europea 2003/613/CE de 6 de agosto de 2003,» Diario Oficial n^o L 212 de 22/08/2003, 2003.
- [13] A. Calvo-Manzano Ruiz, El ruido en la ciudad: Gestión y control, Sociedad Española de Acústica, 1991.

- [14] Secretaría de Estado para las Políticas del Agua y el Medio Ambiente, «Guías metodológicas para la elaboración de estudios de impacto ambiental: Carreteras y ferrocarriles,» Ministerio de Obras Públicas, Transportes y Medio Ambiente, 1995.
- [15] Z. Maekawa, «Noise Reduction by Screens,» *Applied Acoustics*, 1968.
- [16] «Guide du Bruit (1980). “Préliminaires aux études de bruit : Partie II”,» 1991.
- [17] Dept. of Environment HMSO, «Calculation of Road Traffic Noise,» United Kingdom, 1988.
- [18] «Richtlinien für den Lärmschutz an Straben (RLS 90),» Germany, 1990.
- [19] «Nordic Prediction Method for Road Traffic Noise (Statens Planverk 96),» Nordic Countries, 1996.
- [20] A. E. González, E. Gaja Díaz, A. Jorysz y G. Torres, «Monitoreo de Ruido Urbano: Determinación del Tiempo Mínimo de Muestreo en la Ciudad de Montevideo, Uruguay,» 2000.
- [21] A. Torija, D. Ruiz y A. Ramos-Ridao, «Analysis and Prediction of Noise Descriptors for Characterization of Urban Road Traffic Noise,» *Transportation. 1er Cuartil (16/83 subcampo ENGINEERING, CIVIL)*.
- [22] A. Torija, D. Ruiz y A. Ramos, «Noise pollution level corrected in terms of the incident total sonorous energy,» de *Congreso ICA (International Congress on Acoustics)*, Madrid, 2007.
- [23] S. Haykin, *Neural networks. A comprehensive foundation*, Nueva York: Prentice Hall, 1998.
- [24] G. J. Klir y B. Yuan, *Fuzzy Systems: Fuzzy sets and fuzzy logic: Theory and applications*, Prentice Hall, 1995.
- [25] A. Eiben y J. Smith, *Introduction to Evolutionary Computation.*, 2003.
- [26] Z. Ghahramani, *An introduction to hidden Markov models and Bayesian networks.*, World Scientific P. , 2001.
- [27] F. Rosenblatt, *Principles of Neurodynamics*, Nueva York: Spartan books, 1962.
- [28] F. Rosenblatt, «The perceptron: A probabilistic model for information storage and organization in the brain,» *Psychological Review*, vol. 65, 1958.
- [29] B. Widrow, *Generalization and information storage in networks of Adaline neurons*, Washington DC: Spartan books, 1962.

- [30] P. Werbos, «The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting,» Wiley, Hoboken, New Jersey, 1974.
- [31] D. Rumelhart, G. Hinton y R. Williams, «Learning Representations by BackPropagation Errors,» *Nature*, vol. 323, pp. 533 - 536, 1986.
- [32] G. Carpenter, «Distributed activation, search, and learning by ART and ARTMAP neural networks,» de *Proceedings International Conference on Neural Networks*, Piscataway, New Jersey, 1996.
- [33] B. Kosko, «Bidirectional associative memories,» *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 17, 1987.
- [34] K. Fukushima, «Cognitron: A self-organizing multilayered neural network,» *Biological Cybernetics*, vol. 20, pp. 121 - 136, 1975.
- [35] J. Hopfield, «Neural networks and physical systems with emergent collective computational abilities,» *Proceedings National Academy of Science*, vol. 79, 1982.
- [36] D. H. Ackley, G. H. Hinton y T. J. Sejnowski, «A learning algorithm for Boltzmann machines,» *Cognitive Science*, vol. 9, pp. 147-169, 1985.
- [37] T. Kohonen, *Self-Organization and Associative Memory Sciences*, Nueva York: Springer Verlag, 1984.
- [38] L. Zadeh, «Fuzzy sets,» *Information and Control*, vol. 8, p. 338-353, 1965.
- [39] D. E. Goldberg, *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, 1989.
- [40] A. Markov, «An example of statistical investigation in the text of eugene onyegin, illustrating coupling of tests in chains,» *Proceedings of the Academy of Sciences of St. Petersburg*, 1913.
- [41] L. Baum, «A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *Annals of Mathematical Statistics*,» vol. 41, p. 164-171, 1970.
- [42] H. Peng, L. Fuhui y C. Ding, «Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy,» *IEEE Trans. Pattern Anal. Mach.Intell*, vol. 27, p. 1226-1238, 2005.
- [43] D. Wettschereck, D. W. Aha y T. Mohri, «A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms,» *Artificial Intelligence Review*, vol. 11, pp. 273-314, 1997.
- [44] R. S. a. H. Liu, «Chi2: Feature selection and discretization of numeric attributes,» de

Proceedings of 7th IEEE Int. Conference on Tools with Artificial Intelligence, 1995.

- [45] T. Fountain, H. Almuallim y T. G. Dietterich, «Learning with Many Irrelevant Features,» 1991.
- [46] A. L. Blum y P. Langley, «Selection of relevant features and examples in machine learning,» *Artificial Intelligence*, pp. 245-271, 1997.
- [47] G. H. John, R. Kohavi y Pflieger, «Irrelevant features and the subset selection problem,» de *Proc. of the Eleventh International Conference on Machine Learning*, New Brunswick, NJ, USA, 1994.
- [48] P. Langley, «Selection of relevant features in machine learning,» de *Proceedings of the AAAI Fall Symposium on Relevance*, New Orleans, LA, USA, 1994.
- [49] R. Kohavi y G. H. John, « Wrappers for feature subset selection,» *Artificial Intelligence*, Vols. %1 de %297(1-2), pp. 273-324, 1997.
- [50] G. V. Lashkia y L. Anthony, «Relevant, irredundant feature selection and noisy example elimination,» *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*, vol. 34(2), pp. 888-898, 2004.
- [51] L. Yu y H. Liu, «Efficient Feature Selection via Analysis of Relevance and Redundancy,» *J. Mach. Learn. Res.* 5, pp. 1205-1224, 2004.
- [52] K. Kira y L. A. Rendell, «A practical approach to feature selection,» de *Proceedings of the Ninth International Conference on Machine Learning*, Aberdeen, Scotland, 1992.
- [53] D. Shalak, «Prototype and feature selection by sampling and random mutation hill climbing algorithms,» de *Proceeding of the eleventh International Conference on Machine Learning*, New Brunswick, NJ, 1994.
- [54] P. Narendra y K. Fukunaga, «A Branch and Bound Algorithm for Feature Subset Selection,» *IEEE Transactions on Computers*, vol. C26, nº 9, pp. 917-922, 1977.
- [55] H. Liu, H. Motoda y L. Yu, «Feature Selection with Selective Sampling,» de *Proceedings of the Nineteenth International Conference on Machine Learning (ICML '02)*, 2002..
- [56] H. Liu y H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, Boston: Kluwer Academic Publishers, 1998.
- [57] M. A. Hall, «Correlation-based feature selection for discrete and numeric class machine learning,» de *Proceedings of the Seventeenth International Conference on Machine Learning*, 2000.
- [58] L. Yu y H. Liu, «Redundancy based feature selection for microarray data,» de

Proceedings of the Tenth Conference on Knowledge Discovery and Data Mining, 2004.

- [59] Y. Kim, W. Street y F. Menczer, «Feature selection for unsupervised learning via evolutionary search,» de *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2000.
- [60] R. Kohavi y G. H. John, «Wrappers for feature subset selection,» *Artificial Intelligence*, Vols. 97(1-2), p. 273–324, 1997.
- [61] N. Metropolis y S. Ulam, «The Monte Carlo Method,» *Journal of the American Statistical Association (American Statistical Association)*, vol. 44 (247), p. 335–341, 1949.
- [62] I. Jolliffe, *Principal Component Analysis*, Springer, NY, 2002.
- [63] A. Arauzo, J. M. Benitez y J. L. Castro, «C-FOCUS: A continuous extension of FOCUS,» *Advances in Soft Computing - Engineering, Design and Manufacturing*, p. 225–232, 2003.
- [64] J. Quinlan, « Learning Efficient Classification Procedures and Their Application to Chess End Games. Machine Learning. An Artificial Intelligence Approach,» pp. 463-482, 1983.
- [65] I. Kononenko, « Estimating attributes Analysis and extensions of Relief.,» de *Proceedings of the Seventh European Conference on Machine Learning*, 1994.
- [66] U. d. W. - N. Zelanda, «<http://www.cs.waikato.ac.nz/ml/weka/>,» [En línea]. [Último acceso: 23 06 2014].
- [67] M. A. Hall, *Correlation-based Feature Subset Selection for Machine Learning.*, Department of Computer Science, University of Waikato., 1999.
- [68] H. Liu y R. Setiono, «A probabilistic approach to feature selection - A filter solution.,» de *13th International Conference on Machine Learning*, 1996.
- [69] C. Ding y H. Peng, «Minimum redundancy feature selection from microarray gene expression data,» de *Proc. 2nd IEEE Computational Systems Bioinformatics Conference*, Stanford, CA, 2003.
- [70] F. Joanneum, «Cross-Validation Explained,» 2005-2006.
- [71] R. Kohavi, «A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,» p. 1137—1143, 1995.
- [72] A. Whitney, «A Direct Method of Nonparametric Measurement Selection,» *IEEE Transactions on Computers*, vol. C20, n^o 9, pp. 1100-1103, 1971.

- [73] T. Marill y D. Green, «On the effectiveness of receptors in recognition systems,» *IEEE Transactions on Information Theory*, vol. 9, n° 1, pp. 11-17, 1963.
- [74] S. Stearns, «On selecting features for pattern classifiers,» de *Proceedings of the 3rd International Conference on Pattern Recognition*, 1976.
- [75] P. Pudil, J. Novovičová y J. Kittler, «Floating search methods in feature selection,» vol. 15, n° 11, pp. 1119-1125, 1994.
- [76] F. Garcia-Lopez, M. Garcia-Torres, B. Melian, J. Moreno-Perez y J. Moreno-Vega, «Solving feature subset selection problem by a Parallel Scatter Search,» *European Journal of Operational Research*, vol. 169, n° 2, pp. 477-489, 2006.
- [77] M. Hall y G. Holmes, «Benchmarking attribute selection techniques for discrete class data mining,» *IEEE Transactions on Knowledge and Data Engineering.*, vol. 15(6), pp. 1437-1447, 2003.
- [78] A. W. Moore y M. S. Lee, «Efficient Algorithms for Minimizing Cross Validation Error,» de *Eleventh International Conference on Machine Learning*, 1994.
- [79] Mathworks, «<http://www.mathworks.es/products/matlab/>,» [En línea]. [Último acceso: 23 06 2014].
- [80] H. Peng, F. Long y C. Ding, «<http://penglab.janelia.org/proj/mRMR/>,» [En línea]. [Último acceso: 23 06 2014].
- [81] G. Cammarata, S. Cavalieri y A. Fichera, «A Neural Network Architecture for Noise Prediction,» *Neural Networks*, vol. 8, n° 6, pp. 963-973, 1995.
- [82] Y. Avsar, A. Saral, M. Gönüllü, E. Arslankaya y U. Kurt, «Neural Network Modelling of Outdoor Noise Levels in a Pilot Area,» *Turkish J. Eng. Env. Sci.*, n° 28, pp. 149-155, 2004.
- [83] K. P. Dhananjay y B. N. Prashant, «Artificial Neural Network Modeling of Road Traffic Noise Descriptors,» de *Artificial Neural Network Modeling of Road Traffic Noise Descriptors*, 2008.
- [84] S. Givargis y H. Karimi, «A basic neural traffic noise prediction model for Tehran's roads,» *Journal of Environmental Management*, vol. 91, pp. 2529-2534, 2010.
- [85] K. Kumar, M. Parida y V. Katiyar, «Artificial neural network modeling for road traffic noise prediction,» de *Third International Conference on Computing Communication & Networking Technologies (ICCCNT)*, 2012.
- [86] J. K. Arora y P. V. Mosahari, «Artificial Neural Network Modelling of Traffic Noise in Agra-Firozabad Highway,» *International Journal of Computer Applications*, vol. 56, n° 2, pp. 6-10, October 2012.

- [87] F. Cirianni y L. Giovanni, «Environmental modeling for traffic noise in urban area,» *American Journal of Environmental Sciences*, vol. 8, nº 4, pp. 345-351, 2012.
- [88] N. Al-Mutairi y F. Al-Rukaibi, «Neural Network Models for Traffic Noise Quality Prediction: A Comparative Study,» *J Civil Environment Eng*, vol. 2, nº 1, 2012.
- [89] L. Yu y J. Kang, «Soundscape evaluation in city open spaces using artificial neural network,» de *Proceedings of the UIA—XXII World Congress of Architecture*, Istanbul, Turkey, 2005.
- [90] L. Yu y J. Kang, «Neural network analysis of soundscape in urban open spaces,» *Journal of the Acoustic Society of America*, vol. 117, p. 2591, 2005.
- [91] Zaheeruddin y Garima, «A neuro-fuzzy approach for prediction of human work efficiency in noisy environment,» *Applied Soft Computing*, vol. 6, pp. 283-294, 2006.
- [92] L. Yu, J. Kang y R. Harrisons, «Mapping soundscape evaluation in urban open spaces with artificial neural networks and ordinal logistic regression,» de *Proceedings of the 19th International Congress on Acoustics*, Madrid, 2007.
- [93] L. Yu y J. Kang, «Modeling subjective evaluation of soundscape quality in urban open spaces: An artificial neural network approach,» *Journal of the Acoustical Society of America*, vol. 126, nº 3, pp. 1163-1174, 2009.
- [94] S. Stoeckle, N. Pah, D. Kumar y N. Mclachlan, «Environmental sound sources classification using neural networks,» de *Intelligent Information Systems Conference, The Seventh Australian and New Zealand*, 2001.
- [95] T. Berg, «Classification of environmental noise by means of neural networks,» de *Forum Acusticum*, Sevilla, 2002.
- [96] L. Couvreur y M. Laniray, «Automatic Noise Recognition in Urban Environments Based on Artificial Neural Networks and Hidden Markov Models,» de *The 33rd International Congress and Exposition on Noise Control Engineering. Inter-noise*, Prague, 2004.
- [97] A. Betkowska, K. Shinoda y S. Furui, «Model Optimization for Noise Discrimination in Home Environment,» de *Symposium on Large-Scale Knowledge Resources (LKR2005)*, Tokyo, Japan, 2005.
- [98] L. Fortuna, L. Occhipinti, C. Vinci y M. Xibilia, «A Neuro-Fuzzy model of Urban Traffic,» de *Circuits and Systems. Proceedings of the 37th Midwest Symposium*, 1994.
- [99] M. S. Dougherty y M. R. Cobbett, «Short-term inter-urban traffic forecasts using neural networks,» *International Journal of Forecasting*, vol. 13, pp. 21-31, 1997.
- [100] C. Ledoux, «An urban traffic flow model integrating neural network,» *Transportation*

Research-C, vol. 5, n° 5, p. 287–300, 1997.

- [101] H. Dia, «An object-oriented neural network approach to short-term traffic forecasting,» *European Journal of Operational Research*, vol. 131, pp. 253-261, 2001.
- [102] H. Yin, S. Wong, J. Xu y C. Wong, «Urban traffic flow prediction using a fuzzy-neural approach,» *Transportation Research Part C: Emerging Technologies*, vol. 10, n° 2, pp. 85-98(14), 2002.
- [103] S. Ishak y C. Alecsandru, «Optimizing Traffic Prediction Performance of Neural Networks under Various Topological, Input, and Traffic Condition Settings,» *Journal of Transportation Engineering*, vol. 130, n° 4, 2004.
- [104] E. I. Vlahogianni, M. G. Karlaftis y J. C. Golias, «Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach,» *Transportation Research Part C*, vol. 13, pp. 211-234, 2005.
- [105] X. Jiang y H. Adeli, «Dynamic Wavelet Neural Network Model for Traffic Flow Forecasting,» *Journal of Transportation Engineering*, vol. 131, n° 10, pp. 771-779, 2005.
- [106] W. Zheng, D.-H. Lee, M. ASCE y Q. Shi, «Short-Term Freeway Traffic Flow Prediction-Bayesian Combined Neural Network Approach,» *Journal of transportation engineering*, vol. 132, pp. 114-121, 2006.
- [107] C. Quek, M. Pasquier y B. Lim, «POP-TRAFFIC: A Novel Fuzzy Neural Approach to Road Traffic Analysis and Prediction,» *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, vol. 7, n° 2, 2006.
- [108] C. Quek, M. Pasquier y B. Lim, «A novel self-organizing fuzzy rule-based system for modelling traffic flow behaviour,» *Expert Systems with Applications*, vol. 36, pp. 12167-12178, 2009.
- [109] B. Gültekin Çetiner, M. Sari y O. Borat, «A neural network based traffic-flow prediction model,» *Mathematical and Computational Applications*, vol. 15, n° 2, pp. 269-278, 2010.
- [110] K. Chan, T. Dillon, J. Singh y E. Chang, «Traffic Flow Forecasting Neural Networks Based on Exponential Smoothing Method,» de *6th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2011 .
- [111] G. Cammarata, A. Ficher, S. Graziani y L. Marletta, «Fuzzy logic for urban traffic noise prediction,» *J. Acoust. Soc. America*, vol. 98, n° 5, pp. 2607-2612, 1995.
- [112] J. Aguilera de Maya, «Método de predicción de ruido urbano basado en Teoría Fuzzy,» de *XXVIII Jornadas Nacionales De Acústica Y Encuentro Ibérico De Acústica*,

Oviedo, 1997.

- [113] R. Caponetto, M. Lavorgna, A. Martinez y L. Occhipinti, «GAS for fuzzy modeling of noise pollution,» de *First International Conference on Knowledge-Based Intelligent Electronic Systems*, 1997.
- [114] V. K. J. Zaheeruddin y G. V. Singh, «A Fuzzy Model for Noise-Induced Annoyance,» *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS*, vol. 36, nº 4, pp. 697-705, 2006.
- [115] W. Peng y R. V. Mayorga, «Assessing traffic noise impact based on probabilistic and fuzzy approaches under uncertainty,» *Stoch Environ Res Risk Assess*, vol. 22, pp. 541-550, 2008.
- [116] M. Ferri, J. A. Martínez, J. Alba, J. Ramis y C. Llinares, «A fuzzy set based approach for the estimation of loudness of environmental noise,» vol. 57, nº 5, pp. 561-568, 2009.
- [117] D. Botteldooren y A. Verkeyn, «Fuzzy modelling of traffic noise annoyance,» de *Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, 2001.
- [118] D. Botteldooren y A. Verkeyn, «Fuzzy models for accumulation of reported community noise annoyance from combined sources',» *Journal of Acoustic Society America*, vol. 112, p. 1496-1508, 2002.
- [119] D. Botteldooren y V. A., «An iterative fuzzy model for cognitive processes involved in environment quality judgement,» de *Proceedings of FUZZ-IEEE 2002*, Hawaii, 2002.
- [120] D. Botteldooren, A. Verkeyn, C. Cornelis y M. De Cock, «'On the meaning of noise annoyance modifiers: A Fuzzy Set Theoretical Approach,» de *Acta Acust. united with Acustica*, 2002.
- [121] D. Botteldooren, A. Verkeyn y P. Lercher, «Noise annoyance modelling using fuzzy rule based systems,» *Noise Health*, vol. 4, p. 27-44, 2002.
- [122] D. Botteldooren y A. Verkeyn, «Aggregation of specific noise annoyance to a general noise annoyance rating: a fuzzy model,» de *Proceedings of the 9th International Congress on Sound and Vibration*, Orlando, Florida, 2002.
- [123] D. Botteldooren y A. Verkeyn, «Annoyance Prediction with Fuzzy Rule Bases,» de *Proceedings of the 5th International FLINS Conference*, World Scientific, Singapore, 2002.
- [124] D. Botteldooren y A. Verkeyn, «The effect of Land-Use variables in a Fuzzy Rule Based Model for Noise Annoyance,» de *The 2002 International Congress and Exposition on Noise Control Engineerin*, Dearborn, MI, USA, 2002.
- [125] D. Botteldooren, A. Verkeyn y M. K. De Cock, «Generating Membership functions for

a Noise Annoyance Model from Experimental Data,» *Studies in Fuzziness and Soft Computing*, vol. 127, pp. 51-67, 2003.

- [126] D. Botteldooren, A. Verkeyn y P. Lercher, «A fuzzy rule based framework for noise annoyance modelling,» *Journal of Acoustic Society America*, vol. 114, p. 1487– 1498, 2003.
- [127] D. Botteldooren y P. Lercher, «Soft-Computing base analysis of the relationship between annoyance and coping with noise and odor,» *Journal of Acoustic Society America*, vol. 115 (6), pp. 2974-2985, 2004.
- [128] F. Beritelli, S. Casale y G. Ruggeri, «New Results in Fuzzy Pattern Classification of Background Noise,» de *Proceedings of ICSP2000*, 2000.
- [129] B. Uzkent y B. D. Barkana, «Automatic environmental noise source classification model using fuzzy logic,» *Expert Systems with Applications*, vol. 38, p. 8751–8755, 2011.
- [130] L. Dimitriou, T. Tsekeris y A. Stathopoulos, «Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow,» *Transportation Research Part C*, vol. 16, p. 554–573, 2007.
- [131] L. Dimitriou, T. Tsekeris y A. Stathopoulos, «Fuzzy Modeling Approach for Combined Forecasting of Urban Traffic Flow,» *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, 2008.
- [132] Ö. Gündođdu, M. Gökdag y F. Yüksel, «A traffic noise prediction method based on vehicle composition using genetic algorithms,» *Applied Acoustics*, vol. 66, p. 799–809, 2005.
- [133] S. Rahmani, S. Mousavi y M. Kamali, «Modeling of road-traffic noise with the use of genetic algorithm,» *Applied Soft Computing*, vol. 11, pp. 1008-1013, 2011.
- [134] C. Couvreur, V. Fontaine, P. Gaunard y C. Mubikangiey, «Automatic Classification of Environmental Noise Events by Hidden Markov Models,» *Applied Acoustics*, vol. 54, nº 3, pp. 187-206, 1998.
- [135] P. Gaunard, C. Mubikangiey, C. Couvreur y V. Fontaine, «Automatic classification of environmental noise events by Hidden Markov models,» de *Acoustics, Speech and Signal Processing. Proceedings of the 1998 IEEE International Conference*, 1998.
- [136] L. Ma, D. Smith y B. Miller, «Environmental Noise Classification for Context-Aware Applications,» de *Proc. DEXA, (LNCS 2736)*, 2003.
- [137] L. Ma, D. Smith y B. Miller, «Context Awareness using Environmental Noise Classification,» de *Proc. Eurospeech 2003*, Geneva, Switzerland, 2003.

- [138] M. Burgess, «Noise prediction for urban traffic conditions-related to measurements in the Sydney Metropolitan Area,» *Applied Acoustics*, vol. 10, pp. 1-7, 1977.
- [139] R. Josse, *Notions d acoustique*, París: Eyrolles, 1972.
- [140] D. Bertoni, A. Franchini y M. Magnoni, *Il Rumore Urbano e l'Organizzazione de Territorio*, Bologna: Pitagora Editrice, 1987, pp. 45-69.
- [141] D. Specht, «A general regression neural network,» *IEEE Trans. Neural Networks*, vol. 2, pp. 568-576, 1991.
- [142] A. García y D. Bernal, «The prediction of traffic noise levels in urban areas,» de *Proceedings of the International Conference on Noise Control Engineering*, Munich, Germany, 1985.
- [143] F. Cirianni y G. Leonardi, «Road traffic noise prediction models in the metropolitan area of the Strait of Messina,» de *Proc. Inst. Civil Eng. Transport*, 2011.
- [144] J.-S. Jang, «ANFIS: adaptive-network-based fuzzy inference system,» *IEEE Trans. Syst. Man Cyb.*, vol. 23, p. 665-685, 1993.
- [145] H. Hermansky, «Perceptual Linear Predictive (PLP) Analysis of Speech,» *J. Acoustic Soc. Am.*, vol. 87, nº 4, pp. 1738-1752, 1990.
- [146] A. Waibel, «Modular Construction of Time Delay Neural Networks for Speech Recognition,» *Neural Computation*, vol. 1, 1989.
- [147] R. L. Mantara, *Approximate Reasoning Models*, London, U.K.: Ellis Horwood Limited, 1990.
- [148] A. Calixto, F. Diniz y P. Zannin, «The statistical modeling of road traffic noise in an urban setting,» *Cities 2003*, vol. 20, pp. 23-29, 2003.
- [149] W. Lam y M. Tam, «Reliability analysis of traffic noise estimates in Hong Kong,» *Transport Res D*, vol. 3, pp. 239-248, 1998.
- [150] K. Shinoda y K. Iso, «Efficient reduction of gaussian components using MDL criterion for HMM-based speech recognition,» de *Proc. ICASSP2002*, 2002.
- [151] A. González, E. Gaja Díaz, A. Jorysz y G. Torres, «Desarrollo de un modelo predictivo de ruido urbano adaptado a la realidad de la ciudad de Montevideo, Uruguay,» *Tecnoacústica Madrid*, 2000.
- [152] A. Torija, N. Genaro, D. Ruiz, A. Ramos-Ridao, M. Zamorano y I. Requena, «Priorization of acoustic variables: Environmental decision support for the physical characterization of urban sound environments,» *Building and Environment*, vol. 45, nº 6, pp. 1477-1489, 2010.

- [153] N. Genaro, A. Torija, A. Ramos, I. Requena, D. P. Ruiz y M. Zamorano, «Modeling Environmental Noise Using Artificial Neural Networks,» de *Intelligent Systems Design and Applications, 2009. ISDA '09. Ninth International Conference on*, Pisa, Italia, 2009.
- [154] N. Genaro, A. Torija, A. Ramos-Ridao, I. Requena, D. Ruiz y M. Zamorano, «A neural network based model for urban noise prediction,» *The Journal of the Acoustical Society of America*, vol. 128, pp. 1738-1746, 2010.
- [155] A. Araúzo Azofra, *Un Sistema Inteligente para Selección de Características en Clasificación.*, Granada, 2006.
- [156] H. Almuallim y T. G. Dietterich, «Learning boolean concepts in the presence of many irrelevant features,» *Artificial Intelligence*, Vols. 1 de 269(1, 2), p. 279-305, 1994.
- [157] K. P. Dhananjay y B. N. Prashant, «Urban Noise: Ann Modeling Of Noise Levels Due To Vehicular Traffic Flow In Indian Intermediate City,» de *15th International Congress on Sound and Vibration*, Daejeon, Korea, 2008.
- [158] «Directiva 2002/49/CE del parlamento europeo y del consejo de 25 de junio de 2002 sobre evaluación y gestión del ruido ambiental».

