



UNIVERSIDAD MIGUEL HERNANDEZ DEPARTAMENTO DE INGENIERIA DE  
SISTEMAS INDUSTRIALES DIVISION DE INGENIERIA ELECTRICA

Escuela Politécnica Superior de Elche

DESARROLLO Y EVALUACION DE TECNICAS Y MODELOS PARA LA  
IDENTIFICACION Y CLASIFICACION DE SEGMENTOS DE CONSUMIDORES DE  
ENERGIA. APLICACIÓN AL ANALISIS DE LAS POSIBILIDADES DE  
PARTICIPACION DE SECTORES DE LA DEMANDA DE ENERGIA ELECTRICA EN  
MERCADOS COMPETITIVOS

Autor

**Sergio Valero Verdú**

Director

**D. Antonio Gabaldón Marín**

**Dr. Ingeniero Industrial**

**Elche, Julio de 2007**



# Tribunal

Presidente:

Vocal:

Vocal:

Vocal:

Vocal Secretario:



Calificación de la Tesis: \_\_\_\_\_



# AUTORIZACIÓN DE PRESENTACION DE TESIS DOCTORAL

Director: Dr. D. Antonio Gabaldón Marín

Título de la tesis: ***Desarrollo y Evaluación de Técnicas y Modelos para la Identificación y Clasificación de Segmentos de Consumidores de Energía. Aplicación al Análisis de las Posibilidades de Participación de Sectores de la Demanda de Energía Eléctrica en Mercados Competitivos***

Autor: Sergio Valero Verdú

Departamento: Ingeniería de Sistemas Industriales  
Universidad Miguel Hernández

EL director de la Tesis reseñada CERTIFICA QUE HA SIDO REALIZADA BAJO MI DIRECCIÓN POR d. Sergio Valero Verdú en el Departamento de Ingeniería de Sistemas Industriales de la Universidad Miguel Hernández y autorizo su presentación de acuerdo con lo dispuesto en el artículo 8º del R.D. 778/1998, 30 de Abril, por el que se regula el tercer ciclo de estudios universitarios.

Elche, a            de            del 2006-05-21

Firmado: Dr. D. Antonio Gabaldón Marín



DEPARTAMENTO DE INGENIERIA DE SISTEMAS INDUSTRIALES

D. Oscar Reinoso García, Titular de Universidad y Director del Departamento de Ingeniería de Sistemas Industriales de la Universidad Miguel Hernández.

**Certifica**

Que el trabajo realizado por D. Sergio Valero Verdú titulado “**Desarrollo y Evaluación de Técnicas y Modelos para la Identificación y Clasificación de Segmentos de Consumidores de Energía. Aplicación al Análisis de las Posibilidades de Participación de Sectores de la Demanda de Energía Eléctrica en Mercados Competitivos**”, ha sido dirigido por el Dr. D. Antonio Gabaldón Marín y realizado en el Departamento de Ingeniería de Sistemas Industriales, se encuentra en condiciones de ser leído y defendido como Tesis Doctoral ante el correspondiente tribunal en la Universidad Miguel Hernández.

Lo que firmo para los efectos oportunos en Elche a            de            del 2007

Fdo. D. Oscar Reinoso García  
Director del Departamento de Ingeniería  
de Sistemas Industriales





# Resumen

El trabajo de investigación que ha dado como fruto esta Tesis nace de la repercusión que tiene y tendría en un futuro, la participación activa de la demanda en los sistemas y mercados eléctricos.

Entre las principales barreras de la participación real de la demanda en los mercados destacan las relativas a la generación distribuida y el almacenamiento. Actualmente y a la vista de lo sucedido en otros mercados eléctricos internacionales (EEUU, Noruega, Inglaterra, Australia, etc..) los pequeños y medianos consumidores eléctricos solo pueden participar en el mercado a través de empresas comercializadoras o las denominadas empresas agregadoras. La misión de estas compañías es la de ofrecer a sus clientes paquetes o programas, previamente definidos (estudiados) para facilitar la participación de estos en el mercado. Para ello estas empresas y los mismos clientes necesitan conocer los diferentes segmentos de demanda eléctrica existentes para determinadas zonas geográficas. Solo así es posible crear políticas y programas específicos de participación. Del mismo modo esta identificación y clasificación de sectores permitirá a dichas empresas la definición de tarifas específicas adaptadas a determinados clusters de clientes.

En este sentido la Tesis analiza, evalúa y define en una primera fase de la investigación una herramienta para la identificación y segmentación de consumidores eléctricos a partir de datos de perfiles consumo diario (curvas de carga) de los distintos consumidores que forman el espectro de estudio. Posteriormente, en la segunda fase de la investigación, esta herramienta basada en los mapas autoorganizados de kohonen, permitirá definir un modelo de participación de la demanda en mercados eléctricos competitivos, utilizando modelos físicos de carga y consiguiendo demostrar los beneficios conseguidos en cuanto a ahorros de consumo y reducción de costes.

Por último, destacar que los resultados de la investigación fruto de esta Tesis han sido publicados en dos prestigiosas revistas internacionales del área de la Ingeniería Eléctrica catalogadas con índice de impacto SCI. Un tercer artículo ha sido aceptado oficialmente para su publicación en una tercera revista también de ámbito internacional. Las copias de los dos artículos publicados y de la carta oficial de aceptación a publicación del tercer artículo han sido añadidas al final de la Tesis en el punto ANEXO.



# INDICE GENERAL

## CAPÍTULO 1: OBJETIVOS Y ESTRUCTURA DE LA TESIS

1.1.	OBJETIVOS DE LA TESIS.....	21
1.2.	MOTIVACION.....	25
1.3.	METODOLOGIA.....	26
1.4.	ESTRUCTURA DE CAPITULOS.....	28
2.	CAPITULO 2.....	31
2.1.	INTRODUCCIÓN .....	33
2.2.	LOS MERCADOS ELÉCTRICOS.....	35
2.2.1.	ORIGENES Y EVOLUCIÓN.....	35
2.2.1.1.	MODELO DE MONOPSONIO O COMPRADOR ÚNICO.....	37
2.2.1.2.	MODELO DE COMPETENCIA MAYORISTA .....	37
2.2.1.3.	MODELO DE LIBERTAD DE ELECCIÓN PARA TODOS LOS CONSUMIDORES .....	38
2.2.2.	LA LIBERALIZACION COMO PROCESO.....	38
2.2.3.	CARACTERÍSTICAS DEL CONSUMO ELÉCTRICO.....	40
2.2.4.	CONTEXTO ECONOMICO.....	43
2.2.4.1.	EL SECTOR ELECTRICO Y LA ACTIVIDAD ECONOMICA .....	43
2.2.5.	CONTEXTO REGULADOR.....	44
2.2.5.1.	REGULACIÓN TRADICIONAL Y REGULACIÓN EN COMPETENCIA.....	44
2.2.5.2.	LA NUEVA REGULACION ELECTRICA.....	45
2.2.6.	EL MERCADO ELÉCTRICO ESPAÑOL .....	46
2.2.6.1.	SITUACION ACTUAL .....	46
2.3.	LA DEMANDA Y SU PARTICIPACIÓN EN ENTORNOS REGULADOS Y LIBERALIZADOS .....	49
2.3.1.	INTRODUCCION .....	49
2.3.2.	LA DECADA DE LOS SETENTA .....	50
2.3.3.	LA DECADA DE LOS OCHENTA .....	51
2.3.4.	LA DECADA DE LOS 90.....	52
2.3.5.	LA GESTION DE LA DEMANDA ELECTRICA .....	54
2.3.6.	PRINCIPIOS DE LA RESPUESTA DE LA DEMANDA .....	55
2.3.7.	INTEGRACION DE LA RESPUESTA DE LA DEMANDA EN LOS MERCADOS, EN LOS ULTIMOS AÑOS.....	56
2.3.7.1.	AGENTES INTERESADOS .....	57
2.3.7.2.	COMERCIALIZADORES Y AGREGADORES .....	57
2.3.7.3.	LOS CONSUMIDORES .....	58
2.3.8.	ENTORNOS LIBERALIZADOS.....	58
2.3.8.1.	EL MERCADO DE NUEVA INGLATERRA (EE.UU.).....	58
2.3.9.	CARACTERÍSTICAS DE LAS POSIBILIDADES Y REQUISITOS ACTUALES DE PARTICIPACION.....	61
2.4.	SEGMENTACION Y AGREGACION.....	63
2.4.1.	CONSUMIDORES CON POTENCIAL DE PARTICIPACION EN EL MERCADO .....	66
2.4.2.	CONCLUSIONES.....	70
3.	Revisión de técnicas de segmentación .....	72
3.1.	REVISIÓN DE TÉCNICAS .....	74
3.1.1.	INTRODUCCION .....	74

3.1.2.	CARACTERÍSTICAS GENERALES DE LOS PROCESOS DE CLASIFICACIÓN.....	74
3.1.3.	ADQUISICIÓN Y GESTIÓN DE DATOS.....	78
3.1.3.1.	PRETRATAMIENTO O ACONDICIONAMIENTO DE DATOS.....	79
3.1.3.2.	OBTENCIÓN DE CURVAS EN P.U.....	80
3.1.3.3.	DETECCIÓN DE DATOS ANÓMALOS O ATÍPICOS.....	81
3.1.3.4.	PRECLASIFICACION.....	82
3.2.	TÉCNICAS DE CLASIFICACIÓN.....	83
3.2.1.	MÉTODOS DE ANÁLISIS DE PERFILES DE CARGA.....	83
3.3.	TÉCNICAS DE AGRUPACIÓN.....	85
3.3.1.	ALGORITMOS DE AGRUPACIÓN.....	86
3.3.2.	TÉCNICAS ESTADÍSTICAS.....	88
3.3.2.1.	ANALISIS DISCRIMINANTE.....	89
3.3.2.2.	ESCALAS MULTIDIMENSIONALES.....	90
3.3.2.3.	MANOVA (ANÁLISIS MULTIVARIANTE DE VARIANZA).....	91
3.3.2.4.	ALGORITMO K-MEANS.....	92
3.3.2.5.	ÁRBOLES DE DECISION.....	93
3.3.3.	REDES NEURONALES.....	94
3.3.3.1.	ADALINA.....	96
3.3.3.2.	PERCEPTRÓN MULTICAPA (MLP).....	97
3.3.3.3.	RED COMPETITIVA LVQ.....	97
3.3.3.4.	REDES DE ELMAN.....	98
3.3.3.5.	RED DE CAPA COMPETITIVA.....	98
3.3.3.6.	REDES DE BASE RADIAL.....	99
3.3.3.7.	RED DE REGRESION GENERALIZADA.....	100
3.3.3.8.	RED NEURONAL PROBABILISTICA.....	100
3.3.4.	LÓGICA DIFUSA.....	101
3.3.4.1.	ANFIS (SISTEMA DE INFERENCIA NEURO-FUZZY ADAPTATIVO).....	101
3.3.4.2.	CLUSTERING SUBSTRACTIVO.....	102
3.3.4.3.	FUZZY C-MEANS CLUSTERING.....	103
3.3.5.	LA BONDAD DE LA AGRUPACIÓN REALIZADA.....	103
3.4.	CONCLUSIONES.....	106
4.	Análisis de técnicas de segmentación.....	108
4.1.	INTRODUCCIÓN.....	110
4.2.	TÉCNICAS ANALIZADAS.....	110
4.3.	HERRAMIENTAS Y SOFTWARE DISPONIBLE.....	111
4.3.1.	IMPLEMENTACION DE REDES NEURONALES ARTIFICIALES.....	111
4.3.2.	PROGRAMAS DE SIMULACION DE REDES NEURONALES.....	112
4.3.2.1.	PROGRAMAS COMERCIALES.....	112
4.3.2.2.	PROGRAMAS DE LIBRE DISTRIBUCION.....	113
4.3.3.	IMPLEMENTACION DE SISTEMAS BORROSOS.....	114
4.3.3.1.	ENTORNOS DE LOGICA BORROSA.....	115
4.3.4.	IMPLEMENTACION DE TECNICAS ESTADISTICAS.....	116
4.3.5.	SOFTWARE SELECCIONADO.....	116
4.4.	ESPECTRO DE DATOS DE ESTUDIO.....	117
4.4.1.	SELECCIÓN DE LOS RASGOS ADECUADOS PARA LA CARACTERIZACIÓN DE LOS CLIENTES.....	117
4.4.2.	PRIMER CONJUNTO DE DATOS (DOS CONSUMIDORES).....	118
4.4.3.	SEGUNDO CONJUNTO DE DATOS (VEINTISÉIS CONSUMIDORES) 120	
4.4.4.	ANÁLISIS DE LAS TÉCNICAS DE CLASIFICACIÓN.....	127
4.4.5.	IDONEIDAD DE LAS TÉCNICAS.....	127

4.5.	RESULTADOS Y DISCUSIÓN .....	129
4.5.1.	TÉCNICAS DE CLASIFICACIÓN .....	130
4.5.1.1.	TECNICAS ESTADISTICAS .....	130
4.5.1.2.	REDES NEURONALES .....	138
4.5.1.3.	LOGICA FUZZY .....	146
4.5.1.4.	CONCLUSIONES DE LAS TECNICAS DE CLASIFICACION ...	148
4.5.2.	TECNICAS DE CLASIFICACION E IDENTIFICACION.....	149
4.5.2.1.	SEGUNDO CONJUNTO DE DATOS (VEINTISÉIS CONSUMIDORES) .....	150
4.6.	CONCLUSIONES.....	160
5.	CAPITULO 5.....	162
5.1.	INTRODUCCIÓN .....	164
5.2.	ACONDICIONAMIENTO Y PROCESAMIENTO DE DATOS .....	164
5.2.1.	CONCEPTOS .....	164
5.2.2.	APLICACIÓN DE LOS SOMS AL DATA MINING .....	165
5.2.3.	CASO DE ESTUDIO .....	165
5.3.	CLASIFICACIÓN DE PATRONES DE CONSUMO.....	167
5.4.	APLICACIÓN DE SOM AL FILTRADO DE DATOS ANÓMALOS .....	168
5.5.	RESULTADOS TRAS APLICAR EL FILTRADO .....	174
5.6.	CONCLUSIONES.....	177
6.	CAPITULO 6.....	178
6.1.	CLASIFICACIÓN DE PATRONES DE CONSUMO.....	180
6.1.1.	INTRODUCCIÓN .....	180
6.1.2.	ESPECTRO DE DATOS .....	181
6.1.3.	TRANSFORMACIÓN DE LOS DATOS DE LA DEMANDA DEL DOMINIO DEL TIEMPO AL DOMINIO DE LA FRECUENCIA.....	182
6.2.	ENTRENAMIENTOS CON SOM.....	187
6.2.1.	ENTRENAMIENTOS CON DATOS DE ENTRADA EN EL DOMINIO DEL TIEMPO.....	189
6.2.2.	ENTRENAMIENTOS CON DATOS DE ENTRADA EN EL DOMINIO DE LA FRECUENCIA.....	190
6.3.	CLASIFICACIÓN E IDENTIFICACIÓN DE NUEVOS CONSUMIDORES	193
6.4.	ESTUDIO DE RESULTADOS DE CLASIFICACIÓN CON MÁS DE 100 CONSUMIDORES .....	195
6.5.	CONCLUSIONES.....	201
7.	CAPITULO 7.....	202
	Herramienta para la elección de recursos de energía distribuida y consumidores en mercados eléctricos liberalizados .....	202
7.1.	INTRODUCCIÓN .....	204
7.2.	EL MERCADO ELÉCTRICO ESPAÑOL .....	205
7.3.	PROGRAMAS DE RESPUESTA DE LA CARGA. OPORTUNIDADES ...	207
7.3.1.	PROGRAMAS DE RESPUESTA DE LA DEMANDA EN EL MERCADO DIARIO (DAY-AHEAD LOAD RESPONSE) .....	208
7.3.2.	PROGRAMAS DE RESPUESTA DE LA DEMANDA EN EL MERCADO A TIEMPO REAL (REAL-TIME LOAD RESPONSE PROGRAMS) .....	209
7.3.3.	ANÁLISIS DE LOS PROGRAMAS DEL MERCADO DIARIO. ....	210
7.3.4.	ANÁLISIS DE LOS PROGRAMAS DEL MERCADO EN TIEMPO-REAL	215
7.3.4.1.	OBTENCION PATRONES CARACTERISTICOS .....	217
7.4.	IDENTIFICACIÓN DE CONSUMIDORES CON CAPACIDAD DE APLICAR POLÍTICAS DER PARA PARTICIPAR EN PROGRAMAS DE RESPUESTA DE LA DEMANDA EN TIEMPO REAL .....	222
7.5.	SELECCIÓN DE LAS POLÍTICAS DE RESPUESTA EN TIEMPO REAL	226

7.6.	CONCLUSIONES .....	232
8.	CAPITULO 8 .....	234
8.1.	APORTACIONES.....	236
8.2.	LINEAS FUTURAS .....	239
9.	BIBLIOGRAFIA .....	242
9.1.	BIBLIOGRAFIA .....	244
10.	ANEXO .....	250
10.1.	ARTICULOS PUBLICADOS EN REVISTAS CON INDICE DE IMPACTO 252	



# INDICE DE FIGURAS

Figura 1. Modelos de Monopolio.....	36
Figura 2. Evolución del consumo de energía final.....	44
Figura 3. Curva de precios en el mercado de la electricidad.....	49
Figura 4. Consumo vs. Precio de la energía.....	55
Figura 5. Segmentación de curvas de carga de diferentes consumidores.....	64
Figura 6. Comparación curvas de carga de un segmento determinado con curvas de precios del mercado diario (Day-Ahead).....	65
Figura 7. Cuadro de posibilidades, inconvenientes y soluciones de la participación de pequeños consumidores en los mercados competitivos.....	67
Figura 8. Estructura de un Sistema de Participación de la Demanda.....	68
Figura 9. Aportaciones de la Tesis a la Estructura del Sistema de Participación de la Demanda.....	69
Figura 10. Clasificación General de Consumidores eléctricos.....	74
Figura 11. Esquema detallado de técnicas.....	75
Figura 12. Secuencia de los pasos a seguir para realizar la clasificación de los consumidores.....	77
Figura 13. Proceso de gestión de los datos a utilizar.....	79
Figura 14. Ejemplo de Curvas de Carga normalizadas a valor 1.....	81
Figura 15. Ejemplo de método de agrupación no jerárquico y jerárquico.....	85
Figura 16. Ejemplo de resultado con función discriminante bidimensional.....	90
Figura 17. Ejemplo árboles de decisión.....	94
Figura 18. Ejemplo de resultados obtenidos mediante un SOM de Kohonen.....	96
Figura 19. Clasificación de las técnicas analizadas.....	110
Figura 20. Perfiles de carga del consumidor Industria y del consumidor Universidad.....	119
Figura 21. Perfiles Utilizados para el Test de capacidad de identificación de nuevos consumidores.....	120
Figura 22. Curvas de Carga de los consumidores Industria I1 e I2.....	121
Figura 23. Curvas de carga de los consumidores Industrias I3 e I4.....	122
Figura 24. Curvas de carga de los consumidores Industria I5 e I6.....	122
Figura 25. Curvas de carga de los consumidores Industria I7 e I8.....	122
Figura 26. Curvas de carga de los consumidores Industria I9 e Institucional T1.....	123
Figura 27. Curvas de carga de los consumidores Institucional T2 e Institucional T3.....	123
Figura 28. Curvas de carga de los consumidores Institucional T4 y Comercial C1.....	123
Figura 29. Curvas de Carga de los consumidores Comercial C2 y Comercial C3.....	124
Figura 30. Curvas de Carga de los consumidores Comercial C4 y C5.....	124
Figura 31. Curvas de Carga de los consumidores Institucional T5 y Comercial C6.....	124
Figura 32. Curvas de Carga de los consumidores Comercial C7 y Comercial C8.....	125
Figura 33. Curvas de Carga de los consumidores Comercial C9 y Comercial C10.....	125
Figura 34. Curvas de Carga del consumidor Comercial C11 y Residencial R1.....	125
Figura 35. Curvas de carga utilizadas posteriormente en los Testeos para comprobar la capacidad de identificación de nuevos consumidores de cada una de las técnicas. Curvas de un consumidor Industrial y otro Institucional.....	126
Figura 36. Curvas de carga utilizadas posteriormente en los Testeos para comprobar la capacidad de identificación de nuevos consumidores de cada una de las técnicas. Curvas de un consumidor Comercial y otro Residencial.....	126
Figura 37. Escalado Multidimensional aplicado al primer conjunto de datos.....	131
Figura 38. Aplicación de la Técnica de escalado multidimensional al segundo conjunto de datos.....	132
Figura 39. Mapa obtenido aplicando la técnica de MANOVA al primer conjunto de datos.....	133
Figura 40. Mapa obtenido aplicando la técnica de MANOVA al segundo conjunto de datos.....	134
Figura 41. MANOVA, segundo conjunto de datos (dendograma).....	135
Figura 42. K-means, primer conjunto de datos (dos clusters).....	136
Figura 43. K-means, primer conjunto de datos (silueta, 2 clusters).....	137
Figura 44. K-means, segundo conjunto de datos (8 clusters).....	138
Figura 45. Aplicación de la Técnica de escalado multidimensional a los resultados obtenidos de entrenar el primer conjunto de datos con un red neuronal competitiva.....	140
Figura 46. Aplicación de la Técnica de escalado multidimensional a los resultados obtenidos de entrenar el primer conjunto de datos con un red neuronal competitiva.....	140
Figura 47. Aplicación de la técnica de escalado multidimensional a los resultados obtenidos de entrenar el segundo conjunto de datos con un red neuronal competitiva. Resultado tras solicitar 5 clusters.....	141
Figura 48. Arquitectura del SOM.....	143
Figura 49. Mapa SOM de etiquetas obtenido tras un entrenamiento con el primer conjunto de datos.....	144
Figura 50. Mapa SOM de etiquetas obtenido tras un entrenamiento con el segundo conjunto de datos.....	145
Figura 51. Fuzzy c-means, primer conjunto de datos (2 clusters).....	146



Figura 52. Fuzzy c-means, segundo conjunto de datos (4 clusters).....	147
Figura 53. Fuzzy c-means, segundo conjunto de datos (12 clusters).....	148
Figura 54. Clasificación de consumidores con datos de entrada sin filtrado .....	168
Figura 55. Mapa de Color entrenado con el Usuario tipo 2 .....	169
Figura 56. Mapa de etiquetas obtenido tras el entrenamiento (Universidad).....	170
Figura 57. Mapa de clusters de la Universidad con límite máximo de búsqueda de clusters 2.....	171
Figura 58. <i>Mapa de Clusters</i> con 10 clusters como máximo .....	172
Figura 59. Curvas de carga de los días de consumo normal. ....	173
Figura 60. Curvas de carga de los días anómalos. ....	173
Figura 61. Curvas de carga de los días de consumo algo diferente.....	173
Figura 62. Clasificación de consumidores con datos de entrada sin filtrar y con datos de entrada filtrados. ....	175
Figura 63. Perfiles de Carga de la Universidad y de la Industria .....	175
Figura 64. Pasos a seguir para la extracción de vectores semanales .....	185
Figura 65a&b. Perfiles de carga semanal y vector FWES para los consumidores etiquetados como 2 y 14.....	187
Figura 66. Mapa SOM entrenado con valores en el dominio del Tiempo.....	189
Figura 67 a&b. Mapa SOM entrenado con datos en el dominio de la frecuencia y sus correspondientes perfiles de carga semanal (consumidores etiquetados con 4 y 8) .....	191
Figura 68 a&b. Clasificación de nuevos consumidores .....	194
Figura 69. Mapa entrenado con datos de entrada convertidos en el dominio del tiempo (119 consumidores)....	197
Figura 70. Mapa SOM entrenado y clusterizado con valores en el dominio del tiempo (12 Clusters y 119 Consumidores).....	198
Figura 71. Comparación gráfica de los Perfiles de carga de los consumidores etiquetados como 10, 11, 12 y 13 .....	199
Figura 72. Mapa SOM entrenado y clusterizado con valores en el Dominio de la Frecuencia.....	199
Figura 73. Perfil de carga semanal del consumidor industrial con etiqueta 47 .....	200
Figura 74. Curvas de carga de los consumidores con Etiquetas 13 y 5.....	211
Figura 75. Curvas de carga de dos consumidores con etiquetas 26 y 21 .....	212
Figura 76. Mapa SOM entrenado con las curvas de carga de los consumidores.....	212
Figura 77. Comparación en valores por unidad de las curvas de precio y demanda .....	214
Figura 78. Mapa SOM entrenado con los precios del mercado de gestión de desvíos.....	216
Figura 79. Ejemplo de curvas de precios de la zona 2 .....	217
Figura 80. Precios de los días correspondientes al cluster denominado zona 2 .....	219
Figura 81. Curva de Estimación e índices de confianza para la zona 2 de precios .....	219
Figura 82. Precios de los días correspondientes al cluster denominado zona 3 .....	220
Figura 83. Curva de Estimación e índices de confianza para la zona 3 de precios .....	220
Figura 84. Precios de los días correspondientes al cluster denominado zona 4 .....	221
Figura 85. Curva de Estimación e índices de confianza para la zona 4 de precios .....	221
Figura 86. Curva de carga y usos finales de un consumidor tipo Universidad .....	223
Figura 87. Ejemplo de políticas DER para un consumidor de tipo Centro Comercial.....	224
Figura 88. Curva de carga diaria con y sin política de control de los calentadores eléctricos de agua caliente...225	225
Figura 89. Aportaciones de este capítulo al esquema de la estructura del sistema de participación de la demanda. ....	227
Figura 90. Mapas SOM entrenados con las estrategias DER (Consumidores: Universidad y Centro Comercial) .....	228
Figura 91. Integración de consumidores y agentes comercializadores en el mercado eléctrico mediante el desarrollo y uso de esta metodología. ....	238



## INDICE DE TABLAS

Tabla 1. Apertura del mercado eléctrico minorista .....	40
Tabla 2. Primer conjunto de datos.....	119
Tabla 3. Segundo conjunto de Datos (Distribución de Etiquetas).....	120
Tabla 4. Espectro de consumidores con sus etiquetas asociadas y la actividad a la que pertenecen.....	121
Tabla 5. Resultados, segundo conjunto de datos, test 1 .....	151
Tabla 6. Resultados, segundo conjunto de datos, test 2 .....	153
Tabla 7. Resultados, segundo conjunto de datos, test 3 .....	155
Tabla 8. Resultados, segundo conjunto de datos, test 4 .....	159
Tabla 9. Resumen de las mejores técnicas analizadas.....	160
Tabla 10. Espectro de Consumidores para el entrenamiento de la Red SOM.....	166
Tabla 11. Errores de Cuantización y Topográfico.....	176
Tabla 12. Espectro de consumidores.....	182
Tabla 13. Clusters obtenidos según consumidores.....	190
Tabla 14. Análisis de la calidad de los mapas SOM .....	193
Tabla 15. Resultados del Testeo de Consumidores .....	195
Tabla 16. Nuevo Espectro de Consumidores .....	196
Tabla 17. Espectro de Consumidores .....	211
Tabla 18. Best Matching Units para cada curva de precio de la energía del mercado diario .....	214
Tabla 19. Coeficientes de evaluación de políticas DER del consumidor Universidad.....	225
Tabla 20. Coeficientes de evaluación de políticas DER del consumidor centro comercial.....	225
Tabla 21. Coeficientes de evaluación de políticas DER del consumidor CT Residencial.....	226
Tabla 22. Las Mejores Y Peores Estrategias Seleccionadas Por El Mapa SOM (Consumidor: Universidad)...	229
Tabla 23. Comparación de los Resultados de la SOM .....	231





# **CAPITULO 1**

---

Objetivos y Estructura de la Tesis

*Mh Miguel  
Hernández*



## 1.1. OBJETIVOS DE LA TESIS

El objetivo básico de esta Tesis ha sido desarrollar y evaluar herramientas que permitan clasificar los diferentes segmentos de demanda de energía eléctrica en una determinada zona geográfica –es decir, alimentada por un determinado sistema de distribución- en función de parámetros que pueden ser obtenidos con relativa facilidad por un comercializador o distribuidor. La investigación se ha basado principalmente en las curvas de carga diarias que reflejan el consumo diario de los distintos consumidores. Con esta segmentación se podrá evaluar las posibilidades de participación de la demanda –de cada uno de los sectores que se hayan definido con la metodología desarrollada- en los nuevos mercados de energía eléctrica, y por tanto seleccionar los grupos de usuarios a dónde deban dirigirse los esfuerzos de promoción de políticas específicas. Para definir los segmentos más adecuados, se considera posible la participación de los usuarios en el mercado a través de una serie de políticas y productos que hoy en día sólo se encuentran desarrollados a nivel de experiencias piloto en otros países como EEUU, Noruega, Inglaterra o Australia. Con esta herramienta se pretende contribuir a incentivar la flexibilidad y la eficiencia en la operación y planificación de los sistemas de distribución de energía eléctrica, y también en alguna medida a incentivar la participación de la demanda y la entrada de la generación distribuida en el sistema eléctrico de una forma más racional.

De hecho, analizando las principales barreras de la participación real de la demanda en los sistemas y mercados eléctricos nos encontramos siempre con una serie factores importantes, entre los que destacan las relativas a la generación distribuida y el almacenamiento. Algunas de estas barreras han sido identificadas y estudiadas en determinados proyectos internacionales destacando los estudios realizados por la Agencia Internacional de la Energía.

Las barreras de integración y participación de la demanda pueden clasificarse en:

- Tecnológicas. El coste del uso de tecnología específica puede variar de un país a otro, siendo en algunos casos una barrera. Dentro de la implantación de los sistemas de generación distribuida está la dificultad de conseguir instalaciones de pequeña capacidad a precios razonables. En el caso de la monitorización de las cargas y su implantación en pequeños consumidores residenciales y comerciales es otra de las barreras importantes.
- Económicas. Basada en la rentabilidad que pueda obtener la demanda por su participación. Si no existe rentabilidad sobre todo en los sectores comerciales e industriales el consumidor no participará de una manera activa, no habrá respuesta de la demanda.
- Regulatorias y legales. En este caso también depende de la estructura propia del mercado eléctrico de cada país. Dependiendo de la estructura de mercado esta puede ofrecer diferentes oportunidades a la demanda o restricciones a su participación. También y en la mayoría de los mercados los pequeños consumidores encuentran limitaciones para participar directamente en ellos debido a que existen unos niveles mínimos para poder acceder. En España la posibilidad de los consumidores de permanecer a tarifa regulada, con precios menores a los precios del mercado supone un factor en contra para la participación de la demanda.
- Sociales y culturales. Sobre todo por que el consumidor último sigue percibiendo el suministro eléctrico como un derecho, como un servicio público, y no como un producto donde aprovechar las mejores ofertas.

La hipótesis de partida del trabajo, que se ha contrastado en los estudios preliminares realizados hasta el momento y en la bibliografía científica consultada [1-3], [4], es que la demanda de diferentes usuarios depende de: el tipo general de usuario (residencial, comercial e industrial) [5-7], de su tamaño (pequeño, mediano y grande), de su actividad económica y de factores climatológicos.

La metodología disponible para la resolución de este problema y otros similares es muy variada. Para la segmentación y clasificación de usuarios puede utilizarse desde la clasificación estadística de datos, las técnicas llamadas de “data mining” [8, 9], [10], la utilización de redes neuronales e incluso herramientas de lógica difusa [11].

Una parte importante de las técnicas investigadas en la presente tesis se enmarcan dentro de la denominada Inteligencia Artificial, concretamente en las redes neuronales y en la lógica borrosa. No obstante, también se han considerado a nivel de comparación de resultados de simulación y dentro de la investigación técnicas clásicas de agrupación y segmentación de datos. Específicamente, los objetivos que se han marcado en la realización de esta tesis son los siguientes:

1. Determinación del estado del arte de este problema dentro del contexto de la respuesta de la demanda y la segmentación de consumidores. Esta respuesta o participación de la demanda se basa en la decisión y capacidad de consumir o no según el precio de la energía y la posibilidad de obtener beneficios de los precios de la electricidad. El consumidor decidirá consumir si los beneficios obtenidos de este consumo son superiores al precio a pagar. Entre las posibles opciones para esta participación se encuentra la utilización de contratos de riesgos compartidos, la participación directa en el mercado y la participación a través de terceros. La participación directa consiste en acudir a los mercados energéticos con una serie de ofertas de compra o de venta. Las posibilidades exactas dependen de la estructura del mercado. La participación a través de terceros consiste en un proceso en el que un agregador o comercializador participan en el mercado como agente con las ofertas de compra y de venta conseguidas a través de productos estructurados que han ofrecido a los consumidores. En esta última situación el agregador necesita de herramientas capaces de identificar segmentos de demanda a través de la agrupación de consumidores eléctricos con comportamientos similares con la finalidad de lograr un nivel de agregación de potencia lo suficientemente alto como para poder participar en determinados mercados.

2. Revisión de las diferentes metodologías de clasificación de datos: es necesario conocer las herramientas susceptibles de ser utilizadas en este problema para poder determinar las entradas de datos, así como las posibles limitaciones de cada una de ellas. El objetivo de la investigación no será limitarse al análisis de técnicas clásicas de tratamiento de datos, sino justificar y profundizar en la aplicación de herramientas de inteligencia artificial y concretamente en las redes neuronales.
3. Recopilación de datos. Es necesario definir qué tipo de información debe estar disponible para cada uno de los usuarios y si es posible obtenerla de las empresas distribuidoras-comercializadoras de energía eléctrica e incluso de otros suministradores de energía primaria. También es necesaria una selección de un conjunto de parámetros que defina a los diferentes tipos de usuarios. Fundamentalmente se tratará de parámetros de consumo de energía, actividad o ubicación geográfica.
4. Segmentación, clasificación y agregación de usuarios. En esta etapa se pretende definir para ciertos segmentos de la demanda unos comportamientos típicos del usuario en cada sector. Se aplicarán diferentes técnicas para tratar las curvas y comportamientos estándar.
5. Aplicación de políticas de participación de la demanda. Una vez definidas las curvas tipo de algunos segmentos de usuarios abordada en el punto anterior, se aplicarán a los mismos políticas de participación (productos de mercado a corto y medio plazo). Se analizará la flexibilidad de la demanda de ciertos usos finales de la energía en estos usuarios, es decir el cambio de la curva de consumo.



## 1.2. MOTIVACION

Uno de los principales problemas a los que se enfrentan las empresas comercializadoras, es el tratamiento de la gran cantidad de información y datos asociados a los consumidores. Estos datos no pueden ser analizados de una manera fácil y eficaz, de modo que se hacen necesarias herramientas avanzadas para el tratamiento y sintetización de esta información.

Dentro del contexto actual de los mercados eléctricos competitivos la participación de la demanda pasa en muchos casos por la participación de los consumidores en programas ofrecidos por empresas comercializadoras. La segmentación de usuarios en clusters, a partir de datos característicos de cada consumidor se plantea como una herramienta fundamental para estas empresas. Es posible así, aplicar distintas políticas de precios o programas de gestión de la demanda en función del número y características de los consumidores que formen cada uno de estos clusters.

Una primera agrupación de consumidores podría basarse en características propias del sector o consumidor, como la actividad de una empresa, el tamaño de la misma, su localización geográfica su volumen anual o mensual de consumo de energía eléctrica, etc..

Profundizando en la característica principal que diferencia a un consumidor eléctrico de otro, nos encontramos con su curva de carga diaria, tal como refleja el Dr. Gianfranco Chicco [12]. Es decir, la evolución horaria de la cantidad de energía que consume en cada hora a lo largo de un día. Este patrón de comportamiento nos permite identificar y diferenciar consumidores que pueden o no pertenecer a sectores o actividades similares. La curva de carga de un consumidor nos permitirá conocer en que momentos del día su consumo es más acusado y en que otros momentos su consumo es más contenido. Utilizando técnicas de agrupación de datos es posible agrupar distintos espectros de tipos de consumidores en grupos con patrones de consumo similares. Esto a su vez permitirá conocer que consumidores son los más adecuados para ofrecerles en cada momento

políticas, estrategias y mejores precios de la energía en función de la capacidad de modificación de su curva de carga.

Otro factor motivador ha sido la evolución de los procesos de liberalización del mercado de la energía en algunos países. Estos no han tenido el éxito que se esperaba en un principio, debido principalmente a los diversos problemas que vienen ocurriendo desde el año 2000. Por ejemplo, los fallos en el suministro eléctrico (ocurridos en Europa, Estados Unidos y Canadá durante el año 2003 han sido el máximo exponente, por su repercusión económica y social. Debido a estas experiencias los operadores de los sistemas eléctricos creen, cada vez más, que los recursos de Generación Distribuida deberían actuar a través de un proceso de integración que tuviera en cuenta algunas políticas de demanda, como es la mejora de la eficiencia del consumo a largo plazo, o políticas de respuesta a precios en el corto y medio plazo.

La contribución efectiva de estos programas, y el desarrollo de nuevos productos para ofrecer a los medianos y pequeños consumidores, no sólo a los grandes, necesita de un conocimiento detallado del potencial de los distintos consumidores a través de la definición de segmentos de demanda, el conocimiento de la capacidad de respuesta de la misma, la agregación de consumidores, y finalmente la caracterización de estos clusters de demanda.

### **1.3. METODOLOGIA**

La metodología que se propone para alcanzar los objetivos mencionados persigue el análisis de técnicas de segmentación aplicables a la clusterización de consumidores eléctricos con el fin de establecer una herramienta eficaz de agrupación de usuarios con patrones de consumo similares, así como, capaz de identificar y clasificar posteriormente nuevos consumidores eléctricos dentro de estos segmentos. Por último, se investigará la aplicación de la herramienta elegida y sus resultados de segmentación al desarrollo de una metodología de participación de la demanda en mercado competitivos.

Antes de aplicar una técnica de segmentación y clasificación de consumidores eléctricos es necesario realizar un filtrado de días anómalos a los datos de entrada de cada uno de los consumidores que forman el espectro a estudiar, estos resultados fueron recogidos en una importante aportación al 7th IASTED International Conference on Power and Energy Systems [13]. Cada consumidor suele tener una curva de carga diaria característica que define el comportamiento de su consumo normal. En largos periodos, como puede ser un año, es posible encontrar un comportamiento típico o usual, siempre y cuando se logren aislar curvas de días con consumos irregulares o anómalos. Por ejemplo periodos de vacaciones, mantenimientos, días festivos, etc... hay un capítulo donde se recogen los experimentos realizados y las conclusiones al respecto.

Para el estudio se han utilizado consumidores eléctricos pertenecientes a la provincia de Alicante principalmente. Los datos de partida han sido las curvas de carga diaria (potencia activa consumida) de cada uno de los consumidores obtenidas en un periodo concreto. A lo largo de la investigación, y dentro de los distintos espectros a segmentar se han incluido empresas de varios tamaños, consumos de zonas residenciales medidos desde el centros de transformación, hoteles, universidades, almacenes, etc..

Una vez conocidos los segmentos o agrupaciones de usuarios para un espectro de datos analizado, es posible establecer estrategias de control de cargas con la finalidad de definir una metodología de participación de la demanda en mercados competitivos. Los parámetros más importantes en la segmentación vienen dados por los valores de las curvas de carga, como por ejemplo son el nivel máximo de potencia y la forma de la curva a lo largo del día dentro de los periodos de tiempo considerados (invierno o verano pero sin considerar vacaciones ni días festivos) [14]. Aquellos grupos o segmentos de consumidores, cuyos perfiles de consumo tengan picos fuertes de demanda cuando los precios de la energía están muy altos, son a priori los primeros en plantearse la posibilidad de aplicar estrategias de control de cargas con el objetivo de reducir el consumo en los momentos de precios altos e intentar pasar este consumo a las zonas de consumo de menor demanda y con precios más bajos. También cabe la posibilidad de que pequeños y medianos consumidores acogidos a programas de gestión de la demanda ofrecidos por

empresas comercializadoras o agregadoras aprovechen estas reducciones para recibir algún tipo de compensación económica. Esto se debería principalmente a la capacidad que tendrían estas empresas de participar en el mercado eléctrico ofreciendo paquetes de reducción de niveles de potencia considerables.

Por otro último destacar que la aplicación de una herramienta de segmentación eficaz puede ayudar a alcanzar dos objetivos claros:

- La identificación de clusters de consumidores con curvas de potencia alta cuando los precios en el mercado de la energía son también muy altos. Es decir, curvas de carga diaria que mantienen un perfil o patrón similar a las curvas de precio de la energía en el mercado diario a 24 horas.
- La identificación de consumidores a los que les conviene una reducción de carga cuando los precios estén altos. Es decir, aquellos consumidores que aun teniendo un patrón de consumo similar al de precios del mercado diario les convendría reducir o modificar su consumo porque sus picos de consumo coinciden con los picos de la energía no solo en el mercado diario sino también en el mercado intradiario, en el caso del mercado eléctrico Español. La finalidad sería comprar a precios del mercado diario y vender energía (por reducción de consumo, disponibilidad) a los precios del mercado intradiario, con la consiguiente ganancia económica.

## **1.4. ESTRUCTURA DE CAPITULOS**

A continuación se presenta brevemente la forma en la que se estructura el presente trabajo, comentando los contenidos y alcance de cada uno de los capítulos en los que se ha dividido. En este primer capítulo se han planteado los objetivos y el contexto general de la Tesis. El capítulo 2 recoge una introducción a los mercados eléctricos [3], [15, 16], y una revisión de los conceptos y definiciones más usuales dentro del funcionamiento de estos

mercados [17], tales como el Demand-Side [18], las perspectivas futuras de los mercados liberalizados.

En el capítulo 3 se realiza una recopilación y una exposición de distintas técnicas de clasificación de Datos. Basándose en una amplia revisión bibliográfica de estas técnicas. Sirviéndonos de la revisión de técnicas del capítulo 3 como primera aproximación a la clasificación y agrupación de datos.

En el capítulo 4 comienza una investigación todavía más metódica que llevará al autor, tras un proceso recogido a lo largo de los siguientes capítulos, a desarrollar las conclusiones y principales aportaciones de esta Tesis. Este capítulo contiene un estudio y análisis de distintos métodos de segmentación de consumidores eléctricos a partir de sus curvas de carga diaria.

El capítulo 5 recoge los resultados de la investigación llevada a cabo con los mapas SOM al emplearlos como herramienta para el filtrado de días anómalos a partir de un conjunto de curvas de carga de dos consumidores. Este paso es necesario y previo a la segmentación para conseguir una correcta agrupación, y así lo han demostrado los experimentos desarrollados y resultados recogidos en este capítulo.

El capítulo 6 está centrado en el análisis de la calidad de segmentación de consumidores eléctricos mediante el empleo de Mapas Autoorganizados entrenados con diferentes tratamientos de los datos de entrada. Se ha pretendido medir la sensibilidad de los Mapas frente a diferentes tratamientos. Para ello las curvas de carga de los diferentes consumidores que forman el espectro de entrada a la red han sido tratadas y convertidas en valores en el dominio de la frecuencia y valores en el dominio del tiempo.

Una de las aportaciones más importantes de la tesis viene recogida en el capítulo 7, y es la aplicación de los mapas Autoorganizados de Kohonen como herramienta de ayuda en la toma de decisiones en mercados eléctricos. Por último el capítulo 8, recoge un resumen con las conclusiones, aportaciones y líneas futuras de investigación.



# CAPITULO 2

---

Mercados eléctricos, la demanda y la segmentación de consumidores eléctricos.







## 2.1. INTRODUCCIÓN

La electricidad es una forma de energía imprescindible y con infinidad de usos, debido a su gran versatilidad y controlabilidad, a la inmediatez en su utilización y a la limpieza en el punto de consumo. La energía eléctrica podría considerarse como un bien de consumo más, que se produce, se transporta, se comercializa y se consume, pero con características especiales por no ser almacenable, debe producirse y transportarse en el mismo momento en que es consumida. Su transporte no se puede dirigir por caminos específicos, sino que la energía eléctrica fluye por las líneas u otras instalaciones de acuerdo a unas leyes concretas de la física. Dichas leyes imponen además una estrecha interdependencia entre las distintas vías de transporte de forma que cualquier perturbación en una línea u equipo de transporte provoca efectos colaterales significativos e inmediatos en los demás caminos alternativos. La electricidad “se inyecta” en la red o “se extrae” de la red en múltiples puntos, pero no es posible establecer una relación bilateral entre lo inyectado en un punto y lo extraído en otro.

Situándonos en el lado de la Demanda, lo que en muchos artículos y libros internacionales se conoce como “*Demand-Side*”, el modificar y flexibilizar la curva de consumo ha sido siempre una medida bien reconocida y aceptada como forma eficaz de controlar y reducir el coste de producir y distribuir electricidad [18].

El hecho de que las compañías eléctricas, bien distribuidoras o comercializadoras animen a sus consumidores a modificar su patrón de demanda les permite obtener una serie de beneficios mutuos (compartidos con los consumidores) al adoptar medidas basadas en la creación de diferentes tarifas y en la creación de contratos de interrumpibilidad de carga.

El problema más evidente desde la implantación de estas medidas en muchos países, ha sido el hecho de que las ventajas o beneficios obtenidos por las compañías comercializadoras no es proporcional a las ventajas o beneficios obtenidos por los consumidores.

Hay que hacer una distinción entre los efectos a largo plazo y a corto plazo de los precios de la energía. A largo plazo, el precio medio afectará al nivel total del consumo. Las amplias diferencias de precios entre el día y la noche o entre los fines de semana y los días laborables también pueden

animar a clientes a instalar sistemas de almacenamiento térmico para ayudarles a evitar consumir electricidad durante las horas de precios máximos. En muchos casos la estimación de tales inversiones y los efectos a largo plazo harán que los consumidores se planteen la eficacia de estos sistemas.

Sin embargo a corto plazo los consumidores tienen la capacidad de reducir o renegociar su demanda en respuesta a los precios de la electricidad. Por ejemplo, si los precios son altos, algunos consumidores de tipo industrial pueden renunciar a la producción si no es lo suficientemente beneficiosa, es decir, si se realiza a costa de consumir energía cuando los precios están realmente altos, disparando los costes de producción. En la bibliografía nos podemos encontrar el caso de Schweppe y su equipo de trabajo [19] donde los consumidores que tienen la capacidad de almacenar la energía o algún producto intermedio pueden reorganizar su producción. Ellos propusieron un sistema donde los clientes ajustarían su demanda por encima o por debajo dependiendo del precio marcado para la transacción. Los precios para la transacción estarían actualizados en tiempo real para tener en cuenta estos ajustes de carga. Schweppe consiguió demostrar que la volatilidad en los precios ayuda a aumentar el valor de las instalaciones de almacenaje por ser estas de gran ayuda en periodos de precios altos. Esto animaría a algunos consumidores a invertir más dinero en tales instalaciones. El trabajo de Schweppe se consideró como una base teórica para la liberalización del comercio de la electricidad. Sin embargo, ninguno de los mercados competitivos que han puesto en práctica estas medidas han ofrecido oportunidades reales y ventajosas a los consumidores para reducir su demanda en respuesta a los precios marcados para dichas transacciones, principalmente porque esta respuesta puede afectar a los precios.

## 2.2. LOS MERCADOS ELÉCTRICOS

### 2.2.1. ORIGENES Y EVOLUCIÓN

En sus comienzos, a finales del siglo XIX, la industria eléctrica era un sector fragmentado, de carácter local, privado en su mayor parte y, en muchos países, no sujeto a regulación [20]. A principios del siglo XX, los gobiernos empezaron a considerar el suministro eléctrico como un servicio público, y la electrificación llegó a las zonas rurales. Tras la Segunda Guerra mundial, comenzó la aplicación de conceptos económicos a la industria eléctrica, a la vez que aumentó el tamaño mínimo eficiente de las centrales de generación y se completaron las redes de transporte y distribución en los países industrializados. El aumento de las economías de escala hizo que muchas pequeñas empresas eléctricas dejaran de ser rentables. La necesidad de elevadas inversiones obligó a numerosos gobiernos a unir a las pequeñas compañías en una única compañía nacional con carácter de monopolio o, al menos, a crear varias compañías regionales de carácter monopolístico. Es decir, los gobiernos empezaron a considerar al sector eléctrico como un monopolio natural. Australia, Nueva Zelanda y numerosos países europeos, con la excepción de España, optaron por un modelo de monopolio público tras la nacionalización del sector.

Por lo tanto, el modelo de monopolio, ya fuera de propiedad pública o privada, ha sido el comúnmente adoptado en todo el mundo antes de la introducción de medidas liberalizadoras. En el modelo de monopolio no existe competencia ni capacidad de elección por parte del consumidor final, del que se dice entonces que permanece cautivo. Según los países, puede estar caracterizado por una única compañía vertical y horizontalmente integrada (que será entonces de carácter público) o por varias compañías verticalmente integradas con zonas de operación definidas (de propiedad pública o privada). También admite variantes en cuanto a la integración vertical, existiendo casos, como EEUU antes del proceso de liberalización, en los que la actividad de distribución es realizada en régimen de monopolio por compañías independientes de las que realizan las actividades de generación y transporte. La Figura 1 muestra esquemáticamente diversas estructuras de monopolio.

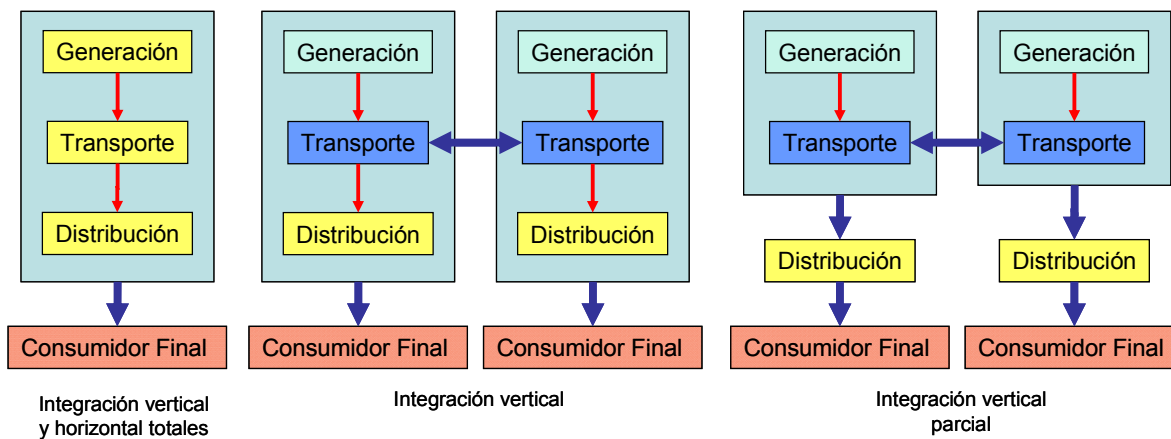


Figura 1. Modelos de Monopolio

Durante las décadas de los 80 y los 90, diferentes factores motivaron las reformas liberalizadoras del sector eléctrico en diversos países. Pueden clasificarse estos factores en tres tipos [21]:

- **Económicos:** Aumento del tamaño relevante de los mercados energéticos, motivado por la globalización de los productos energéticos y el desarrollo de la capacidad de interconexión de los sistemas eléctricos y gasistas.
- **Técnicos:** Disminución del tamaño mínimo eficiente en generación, gracias al desarrollo de la turbina de gas de ciclo combinado. Además, la reducción del coste de medida y control gracias al desarrollo de las tecnologías de información facilita el suministro descentralizado.
- **Políticos:** En los países desarrollados, la búsqueda de la eficiencia económica incentiva la creación de un marco competitivo en el sector eléctrico. En los países en vías de desarrollo, la creación de competencia se veía como una forma de obtener recursos provenientes de la privatización y/o de atraer nuevas inversiones.

Existen tres modelos de organización alternativos al modelo de monopolio que permiten la existencia de competencia [22], aunque cada uno de ellos con posibles variaciones. Elegir uno de estos modelos es el primer paso en el diseño de un sector eléctrico en competencia. Adicionalmente, también pueden verse como modelos secuenciales que, partiendo del modelo de monopolio, introducen progresivamente mayor competencia.

### 2.2.1.1. MODELO DE MONOPSONIO O COMPRADOR ÚNICO

En este modelo el comprador único, típicamente la empresa eléctrica local, realiza las compras de energía a las centrales disponibles, de forma que sólo existe competencia en generación. En el modelo aún perviven una o varias compañías verticalmente integradas pero se permite la existencia de generadores independientes con el objetivo de atraer la inversión. El agente comprador actúa como un monopolio en las actividades de transporte y distribución, permaneciendo cautivos los consumidores finales.

Los principales inconvenientes del modelo son el posible trasladado del riesgo a los consumidores cautivos a través del comprador único y la posible discriminación del comprador único a los generadores independientes frente a los generadores de su propiedad, habitualmente pública. En general, puede decirse que el modelo de comprador único es un buen modelo transitorio y puede ser definitivo en países en vías de desarrollo en los que proponer un modelo de mercado más desarrollado no tenga sentido [23].

### 2.2.1.2. MODELO DE COMPETENCIA MAYORISTA

Este modelo permite la competencia en dos ámbitos. Por un lado, las empresas generadoras compiten entre ellas y con posibles nuevos entrantes por producir la energía necesaria, lo que mejora la eficiencia económica, dejando el riesgo de las decisiones de inversión al libre mercado. Por otro, las comercializadoras a consumidores cautivos (habitualmente las distribuidoras tradicionales) mantienen el carácter de monopolio local, pero ahora pueden comprar libremente a los generadores. La compraventa de energía suele establecerse total o parcialmente a través de un mercado organizado por el Operador del Mercado Eléctrico. En España este operador se llama OMEL [24] y asume las funciones necesarias para realizar la gestión económica referida al eficaz desarrollo del mercado de producción de electricidad y ejerce las mismas respetando los principios de transparencia, objetividad e independencia.

### 2.2.1.3. MODELO DE LIBERTAD DE ELECCIÓN PARA TODOS LOS CONSUMIDORES

Este es el modelo existente en España desde el 1 de enero de 2003. Es similar al anterior, sólo que la competencia se ha introducido ya en todos los niveles del sector, no existiendo consumidores cautivos. Es habitual que la apertura del sector a la competencia sea de forma gradual y que los últimos consumidores en tener capacidad de elección sean también los de menor consumo. Idealmente y con la regulación adecuada que minimice las imperfecciones del mercado, este modelo optimiza la eficiencia económica. Como en el modelo anterior, este modelo implica la desaparición de la posibilidad del Estado de satisfacer ciertas políticas sociales utilizando los ingresos de generación, algo completamente lógico en un contexto de libre mercado.

### 2.2.2. LA LIBERALIZACION COMO PROCESO

En 1979, Margaret Thatcher pensó que los planes de privatización de su partido eran un tema demasiado controvertido para mencionarlo durante la campaña electoral. Veinte años más tarde, lo dan por hecho los gobiernos y líderes de opinión, y en más de cien países ya se han privatizado varias empresas públicas [25].

La liberalización ha presenciado como la meta de contar con “precios confiables y universales en el servicio eléctrico a un precio razonable” es remplazada por la meta de “lo que el consumidor elija” [26]. Sin embargo, esa elección del consumidor se refiere a un rango muy reducido de proveedores minoristas en el que todos ofrecen el mismo producto, “*la electricidad será un artículo*” producido, transportado y vendido por empresas privadas que usan su influencia para elevar al máximo sus ventas y expansión, o bien, si la electricidad debe ser considerada como un servicio no lucrativo regido por las políticas que buscan el menor costo, eficiencia y la satisfacción de necesidades de energía eléctrica no fomentadas ni empleadas [27].

El término desregulación resulta algo inapropiado, debido a que los cambios que pretende designar en realidad no consisten en deshacerse de reglamentaciones, sino en sustituir las que

protegen al público y al ambiente por otras que aseguren un funcionamiento sin tropiezos en el mercado y el sistema eléctrico. “Privatización” es el término más exacto, puesto que es lo que está sucediendo en el caso de la desregulación consiste en la privatización del control sobre el suministro eléctrico. El uso de este término lo avala el Departamento de Energía de Estados Unidos, al decir: “Tratamos la privatización [...] como a cualquier otro movimiento que disminuye la propiedad y el control público, e incrementa la propiedad y el control privado” [28].

En Europa la Comisión Europea ha estado presionando para lograr la liberalización de la electricidad. En diciembre de 1996 aprobó una directiva que requería que las naciones integrantes abrieran sus mercados a la competencia a partir de 1999, con el propósito de crear un solo mercado eléctrico en todo el continente europeo. La directiva también demandaba que los países permitieran que las empresas privadas aportaran instalaciones generadoras y que dieran acceso a sus redes de transmisión y distribución sin discriminación alguna [29]. Algunos países incluso se adelantaron a los requisitos de la Comisión. Los beneficios que prometía esta competencia se fundamentaban en los supuestos neoliberales de que mejoraría la eficiencia y el servicio, se promovería la innovación y los precios disminuirían. La prioridad de la Comisión Europea le da a la competencia es evidente en sus documentos, en los cuales exhorta a las naciones miembro a “favorecer los argumentos de las políticas públicas”, como es el caso de la protección medio ambiental, siempre y cuando esto no ponga límite al proceso de liberalización ni “*restringa el comercio y la competencia más de lo necesario*”.

Algunos países, entre los que sobresale Francia, resistieron las presiones de otras naciones de la Unión Europea y de líderes comerciales dentro de Francia para privatizar su electricidad. Sin embargo, en 2002, con la elección de un gobierno de centro-derecha, el gobierno francés finalmente aceptó abrir sus mercados y privatizar parcialmente Electricité de France. La privatización parcial es mencionada por los ministros de gobierno como “apertura de capital” para no verse afectados por los votantes que relacionaban la privatización con el liberalismo económico “anglosajón”. La privatización parcial fue recomendada por un equipo de investigadores especializados de derecha, de la Foundation Concorde, en un informe que le encargó el presidente Jacques Chirac [30].

El la Tabla 1 se muestra a nivel europeo la apertura del mercado eléctrico al mercado minorista [31], [32]. Los datos que recoge esta tabla han sido obtenidos de un informe de la Comisión Europea del año 2002.

Tabla 1. Apertura del mercado eléctrico minorista.

<b>País</b>	<b>Apertura mercado minorista</b>	<b>Apertura completa mercado minorista</b>
Austria	100%	2001
Australia	Algunos estados	No hay fecha
Bélgica	35%	2007
Dinamarca	90%	2003
Finlandia	100%	1997
Francia	30%	No hay fecha
Alemania	100%	1998
Grecia	30%	No hay fecha
Irlanda	30%	2005
Italia	45%	No hay fecha
Holanda	33%	2003
Nueva Zelanda	100%	1994
Noruega	100%	1991
Portugal	30%	No hay fecha
España	54%	2003
Suecia	100%	1998
Reino Unido	100%	1998
Estados Unidos	Algunos estados	No hay fecha

### 2.2.3. CARACTERÍSTICAS DEL CONSUMO ELÉCTRICO

Los índices de consumo eléctrico constituyen uno de los elementos más indicativos del desarrollo industrial de un país. El grado de consumo eléctrico per cápita y sobre todo el nivel de electrificación de un país son claras señales del nivel de bienestar. De los seis mil millones de personas que habitamos la Tierra, la tercera parte no tiene acceso a la electricidad. Sin embargo los países en desarrollo evidencian una cierta saturación en el crecimiento de la demanda eléctrica pero



en ningún caso un estancamiento. En estos mercados y desde hace unos años ha tomado especial relevancia la denominada “*Gestión de la demanda*”, entendida como todas aquellas técnicas y acciones encaminadas a racionalizar y aprovechar mejor el consumo de energía eléctrica. Uno de los objetivos de la Tesis es aportar herramientas y técnicas dentro de una metodología de participación de la demanda en los mercados eléctricos con la finalidad de mejorar la eficiencia de estos desde el punto de vista de minimizar los costes del consumo. La participación activa de la demanda en los mercados mejoraría la operación del sistema eléctrico, y conseguiría un ahorro importante en nuevas inversiones.

A continuación se exponen algunos de los principales factores que dependen del consumo:

- Elasticidad de la demanda: Es entendida como el porcentaje de variación del consumo respecto a una variación unitaria del precio de la electricidad. La demanda eléctrica en general se caracteriza por presentar una elasticidad muy reducida en el corto plazo (reacciona poco al precio). Ver la siguiente fórmula.

$$Elasticidad = \frac{\Delta P / \Delta P_i}{\Delta C / \Delta C_i}$$

- Clima: Este factor es muy importante, debido a que condiciona en gran medida los picos de demanda que se producen a lo largo del año coincidiendo con las diferentes estaciones climáticas. En las épocas invernales derivados principalmente de los altos consumos en calefacción en los días de frío intenso y en épocas estivales derivados de los picos de consumo eléctrico para aire acondicionado y climatización.

Como principales parámetros que caracterizan el consumo se indican a continuación los siguientes:

- Potencia: Media, Máxima, Mínima, factor de carga (Media/Máxima)

- Energía: La energía eléctrica no se almacena y toda instalación debe estar preparada para soportar la máxima potencia que a lo largo de todo el ciclo de consumo va a demandar la instalación.
- Perfil de consumo: Se suele definir también como curva de carga (normalmente los perfiles de consumo se consideran diarios, 24 horas), y representa la potencia consumida en función del tiempo. Los perfiles pueden ser muy planos, con una curva de carga bastante constante en el tiempo, siendo por regla general más baratos de cubrir, o perfiles de consumo variables en el tiempo más caros de cubrir.
- Curva monótona de carga: Representa el valor de la potencia demandada superado para una duración determinada. Su integral representa la energía consumida para el horizonte de tiempo contemplado.

Como factores técnicos adicionales nos podemos encontrar:

- Consumo de potencia reactiva (factor de potencia)
- Condiciones de alimentación: como la tensión y la frecuencia.

Otra de las características del consumo es su sensibilidad a las condiciones técnicas con las que se alimenta eléctricamente. La precisión, la calidad, las prestaciones y el correcto servicio que prestan muchos de los equipos eléctricos dependen de la calidad de la onda de tensión que les alimenta. Los países desarrollados le prestan cada vez más atención a la calidad del suministro, exigiendo a las empresas que proporcionan el servicio una mejor calidad y una mayor fiabilidad. A continuación se exponen algunos de los aspectos que caracterizan y definen la calidad técnica del suministro eléctrico [33], [34], [35]:

- Los cortes de suministro tanto en número, como en duración.
- Las caídas de tensión tanto en valor, como en duración.
- Los armónicos en las ondas de tensión, en número y en valor.
- Las sobretensiones.

Ya se ha comentado anteriormente, la necesidad de la demanda de ser servida instantáneamente de forma que el sistema de alimentación eléctrico debe estar preparado para detectar y responder de forma inmediata a estas variaciones. Para ello el sistema tiene que estar provisto de sofisticados elementos de medida, control, supervisión, y debe tener siempre disponible generación de reserva lista para producir. Ahora bien el uso de esta reserva condiciona de manera importante el precio final de la energía en los mercados eléctricos.

## 2.2.4. CONTEXTO ECONOMICO

### 2.2.4.1. EL SECTOR ELECTRICO Y LA ACTIVIDAD ECONÓMICA

A continuación se muestra una lista de fechas y datos representativos de la evolución del sector eléctrico español en la última década:

- En 1999 la energía eléctrica utilizada en España era igual al 19% del total.
- En el periodo de 1995 a 1999 las inversiones de los principales grupos empresariales eléctricos representaron el 5,2% de la formación bruta de capital fijo total de la economía.
- En el año 1998 el 74% de la energía eléctrica fue consumida por los sectores productivos de la economía.
- En el año 1999 los ingresos estimados de las empresas eléctricas fueron de 10.760 M€ de tarifas reguladoras + 1.200 M€ de peajes a consumidores = 11.960 M€ en total.
- En el periodo de 1996 a 1999 la factura eléctrica Española fue igual al 2,66% del gasto total de las familias.

La Figura 2 muestra la evolución del consumo de energía final desde 1994 hasta el 2004:

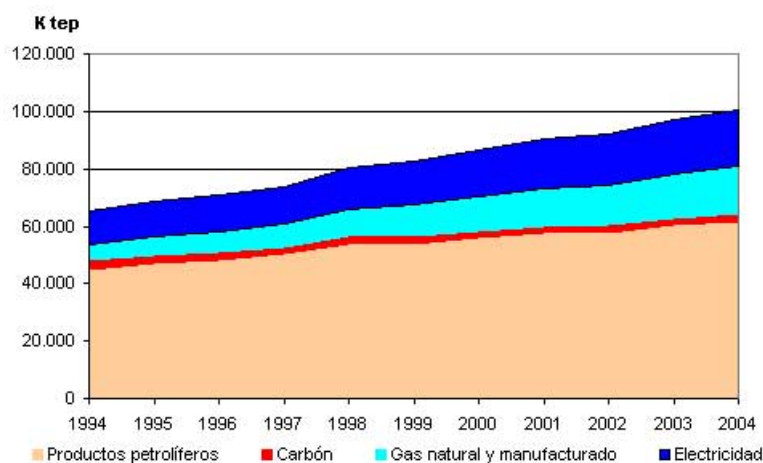


Figura 2. Evolución del consumo de energía final.

## 2.2.5. CONTEXTO REGULADOR

### 2.2.5.1. REGULACIÓN TRADICIONAL Y REGULACIÓN EN COMPETENCIA

Una forma de definir o explicar el concepto de Regulación del mercado es: “Sistema que permite a un gobierno formalizar e institucionalizar sus compromisos de proteger a los consumidores e inversores en un determinado sector industrial”. La regulación tradicional corresponde a un servicio prestado en condiciones de monopolio. En la mayoría de los países se han establecido o están inmersos en el proceso de establecer también marcos regulatorios de libre competencia.

Entre las características tecnológicas y económicas de la industria eléctrica que condicionan el diseño de los planteamientos regulatorios que le son aplicables están las siguientes:

- Las infraestructuras necesarias para producir, transportar y distribuir electricidad tienen un coste muy elevado, son específicas y muy duraderas.
- La energía eléctrica tiene el carácter de esencial para los consumidores, por lo que existe una gran sensibilidad social ante su eventual carencia o falta de calidad.

- La energía eléctrica no es almacenable económicamente en cantidades significativas, por lo que su producción ha de estar adaptada instantáneamente a su demanda.
- La operación real de un sistema eléctrico es el resultado de una compleja cadena jerarquizada de toma de decisiones.
- El suministro de electricidad integra actividades con claras condiciones de monopolio natural (los servicios de red de transporte y de distribución o la operación del sistema) con otras que admite la competencia (la generación y la comercialización).
- Las estructuras de organización y de propiedad de las empresas eléctricas en distintos sistemas eléctricos son muy diversas.

#### 2.2.5.2. LA NUEVA REGULACION ELECTRICA

El marco de regulación eléctrica parte del concepto de que es posible la existencia de un mercado mayorista de energía eléctrica al que todas las entidades generadoras así como todas las entidades consumidoras puedan acudir. Este marco regulatorio ha de ocuparse de muchos aspectos, entre los que destacan:

- La creación de un mercado minorista que permita a los consumidores ejercer el derecho a elegir la compañía suministradora.
- De definir los mecanismos e instituciones de coordinación de los mercados organizados y, de forma particular, de la operación técnica del sistema.
- Las redes de transporte y distribución en sus diversos aspectos de acceso.
- El diseño de un proceso de transición del marco tradicional al de competencia que proteja los intereses legítimos de consumidores y empresas.

Dentro de este marco regulatorio basado en un mercado de la energía eléctrica donde puedan acudir no solo entidades generadoras sino también entidades distribuidoras, comercializadoras y consumidores finales, es donde las empresas de venta de energía han de ofrecer a sus clientes nuevas ofertas, tarifas y programas adaptados a sus patrones de consumo. Es aquí, donde la segmentación de consumidores eléctricos, juega un papel crucial. El uso de herramientas eficaces que permitan agrupar y segmentar consumidores eléctricos bajo patrones y modelos de consumo similares podría

permitir la obtención de beneficios en forma de ahorros de costes en el caso de los consumidores y de beneficios en las operaciones de planificación de compra-venta de paquetes de energía por parte de distribuidoras y comercializadoras.

#### Transacciones en el contexto del mercado mayorista

A nivel mayorista los generadores, los consumidores autorizados y las entidades comercializadoras de cualquier tipo pueden realizar libremente transacciones entre sí, ya sea a través de un mercado spot o por medio de contratos. Los intercambios internacionales constituyen un caso particular de transacción en el mercado mayorista, en donde la novedad es el tratamiento a dar a los agentes externos, en general sometidos a esquemas regulatorios diferentes y con un nivel distinto de apertura a la competencia.

#### Transacciones en el contexto del mercado minorista

La comercialización a consumidores sin capacidad de elección: es una actividad regulada, con retribución según costes reconocidos, y sujeta a estándares de calidad de servicio en la atención al cliente. La comercialización a consumidores con capacidad de elección: es una actividad no regulada, cuyo cometido esencial es la gestión del riesgo.

## **2.2.6. EL MERCADO ELÉCTRICO ESPAÑOL**

### 2.2.6.1. SITUACION ACTUAL

El mercado de producción de energía eléctrica, ofrece a los agentes participantes en el mismo la posibilidad de contratar energía eléctrica en siete sesiones: la primera y principal del mercado diario y seis sesiones posteriores del mercado intradiario, distribuidas a lo largo de las 24 horas del día. Las liquidaciones, los cobros y los pagos, referidos a los precios finales del mercado por los suministros y producciones realizadas y por tanto, resultado de las transacciones en los mercados diario e intradiario, procesos de gestión técnica y de las mediciones, se han producido durante el año 2005 conforme a las Reglas de Funcionamiento del Mercado. Desde febrero de 1999 se presta a los agentes el servicio de facturación mensual de la energía comprada y vendida a través de OMEL [24], que ha sido adaptado desde 2003 a las solicitudes de los productores en régimen especial, con la

doble modalidad de liquidación por agente vendedor agregada o individual, cuando éste actúa exclusivamente como representante en la presentación de ofertas.

El avance registrado en la recepción e incorporación de medidas de la generación de electricidad ha permitido que los datos de medida de los productores participantes en el mercado estén disponibles al día siguiente al del suministro, posibilitando así que la liquidación con medidas de estos agentes esté disponible con carácter mensual.

En el año 2005 ha sido escenario, en el ámbito de la Unión Europea, de análisis y debate sobre el desarrollo del mercado interior de la electricidad y del gas. Los informes de la Comisión Europea (quinto informe sobre el mercado interior y libro verde de la energía) [36], ponen especial énfasis en la necesidad de eliminar las trabas que lo obstaculicen y en la completa aplicación de las directivas comunitarias sobre el mercado interior del gas y de la electricidad, así como en el fomento de las inversiones en interconexiones entre los Estados Miembros.

En este contexto, la elevación de los precios de las materias primas energéticas, petróleo, gas y carbón, en los mercados internacionales ha tenido su reflejo en los precios al por mayor de la electricidad, que en 2005 se han situado en niveles más elevados que en 2004, factor que ha servido de toque de atención sobre el grado de competencia imperante en dichos mercados, dado el elevado nivel de concentración en la producción de electricidad en la mayoría de los Estados Miembros. Estas cuestiones se han suscitado también en nuestro país con la finalización de los trabajos del “Libro Blanco” [37], las disposiciones sobre competitividad y adecuación al Mercado Ibérico incorporadas en el Real Decreto-Ley 5/2005, el propio Real Decreto-Ley 3/2006 y las adaptaciones previstas de nuestro ordenamiento jurídico a las Directivas 2003/54/CE y 2003/55/CE de 26 de junio de 2003, sobre normas comunes para el mercado interior de la electricidad y del gas natural, respectivamente. Por lo que se refiere a la evolución del mercado, este ejercicio se ha caracterizado por la incorporación al mercado de nueva capacidad de generación, tanto en régimen ordinario (4.800 MW) como en régimen especial (9.745 MW), por el avance de las actividades de representación en el mercado de pequeños productores y por una vuelta a una situación continuada de convergencia de los precios de nuestro mercado con los de los principales mercados europeos desde el segundo semestre del año 2005. También debe destacarse que la actividad de comercialización, que venía desarrollándose de forma muy dinámica hasta abril de 2005, sufre desde

esa fecha un proceso de atonía derivado de la vuelta a la tarifa integral de muchos consumidores, dado que los precios de las tarifas integrales al no incorporar las alzas de precio registradas en el mercado, resultan más competitivas que los precios liberalizados para los consumidores.

Por lo que se refiere a las modificaciones habidas en el marco normativo debe destacarse, en primer lugar la publicación del Real Decreto-Ley 5/2005. Como contenido fundamental del mismo deben destacarse las disposiciones que asignan al Operador del Sistema parte de las responsabilidades que venía desempeñando hasta ese momento el Operador del Mercado relativas a la gestión económica del mercado de producción de energía eléctrica para su adaptación al Convenio Internacional sobre el Mercado Ibérico.

El Real Decreto-Ley 5/2005 también incorpora nuevas normas de transparencia y en especial la determinación de hechos relevantes en la formación de los precios. Una vez se produzca su completo desarrollo normativo se adaptarán al mismo las informaciones publicadas por la Compañía. La citada propuesta de Reglas de Funcionamiento del Mercado contiene un capítulo específicamente dedicado a la información pública y confidencial del mercado.

En el ámbito institucional debe señalarse el normal y fructífero funcionamiento de las sesiones del Comité de Agentes del Mercado, haciéndose pública toda la información relevante sobre el funcionamiento del mismo a través del Web público de OMEL. Durante el año 2005 se han negociado en el mercado 256.620GWh de energía, por un valor de 15.792 millones de euros, lo que ha supuesto un incremento del 6,4% en cuanto a energía y un aumento del 89,6% en cuanto al valor económico respecto del año anterior.

El precio horario final medio se ha situado en el año 2005 en 6,049 cent.€/Kwh., el más alto de los últimos 9 años, con la evolución que figura en el gráfico correspondiente, ver Figura 3.



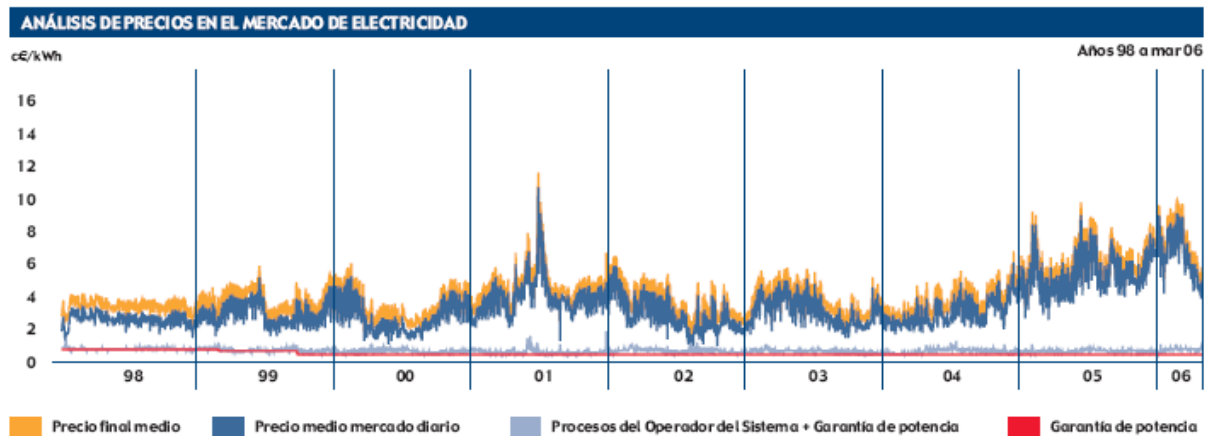


Figura 3. Curva de precios en el mercado de la electricidad

## 2.3. LA DEMANDA Y SU PARTICIPACIÓN EN ENTORNOS REGULADOS Y LIBERALIZADOS

### 2.3.1. INTRODUCCION

En tan solo un siglo, la industria eléctrica actual suministra energía a tres cuartas partes de la población mundial y sigue creciendo diariamente. La electricidad es probablemente la energía más utilizada y con más aplicaciones de las existentes. Los sistemas eléctricos en los países industrializados siempre se han contemplado como una estructura compleja donde interactuaban principalmente, las empresas eléctricas, las empresas proveedoras de equipos y tecnología y los gobiernos.

En las últimas décadas esta industria ha sufrido modificaciones importantes en su estructura tradicional, cambiando hacia sistemas eléctricos en entornos más competitivos y menos regulados.

Existen algunas razones de peso que han llevado a algunos países desde principios de la década de los ochenta a adoptar cambios en las reglas del sistema eléctrico entre las que cabe destacar:

- La mejora en la eficiencia energética del sistema.
- La mejora en la investigación y las operaciones de la industria eléctrica.
- La reducción de costes de operación de las empresas distribuidoras y comercializadoras.
- El aumento de la satisfacción de los consumidores.

Desde los primeros años de la década de los ochenta, pero sobre todo en los años noventa del siglo pasado, algunos países comenzaron la liberalización de sus mercados eléctricos como Argentina, Chile, Reino Unido, Noruega, Finlandia, Suecia, Australia, Nueva Zelanda, España y los Estados Unidos de forma parcial según los diferentes Estados. En España, siguiendo el modelo americano, también en los años 90 se potencia la posibilidad de cogeneración de energía y por lo tanto la presencia de más actores en el sistema.

La idea de que la demanda participe de forma activa en un sistema eléctrico, ya sea este liberalizado o regulado, no es nueva y siempre se ha tenido en mente por todas las partes con intereses en el sistema eléctrico.

### **2.3.2. LA DECADA DE LOS SETENTA**

Hasta esta década la capacidad de respuesta de la demanda eléctrica desde el punto de vista de su elasticidad no se había tenido en cuenta en los escenarios de planificación energética. Sin embargo, la crisis del petróleo de 1973 y sus repercusiones demostraron que la demanda era sensible a los precios elevados.

En el caso de los Estados Unidos esto provocó una importante disminución en el crecimiento de la demanda en EE.UU., hasta el punto, que fue difícil amortizar el exceso de plantas de generación que se habían construido como consecuencia de no considerar la demanda y su flexibilidad en la planificación [38]. Este acontecimiento marcó un antes y un después en los mercados energéticos. A

partir de entonces quedó patente que el suministro energético no es ilimitado y que la demanda no es inelástica ante situaciones límite.

### 2.3.3. LA DÉCADA DE LOS OCHENTA

A finales de la década de los setenta y a lo largo de los ochenta se produjo un notable interés por parte de las compañías eléctricas en aprovechar la elasticidad de la demanda a través de modificaciones de carga [39], [40], [41]. En un principio se trataba de reducir la necesidad de capacidad punta mediante la aplicación de incentivos regulados. En este periodo la gestión de la demanda se entendía como una herramienta de optimización del sistema, aprovechable por la parte suministradora a la hora de elaborar su planificación. Dentro de este contexto se desarrollaron una serie de medidas que se pueden resumir en dos enfoques diferentes:

- Las compañías, a través de los precios ofrecidos, incentivan o penalizan a los clientes para que cambien sus hábitos de consumo. Dependiendo del horario de consumo los clientes pagarían un precio. Se trata de intentar que desplacen sus consumos desde los periodos de precios punta a los periodos de precios valle. Esta solución, basada en el traslado del consumo de un periodo horario a otro, tiene limitaciones y se necesitan medidas adicionales para conseguir que la demanda reaccione a los precios no sólo de manera permanente sino también a precios elevados producidos por ejemplo en situaciones de emergencia.
- Las compañías desarrollaron programas de control directo de cargas en los que las compañías eléctricas controlan de manera remota las cargas del consumidor y las reducen según sus necesidades. Los consumidores que se acogen a este tipo de programas renuncian de alguna manera al control de sus propias cargas en el sentido de que pueden sufrir cortes de suministro un número máximo de veces fijado en el contrato. A cambio reciben recompensas por disponibilidad, traducidas habitualmente en tarifas más baratas de lo habitual para todos los periodos horarios. El problema de este enfoque son las interrupciones, ya que pueden ocurrir en cualquier momento, con los problemas que

conlleven para los clientes acogidos (paradas de procesos productivos, reducción del ritmo de trabajo, etc.). De ahí, el alto abandono de estos programas por parte de muchos clientes. Este tipo de programas se han denominado tradicionalmente “*Programas de gestión de la Demanda*” o DSM (Demand-Side-Management) [42], [43], e incluyen tarifas de interrumpibilidad, control cíclico de aires acondicionados, control de bombas de piscinas, etc..

#### **2.3.4. LA DECADA DE LOS 90**

En la década de los 90 se produce en la mayoría de los países industrializados un cambio en la estructura y funcionamiento de los mercados eléctricos. A semejanza de otros mercados – telecomunicaciones-, se liberalizan los sectores de generación, distribución y comercialización de la energía eléctrica. Este proceso ha generado no poca incertidumbre entre el sector, que unida en los últimos años a una mayor consideración de los efectos medioambientales de las actividades relacionadas con la generación de la energía, y el aumento de la importancia de la generación distribuida –y en particular de la renovable, políticamente incentivada- hace más arriesgadas las perspectivas de negocio y estabilidad en el sector. Asimismo, en la última década se ha producido también un incremento notable de la demanda de energía que no ha sido acompañado por una inversión acorde en la generación y distribución de la energía –precisamente motivada por la incertidumbre asociada a la nueva estructura de mercado, que hace económicamente mucho más arriesgada cualquier inversión a largo plazo que en años previos a la liberalización-. Por estos dos motivos: coste y riesgo de las inversiones a realizar para la adición de nuevos recursos al sistema eléctrico; y por otra parte la menor flexibilidad de la generación –en especial la renovable- surge la necesidad de incentivar la participación de la demanda en el mercado eléctrico para hacerla más flexible que antaño y compensar los factores mencionados con anterioridad. A priori esta participación de la demanda supondría una mejora apreciable en el funcionamiento efectivo y

eficiente de los mercados eléctricos al propiciar la igualdad de oportunidades reales de participación a la generación (Supply-Side) y el consumo (Demand-Side). Estas políticas de participación de la demanda suponen incentivar la gestión de ésta, bien a través de los precios –elasticidad de la demanda de cada usuario-, bien a través de políticas de control de la demanda realizadas por el propio usuario –tanto más factible cuanto mayor sea la demanda del usuario- o bien en último término por los comercializadores y distribuidores de energía –en el caso de pequeños usuarios. Por tanto, el primer paso para que los operadores del sistema, un distribuidor o un comercializador puedan considerar la puesta en marcha de políticas de flexibilización de la demanda es poder distinguir si en una zona geográfica del sistema de distribución de energía eléctrica existen usuarios que pueden “generar” paquetes de demanda suficientes en una hora y día determinados. Para ello es necesario segmentar y clasificar la totalidad de los usuarios de una zona geográfica, analizar las posibilidades de controlar y flexibilizar esa demanda, y por último reagrupar los usuarios y demandas más adecuados. Es necesario, por tanto, encontrar herramientas eficaces para la segmentación y clasificación de consumidores eléctricos en función de su perfil característico de consumo.

Dentro de este contexto, en los nuevos escenarios de libre mercado que se dan hoy en día en la mayoría de países industrializados los consumidores eléctricos tienen, por tanto, la oportunidad de elegir entre varios suministradores, así como la posibilidad de acceder a los nuevos productos y servicios que ofrezcan estas compañías comercializadoras, con tal de captar clientes y ofrecerles el mejor servicio con el máximo ahorro. Por este motivo estas nuevas empresas comercializadoras están cada vez más interesadas en el desarrollo de nuevas estrategias y productos orientados a los consumidores (Demand-Side) con el objetivo de ofrecer (alcanzar) el mejor servicio a cada uno de sus diferentes clientes o con la finalidad de encontrar nuevas oportunidades de mercado, por ejemplo a través de los denominados “*Load Trade Services- LTS*”.

En este sentido, los agentes que forman el lado productor y suministrador de la energía eléctrica, tales como agregadores, comercializadores, distribuidores, etc., necesitan un modelo que describa de forma precisa el tipo de demanda de los consumidores y clientes para evaluar el potencial de nuevos productos y políticas más específicas. Por ejemplo, tarifas a tiempo real, programas de reducción de carga ante contingencias o políticas de ahorro desde el punto de vista del lado de la demanda.

A modo resumen, es a partir de mediados y finales de los noventa donde se entiende la respuesta de la demanda como una participación activa de la misma en los procesos de obtención de precios y mantenimiento de la calidad de suministro. No obstante, durante esta década algunos investigadores pensaron que con la introducción del sistema liberalizado, la demanda entraba en recesión como herramienta útil en la nueva estructura de los Sistemas Eléctricos. Este pensamiento se basaba en la idea de que la participación de la demanda en la operación y planificación de los sistemas eléctricos, tenía interés en un entorno regulado, pues las políticas de Gestión de la Demanda (Demand Side Management) y de Eficiencia Energética (EE) estaban fuertemente incentivadas por los gobiernos y reguladores.

### 2.3.5. LA GESTION DE LA DEMANDA ELECTRICA

La Gestión de la Demanda eléctrica es una actividad económica orientada a que los consumidores tengan en cuenta los costes del suministro de la energía cuando organicen y realicen su demanda eléctrica. La industria eléctrica dispone de herramientas no obstante, que le permiten reducir este coste, entre ellas destacan:

- Aumentar su capacidad de predecir la demanda futura
- Adaptarse a la incertidumbre de su demanda
- Incentivar a los consumidores para que gestionen su consumo.

Esta última iniciativa pretende lograr una mayor implicación de los consumidores en lugar de ser las compañías suministradoras las únicas en buscar soluciones a la mejora de la eficiencia. En este sentido es posible definir dos tipos de gestión de la demanda:

- **Gestión de Demanda Directa.**- en la cual el consumidor permite al operador del sistema programar y/u operar su uso de la energía. Se suele realizar con grandes y medianas industrias y con consumidores residenciales.
- **Gestión de Demanda Indirecta.**- en la cual los consumidores modifican sus equipamientos y hábitos de consumo, favoreciendo la eficiencia del sistema. En este

caso es el usuario el que tiene que analizar y adaptar su consumo aprovechando los periodos u horas del día donde las compañías o el mercado ofrece precios muy bajos, trasladando en este sentido sus picos de potencia de las horas con precios altos a estos periodos valle.

### 2.3.6. PRINCIPIOS DE LA RESPUESTA DE LA DEMANDA

Es evidente que la participación de la demanda puede mejorar la eficiencia de los diferentes procesos de los mercados. Esta participación hay que entenderla como una respuesta de la demanda frente a la volatilidad de los precios y condiciones que en muchas ocasiones impone el lado suministrador [44]. La Figura 4 representa los beneficios conseguidos por el mercado gracias a la flexibilidad de la demanda en un periodo de precios elevado. Si la demanda (D) es inelástica el precio del mercado se fijaría en  $P_1$ , mientras que para la misma situación si la demanda es elástica ( $D'$ ) el precio se fijaría en  $P_2$ .

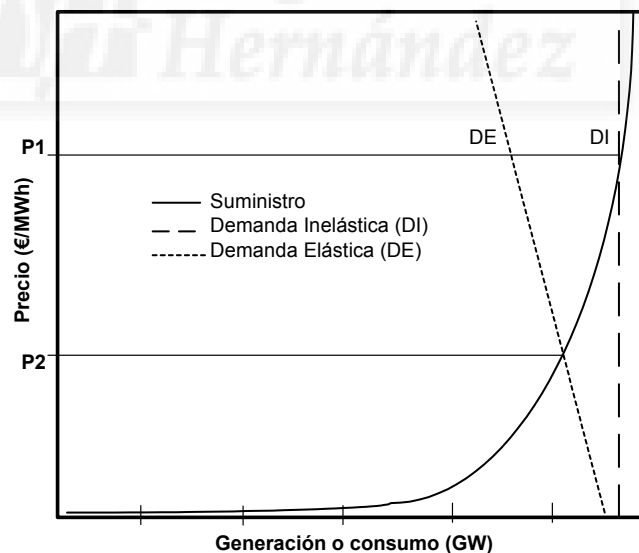


Figura 4. Consumo vs. Precio de la energía

A medida que los mercados han ido evolucionando también lo ha hecho la forma en que ha participado la demanda eléctrica. La energía eléctrica siempre se ha considerado un recurso

obligatorio e ilimitado, donde el crecimiento de la demanda y la regulación excesiva del estado han ido provocando, con los años, la ineficiencia del sistema.

### **2.3.7. INTEGRACION DE LA RESPUESTA DE LA DEMANDA EN LOS MERCADOS, EN LOS ULTIMOS AÑOS.**

En la bibliografía es posible encontrar numerosos estudios sobre la integración de la respuesta de la demanda en los mercados [45], [46], [47].

Para una correcta integración económica de la respuesta de la demanda en los mercados eléctricos, esta ha de financiarse a través de ella, pero teniendo en cuenta una serie de consideraciones.

Es evidente que aquellas estructuras de mercado que permitan la respuesta de la demanda conseguirán beneficios a largo plazo tanto para los consumidores individuales como para la sociedad en general. Por otro lado, si el beneficio neto de la respuesta de la demanda, desde el punto de vista global del sistema, se consigue no sólo de la reducción en los costes de suministro (al reducir la generación) sino que hay que incluir también el aumento de los costes del consumidor debidos a la reducción en el consumo. De este modo si el coste de incurrido por los consumidores al reducir su demanda es mayor a la reducción de costes de suministro producida por esa misma reducción, el sistema perdería eficiencia.

Otro punto a considerar es que en el largo plazo la posible reducción de precios punta se verá compensada por los incrementos en los precios base. También es cierto que no se puede ofrecer a las reducciones de carga que puede realizar la demanda el mismo trato que a la generación. Los



generadores poseen la capacidad de generar por lo que pueden venderla, sin embargo, los consumidores no pueden ofrecer vender reducciones de un consumo a no ser que lo haya comprado con anterioridad.

Por último y como punto importante destacar que el desarrollo y aplicación de mecanismos para la participación activa de la demanda no tienen por qué ser diseñados por los operadores del mercado, del sistema o los reguladores sino que han de ser el resto de los agentes los que deberían desarrollarlos como parte de su estrategia comercial. En este sentido comercializadores y agregadores serían los principales agentes interesados en el desarrollo de programas y políticas de participación de la demanda. No obstante el papel de las operadores y reguladores sería el de asegurar que la estructura del mercado admite la participación de la demanda.

#### 2.3.7.1. AGENTES INTERESADOS

A continuación se nombran los agentes del mercado que estarían potencialmente interesados en una participación activa de la demanda:

La mayoría de agentes del mercado obtendrían beneficios, en mayor o menor medida:

- Reguladores
- Operadores del mercado
- Operadores del sistema
- Operadores de distribución y operadores locales
- Comercializadores, agregadores y otros suministradores
- Consumidores

#### 2.3.7.2. COMERCIALIZADORES Y AGREGADORES

Los comercializadores y agregadores quizás sean los agentes, a los que probablemente les aportaría mayores beneficios la participación activa de la demanda.

Podrían ampliar el abanico de productos para ofrecer a sus clientes, en su mayoría pequeños y medianos consumidores, tales como nuevas políticas de precios con tarifas específicas dirigidas a determinados segmentos de la demanda, programas de participación de los consumidores con la finalidad de llegar a determinados niveles de agregación de potencia y obtener mejores precios en el mercado mayorista. Gestionar mejor los niveles de reducción con la finalidad de proporcionar al mercado o al operador los recursos obtenidos con la respuesta de la demanda, etc.. Estos serían algunos ejemplos de los objetivos y beneficios que lograrían alcanzar estos agentes.

### 2.3.7.3. LOS CONSUMIDORES

Los grandes beneficiados serían los consumidores, los cuales podrían obtener beneficios económicos al lograr reducir los precios del mercado debido a la propia participación masiva de la demanda, mejorar la fiabilidad y eficacia del mercado y del sistema eléctrico debido a la participación masiva de la demanda, y por último reducir el coste por consumo eléctrico, tanto por la reacción a los precios como a las posibles compensaciones por reducciones de carga.

## 2.3.8. ENTORNOS LIBERALIZADOS

### 2.3.8.1. EL MERCADO DE NUEVA INGLATERRA (EE.UU.)

A modo de ejemplo, en este apartado se va a exponer los diferentes programas que se han desarrollado en el mercado eléctrico de Nueva Inglaterra y las características y requisitos necesarios para su correcto funcionamiento y aprovechamiento.

Los programas de respuesta de carga desarrollados en este mercado han permitido que la parte consumidora influya de manera notoria en la demanda de recursos eléctricos, ya que las compañías de electricidad a través de normas y programas específicos ofrecen al consumidor una serie de propuestas para la reducción de la carga consumida. Estos programas van a compensar al consumidor económicamente, dependiendo del tipo de programa al que se haya acogido.

Los diferentes programas a los que se puede acoger el consumidor para la reducción de la carga son varios, entre los que se encuentran los siguientes:

- Programa de respuesta de carga (Day-Ahead).
- Programa de respuesta en tiempo real de la demanda (Real-Time Demand Response Program)
- Programa de respuesta a precios en tiempo real (Real-Time Price Response Program).
- Programa de respuesta de perfil en tiempo real (Real-Time Profiled Response Program).

#### **Programa de Respuesta de carga (Day-Ahead).**

En este programa hay que destacar que si las reducciones de carga por parte del consumidor coinciden con los picos de demanda de energía, el precio al cual se les pagara los intervalos de reducción de carga por parte del consumidor serán aquellos que corresponderán al precio de cabeza de día de la zona. En este tipo de programas se tiene que proporcionar una reducción de un mínimo de 100 kW.

#### **Programa de Respuesta de la Demanda en Tiempo Real (Real-Time Demand Response Program).**

Dentro de él encontramos dos subprogramas, los cuales se diferencian en el tiempo en el que se avisa a los consumidores que deben reducir el consumo de energía (30 minutos y 2 horas), cuando ocurran una serie de eventos que necesiten una reducción de la energía consumida. En este tipo de programas se tiene que proporcionar una reducción del consumo entre 100 Kw. y 5 Mw.

#### **Programa de Respuesta a Precios en Tiempo Real (Real-Time Price Response Program).**

En este programa el cliente (consumidor eléctrico) va a reducir su consumo de energía eléctrica durante ciertos periodos de tiempo. Estas reducciones se permitirán en unos horarios pronosticados por una entidades, de forma que garantizan que el precio por megavatio-hora que haya dejado de consumir sea de 100 dólares. En este tipo de programas se tiene que proporcionar una reducción del consumo entre 100Kw. y 5Mw.

### **Programa de modificación del perfil de consumo en Tiempo Real (Real-Time Profiled Response Program).**

En este caso los clientes acogidos a este programa son avisados de cuando deben reducir su consumo eléctrico de forma obligatoria. Estos avisos ocurrirán cuando se produzca una deficiencia en el procedimiento de abastecimiento. En este tipo de programas el cliente tiene que proporcionar una reducción del consumo de un mínimo de 200 kW.

Tras resumir las diferentes alternativas que se les pueden ofrecer a los consumidores, es posible afirmar que los consumidores disponen de opciones para intervenir de forma activa. Estos planes les ofrecen una serie de beneficios económicos con la finalidad de que reduzcan su consumo en ciertas franjas o periodos de tiempo. De esta forma el mercado puede funcionar sin deficiencias de abastecimiento eléctrico en momentos críticos.

### **¿Quien puede participar?**

Cualquier participante en el mercado que pueda proporcionar una reducción de la carga de cómo mínimo 100 Kw. y como máximo 5Mw. Los consumidores que se inscriben en los programas de reducción de la carga se subdividen en 4 grupos, dependiendo del tipo de programa al que se acojan:

#### **Programa de Respuesta de carga a principio del día.**

Los participantes en este tipo de programas pueden sufrir reducciones que concurren en el mercado de energía a principio del día. Las reducciones de la carga sucederán de lunes a viernes (no en vacaciones) y se realizarán entre las 7:00 AM y las 6:00 PM.

#### **Programa de Respuesta en tiempo real de la demanda.**

Dentro de este tipo de programas existen dos, los cuales se diferencian en el tiempo de aviso para la reducción de la carga (30 minutos y 2 horas). Estos avisos se realizaran mediante los sistemas de comunicación a través de Internet. Las reducciones de la carga sucederán de lunes a viernes y no en vacaciones y se realizaran entre las 7:00AM y las 6:00 PM. Dichas interrupciones de carga se les abonarán a los consumidores, de forma que sea el máximo entre dos opciones, teniendo en cuenta la existencia de dos subprogramas, dentro de este.

**Programa de respuesta de precios en tiempo real.**

Aquellos consumidores que se acojan a este tipo de programas deberán utilizar los precios acordados para decidir la reducción de carga voluntaria. Serán avisados cuando el precio pronosticado para la franja horaria sea igual o mayor que 100\$/MWh de lunes a viernes (no en vacaciones), entre las 7:00 AM y las 6:00 PM. Los participantes inscritos recibirán el máximo del precio en tiempo real de la zona por interrupción del consumo o un pago mínimo garantizado de 100\$/MWh cuando el periodo elegido esta abierto.

**Programa de Respuesta del perfil en Tiempo Real.**

Aquellos consumidores que se acojan a este programa deben ser capaces de interrumpir su carga durante 2 horas, de lunes a viernes (no en vacaciones), entre las 7:00 AM y las 6:00 PM. La interrupción de la carga esta bajo el control del usuario a excepción de que ocurra un evento. El participante inscrito recibirá el máximo del precio en tiempo real o recibirá un pago de 100\$/MWh.

**Agregación de la Carga**

En todos los programas se permite la agregación de carga con el objetivo de alcanzar o exceder un nivel mínimo para el programa específico. Entre los requisitos principales para alcanzar este nivel cabe destacar:

- Que los recursos de demanda agregada tienen que ser en la misma zona de carga.
- Que ha de existir una agrupación de consumidores por zona de carga, por programas y por avisos de requerimientos.
- Que una carga estipulada y medida no puede estar suscrita a más de un programa.

**2.3.9. CARACTERÍSTICAS DE LAS POSIBILIDADES Y REQUISITOS ACTUALES DE PARTICIPACION**

Para que un consumidor pueda participar en un programa de reducción de la carga tiene que cumplir una serie de requisitos:

- Ser un participante del mercado o firmar un acuerdo con otro participante que actuara de intermediario entre la ISO (Independent System Operator) y el consumidor del LRP (Load Response Program).
- Los programas de respuesta del perfil en tiempo real requieren una representación estadística de las mediciones o tecnología equivalente que esté aceptada por la ISO para determinar la respuesta.
- Para cada intervalo medido o aprobado por el plan el usuario ha de ser capaz de proporcionar un mínimo de 100 kW de reducción o estar de acuerdo con los intervalos de medida para obtener un mínimo de 100 kW de reducción por zona durante los momentos de respuesta frente a eventos de carga. Para los programas de respuesta de carga a principio de día el tamaño mínimo es de 100 kW y los consumidores deben estar de acuerdo para alcanzar o exceder el mínimo de 100 kW. Para los programas de respuesta del perfil en tiempo real el tamaño mínimo es de 200 kW.
- Los consumidores que poseen respuesta de la demanda en tiempo real tiene que ser capaces de conectarse a los sistemas de monitorización en tiempo real ICBS (Internet-Based Communication System).
- La respuesta de la demanda de los consumidores en tiempo real tiene que ser capaz de reducir la carga en 30 minutos o 2 horas después de la notificación.
- Los consumidores inscritos en los programas de respuesta del perfil en tiempo real han de ser capaces de recibir las instrucciones a través de los ICBS.
- Responsabilidades de los agentes participantes (Empresas comercializadoras, agregadoras, etc..)

Siguiendo con el ejemplo del mercado de Nueva Inglaterra las empresas comercializadoras (agregadores) actúan como intermediarios entre el ISO y los consumidores acogidos a los programas ofrecidos por estos agentes. Las responsabilidades que han de cumplir estas empresas son:

- Informar a los consumidores (clientes) del gasto o coste que conlleva ser una agente participante en el mercado.

- Proporcionar los recursos adecuados y necesarios para garantizar la instalación adecuada del hardware y software.
- Consultar a los consumidores la forma de mejorar la respuesta durante los momentos de reducción de carga.
- Asistir a los consumidores en el desarrollo de estrategias de reducción de cargas, y en la determinación de la cantidad de reducción posible frente a los diferentes niveles de precios aplicables.
- Resolver cuestiones referentes a avisos, niveles básicos, informes y registros.
- Informar detalladamente a los clientes de las disponibilidades y oportunidades de los diferentes programas.

## 2.4. SEGMENTACION Y AGREGACION

Tras analizar estos requerimientos se llega a la conclusión de que la utilización de una herramienta capaz de segmentar consumidores eléctricos a partir de su perfil de carga permitiría obtener patrones y modelos de consumo. Esta información sería de gran utilidad en los programas de respuesta del perfil (Real-Time Profiled Response Program) donde se requiere una representación de las mediciones que sea aceptada por el ISO (Independent System Operator) para determinar la respuesta. Mediante la segmentación de consumidores estamos obteniendo el perfil de consumo característico de un conjunto de consumidores, su patrón y como se comportan, además de conocer su nivel de potencia y forma típica. Todo esto permitirá conocer a que consumidores es posible dirigir programas de participación de la demanda y hasta que punto un segmento de consumidores podría proporcionar un nivel de agregación de potencia necesario para permitir a las empresas comercializadoras una participación efectiva en los mercados.

Actualmente en el mercado eléctrico español no existe la posibilidad de participación del pequeño y mediano consumidor de una forma directa o indirecta, como ocurre en otros países. Sin embargo, tomando como base la experiencia de otros mercados internacionales en la puesta en marcha de programas de respuesta de la demanda (Load Response Programs), se puede plantear la

importancia de definir y establecer una herramienta capaz de extraer características y patrones de consumo similares, ver Figura 5.

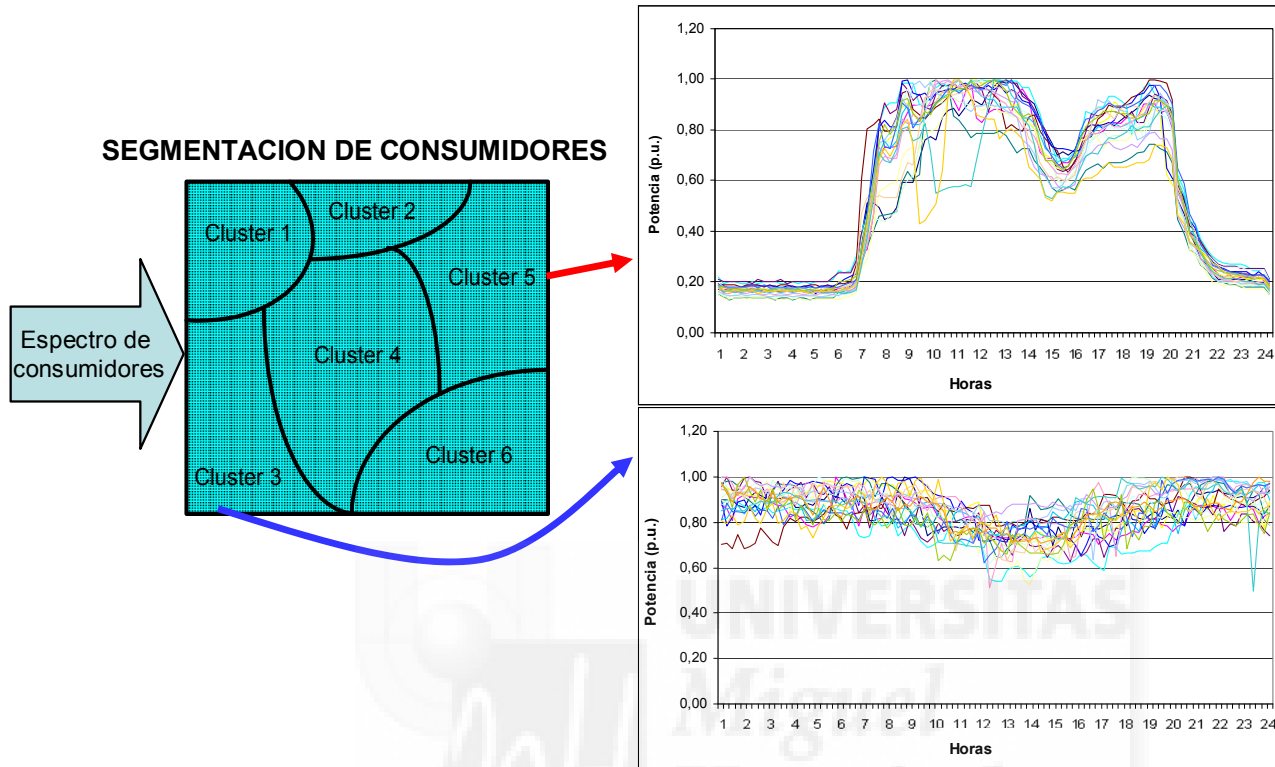


Figura 5. Segmentación de curvas de carga de diferentes consumidores.

El objetivo de clasificar consumidores es agrupar clientes con comportamientos de consumo similares, de esta forma sería más fácil determinar que usuarios tienen un perfil de consumo (curva de carga diaria) lo suficientemente coincidente con la evolución de los precios en el mercado. Por ejemplo, y a nivel internacional, con los precios dados en el “Day-Ahead Market”. En el mercado Español esta coincidencia habría que buscarla en el mercado diario. Aquellos consumidores con picos de consumo coincidentes cuando los precios pagados por la energía son muy altos, serían los primeros en plantearse la posibilidad de participar en programas de respuesta de la carga o de modificación del perfil de consumo, ver Figura 6.



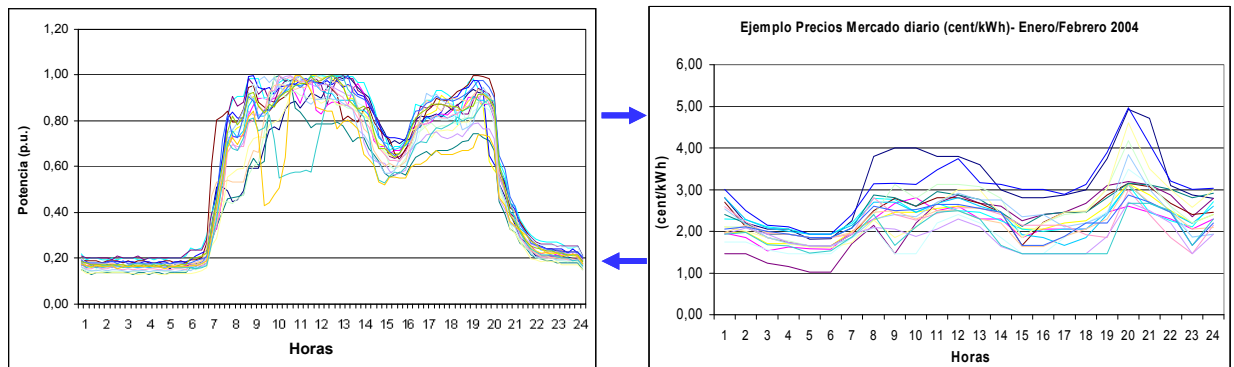


Figura 6. Comparación curvas de carga de un segmento determinado con curvas de precios del mercado diario (Day-Ahead).

Entre las políticas aplicables a los usuarios hay que diferenciar, por un lado las encaminadas a conseguir una modificación de consumo en términos de eficiencia energética a medio/largo plazo mediante la implantación de sistemas G.D., y por otro lado las políticas basadas en la respuesta de la demanda a los precios del mercado, en este caso en el corto plazo. Esto supone un nuevo escenario donde los consumidores y los suministradores pueden competir en igualdad de condiciones en los mercados energéticos competitivos.

A modo de ejemplo en EE.UU. la Comisión de la Energía del Estado de California ha financiado una serie de programas de eficiencia energética destinados a lograr una reducción del consumo entorno a 6000 GWh en el año 2008 [48]. La contribución eficaz de estos programas y la necesidad de ofrecer a los consumidores nuevos productos pasa nuevamente por el conocimiento detallado de los segmentos de demanda, entendiendo como tal, grupos de consumidores que forman clusters de características similares, desde el punto de vista del uso de la energía.

Dentro de este nuevo marco de mercados liberalizados ha surgido la necesidad, tanto en consumidores como en los propios suministradores de disponer de los sistemas adecuados para realizar correctamente las actividades de control, monitorización y medición de la energía. La necesidad de almacenamiento de datos por parte de las grandes compañías suministradoras siempre ha sido muy elevada, esto supone un problema a la hora de tratar esta información. Sin embargo, las compañías distribuidoras y comercializadoras tienen la oportunidad de extraer importantes conclusiones de estos datos, sobre todo, de los referentes a los consumidores.

El problema de este tipo de bases de datos multidimensionales, en la mayoría de casos, es la dificultad en la obtención de la información requerida de forma fácil y rápida, por lo que es necesario desarrollar herramientas para el tratamiento y extracción de características a partir de estos datos. Este objetivo queda cubierto por las técnicas de “*Data Mining*” [9], [8], algunas de las cuales se pueden utilizar para la segmentación de consumidores eléctricos y la obtención de patrones comunes.

### **2.4.1. CONSUMIDORES CON POTENCIAL DE PARTICIPACION EN EL MERCADO**

Las posibilidades de participación en los mercados eléctricos de pequeños y medianos consumidores plantea a los comercializadores y operadores la necesidad de desarrollar y ofrecer nuevos servicios y productos a los consumidores.

La complejidad de las operaciones de los mercados eléctricos dificulta la entrada y la participación directa de estos consumidores, es decir, limita la capacidad de estos para decidir cual es la mejor opción en cada momento. Por este motivo, es de vital importancia el abanico de servicios que puedan ofrecer otros agentes del mercado, actuando como intermediarios, en pro de facilitar a los consumidores la toma de decisiones y su eficiente participación.

El desarrollo de estos productos se basa en la utilización de herramientas que permitan a los agentes seleccionar e identificar potenciales consumidores, los cuales estarían interesados en participar en los mercados eléctricos. Esta selección distinguiría a los consumidores con capacidad de modificar su demanda (reducciones, cambio del perfil de consumo, etc...) o implantar sistemas de eficiencia energética, generación distribuida etc., para ofrecer una respuesta en determinadas situaciones del mercado y sistema eléctrico, ver Figura 7.



Figura 7. Cuadro de posibilidades, inconvenientes y soluciones de la participación de pequeños consumidores en los mercados competitivos

Para la caracterización de la demanda se muestra en la Figura 8 un esquema de los pasos a seguir con la finalidad de facilitar la participación de la demanda. La contribución eficaz de estos programas y la oportunidad de ofrecer a los consumidores lo que realmente necesitan pasa por el conocimiento preciso del comportamiento de estos, agrupando consumidores en segmentos o clusters con características y patrones de consumo similares. En el esquema de la Figura 8 muestra las diferentes etapas para desarrollar un sistema eficaz de participación de la demanda.

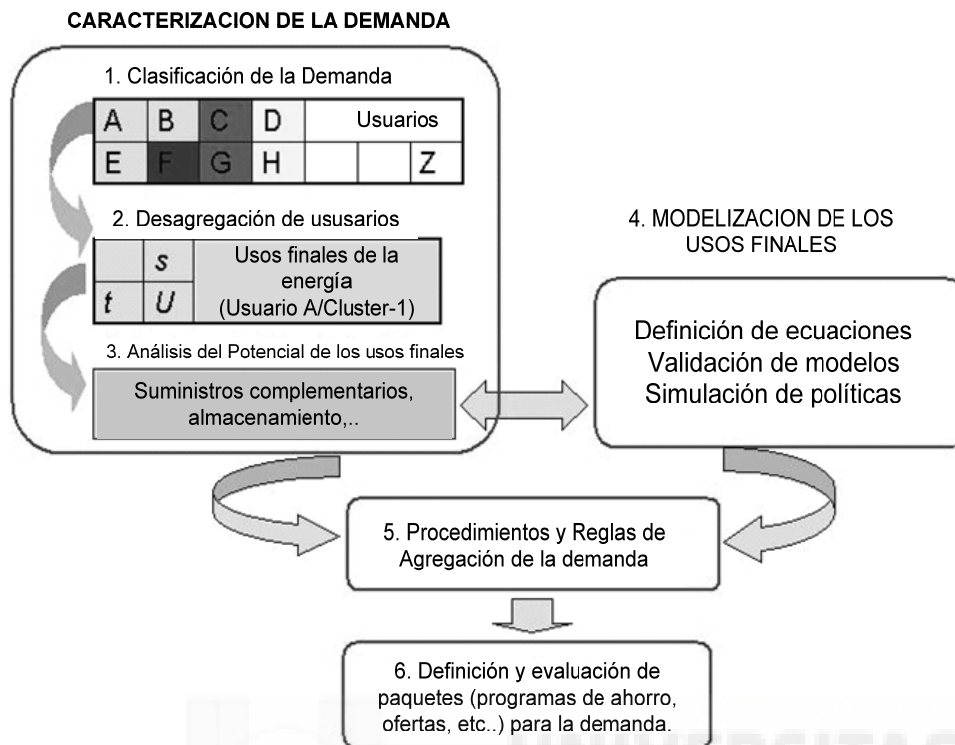


Figura 8. Estructura de un Sistema de Participación de la Demanda

La investigación que ha dado como fruto esta Tesis se ha centrado, principalmente, en la primera etapa del esquema de la Figura 8, “*La Clasificación de la Demanda*”. Clasificación entendida como segmentación de consumidores eléctricos en grupos de características de consumo similares.

El análisis del potencial de los usos finales de la Etapa 3 ha sido analizado parcialmente a lo largo de la investigación, desde el punto de vista de las distintas posibilidades u opciones a la que puede recurrir la demanda, pero no se ha profundizado en los mecanismos de su implantación.

Finalmente las etapas 5 y 6 han recibido aportaciones en lo referido a la agregación de la demanda y a la definición de una metodología de participación de la demanda, tal como se verá en los últimos capítulos. A modo resumen esta Tesis tiene aportaciones en las etapas que aparecen en color en la Figura 9.

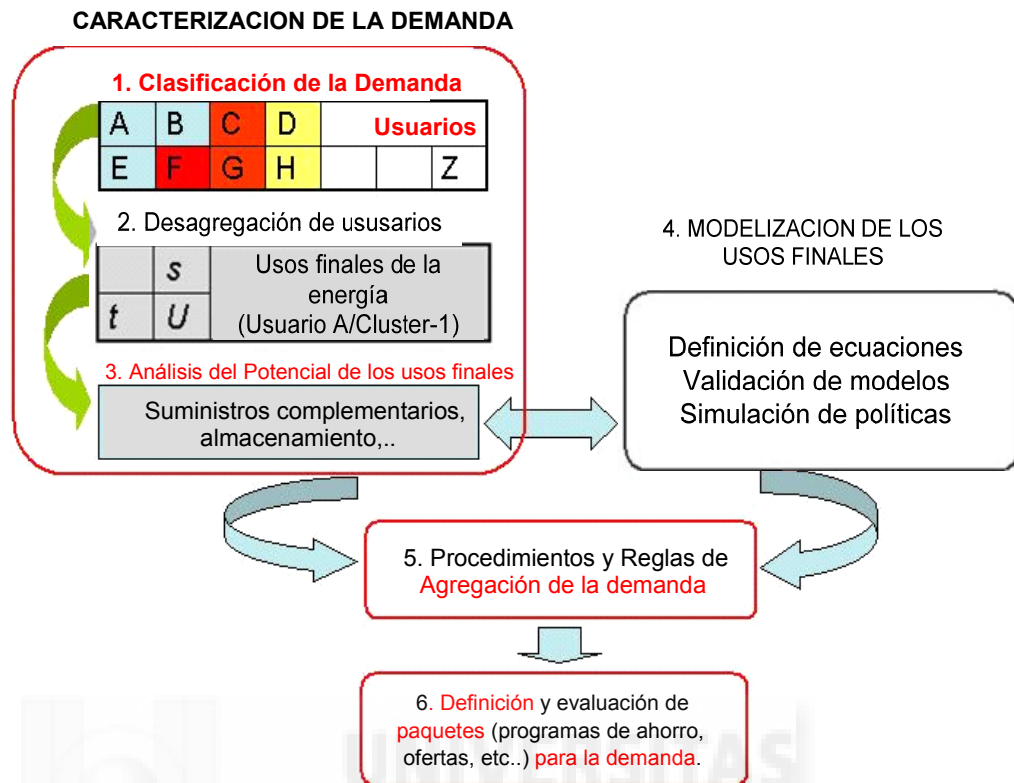


Figura 9. Aportaciones de la Tesis a la Estructura del Sistema de Participación de la Demanda.

A lo largo de la investigación se utilizarán espectros de datos formados por curvas de carga de consumidores Españoles de varios sectores y actividades, todos ellos situados en el arco mediterráneo. De igual modo y como se detallará en el capítulo 7, para el desarrollo de una metodología de participación de la demanda en el mercado, en función de los precios diarios y de balances, se han utilizado precios reales obtenidos de las bases de datos que tiene disponibles el operador del mercado eléctrico español en su página web [24].

## 2.5. CONCLUSIONES

Tras analizar la situación de otros países es previsible una participación de la demanda. La puesta en marcha en España de programas similares necesitaría del desarrollo de nuevas técnicas y conocimientos por parte de las compañías comercializadoras, lo cual facilitaría la participación de los pequeños consumidores. Los beneficios de estas políticas serían muy importantes para ambas partes; los agentes intermediarios, gracias a la agregación de carga, lograría niveles de potencia suficientes para participar en los mercados y obtener beneficios en las operaciones de compra-venta. Del mismo modo, estos agentes al identificar segmentos de demanda con patrones de consumo muy definidos podrían ofrecer a los consumidores de estos segmentos tarifas de consumo que les permitiera obtener ahorros importantes. Los pequeños consumidores serían asesorados por los agentes para la puesta en marcha de sistemas de Generación Distribuida, participación en programas, etc..

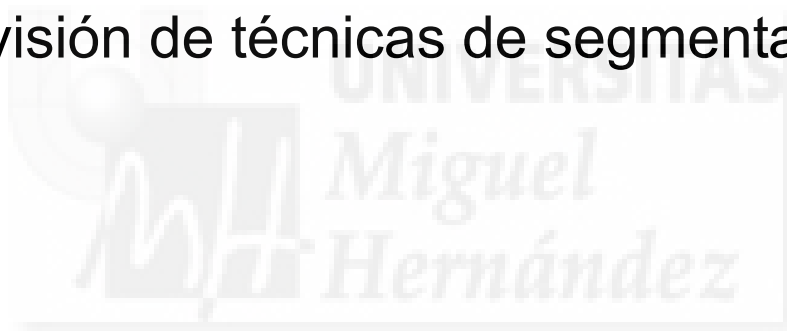




# CAPITULO 3

---

Revisión de técnicas de segmentación







## 3.1. REVISIÓN DE TÉCNICAS

### 3.1.1. INTRODUCCION

En el capítulo anterior se ha visto como los procesos de liberalización de los mercados eléctricos y energéticos han impulsado a las compañías energéticas – principalmente las compañías eléctricas – a la búsqueda de un mejor conocimiento de los comportamientos y preferencias de sus clientes. Esta búsqueda de conocimiento, que se enmarca dentro del estudio de los consumos, supone una solución alternativa al empleo generalizado de los actuales aparatos de medición de potencia, que permita la adquisición de datos en tiempos muchos menores que los actuales registradores de consumo de energía. Así pues, el empleo de técnicas de estudio y obtención de perfiles de carga y su clasificación se muestra como un camino intermedio para adecuar la facturación energética, [8] y [49], del usuario dependiendo del tipo de perfil que presente. Al mismo tiempo las técnicas empleadas se utilizan también con otros propósitos diferentes en los sistemas de potencia – tales como la predicción de consumos, cuestiones de seguridad–, propósitos que no serán expuestos en el presente capítulo.

### 3.1.2. CARACTERÍSTICAS GENERALES DE LOS PROCESOS DE CLASIFICACIÓN

Una primera división de las principales técnicas empleadas en la clasificación de consumidores eléctricos [50] sería la mostrada en la Figura 10.

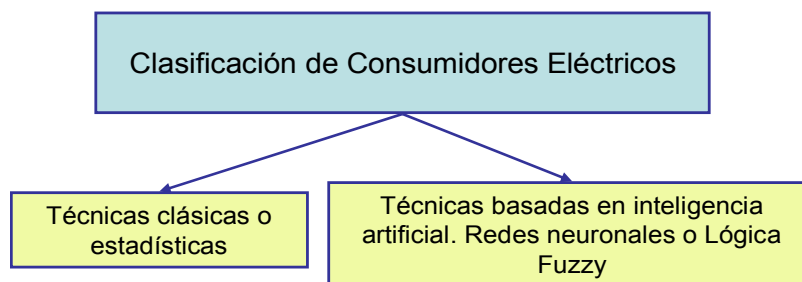


Figura 10. Clasificación General de Consumidores eléctricos

Dentro de esta primera clasificación nos encontramos con:

- Técnicas clásicas: abarcan las técnicas de agrupación (clustering), tanto estadísticas como deterministas.
- Técnicas basadas en inteligencia artificial: se incluyen aquí el empleo de redes neuronales y de la lógica difusa los denominados sistemas fuzzy.

También es posible hacer una clasificación si atendemos al tipo de datos empleados [50], en este caso podemos hablar de:

- Técnicas en el dominio del tiempo: cuando, directamente, son utilizados datos que varían en función del periodo de tiempo considerado, por ejemplo la utilización de curvas de carga diaria.
- Técnicas en el dominio de la frecuencia: cuando los datos son, previamente, tratados mediante análisis de armónicos y de pequeñas componentes de onda.

Centrando la atención en las técnicas clásicas y en las de inteligencia artificial es posible realizar una clasificación más detallada y profunda de las distintas metodologías, ver Figura 11.

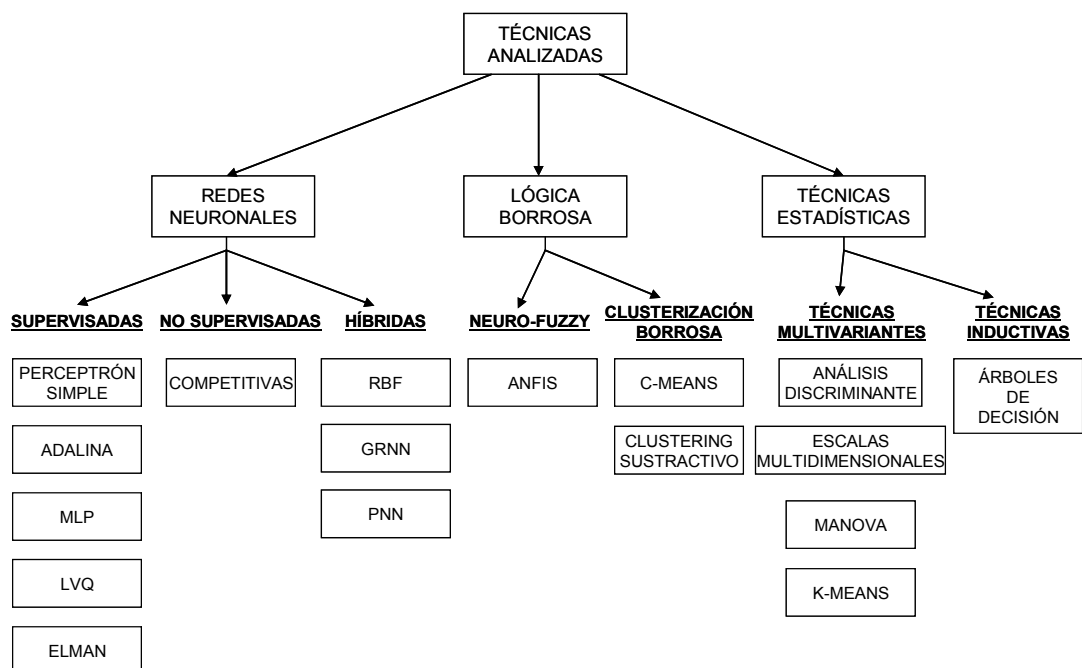


Figura 11. Esquema detallado de técnicas

Un primer paso a la hora de realizar una clasificación de consumidores eléctricos podría ser la distinción de tres macro-categorías: los pertenecientes a sectores residenciales, comerciales e industriales. Otra división básica es aquella que distingue entre los denominados clientes subjetivos – división entre residenciales y comerciales, agrupando modelos de consumo y curvas de carga no homogéneos dentro de la misma clase (y objetivos) aquellos que presentan curvas de carga y comportamientos similares, justamente aquellos que se pretenden agrupar en una misma clase.

Consultando las investigaciones realizadas anteriormente por otros autores como las llevadas a cabo por el Dr. Gianfranco Chicco [50], los pasos básicos a seguir para realizar una clasificación de clientes serían los siguientes:

- Adquisición y gestión de datos.
- Selección de los rasgos adecuados para la caracterización de los clientes y aplicación de las técnicas de clasificación.
- Búsqueda y empleo de índices y métricas diversas que estudien la idoneidad de la técnica de clasificación empleada, así como la precisión en la identificación y asignación de nuevos consumidores en uno de los clusters o segmentos encontrados durante el proceso de aprendizaje.
- Obtención de perfiles típicos para cada clase de cliente – cada grupo obtenido en el proceso de clasificación – y construcción de modelos de curva de carga para la predicción – con estimación de confianza – de la forma de curva probable de un cliente potencial para una compañía, bajo ciertas condiciones de carga.

En realidad las técnicas de clasificación son empleadas tanto en la fase de gestión de los datos como en la propia fase de clasificación. El proceso común de trabajo con bases de datos pasa por un primer uso de los métodos de clasificación consistente en identificar los perfiles que se corresponden con determinadas condiciones de carga – como perfiles con presencia de datos anómalos, o perfiles correspondientes al mismo día tipo, mismas condiciones climatológicas, etc.. En segundo lugar, se aplican esas técnicas o métodos de agrupación a los perfiles que presentan las mismas condiciones de carga o combinaciones de ellas, siguiendo

un criterio definido por los parámetros que caractericen los rasgos buscados. Siguiendo este proceso se pueden diferenciar 3 pasos, ver Figura 12:

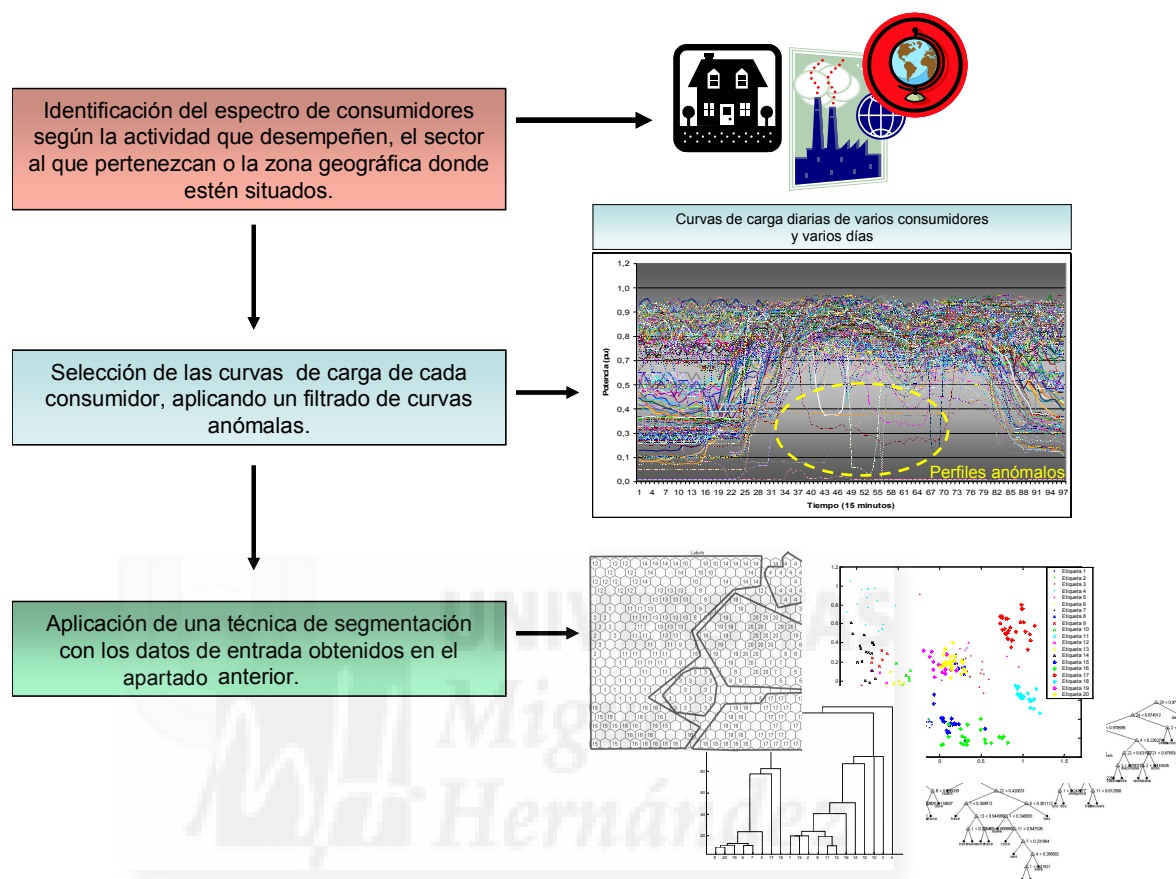


Figura 12. Secuencia de los pasos a seguir para realizar la clasificación de los consumidores.

El proceso no es cerrado ni sigue una serie de pasos estrictos, sino que existen variantes dependientes de las técnicas empleadas y la finalidad buscada en el estudio, obtención de patrones para cálculo de tarifas, obtención de modelos para predicción, etc. Existe además la total libertad de combinar diferentes técnicas y diferentes rasgos de clasificación en el proceso, con lo cual existe una amplia variedad de formas de obtener la segmentación de los datos de entrada.

### 3.1.3. ADQUISICIÓN Y GESTIÓN DE DATOS

El primer paso consistiría en la identificación tanto de los datos que puedan servir para la clasificación, como de la importancia relativa de las diversas variables, tales como temperatura, meteorología, características del consumidor, cuyo efecto sobre el comportamiento del cliente se pretenda determinar. Por tanto, es necesario identificar las fuentes adecuadas y establecer la disponibilidad de datos, así como las medidas complementarias que sea necesario realizar, campañas de medida, etc.... En caso de existir esa necesidad de tomar medidas o de recopilar información a través de fuentes diversas como pudiesen ser las encuestas, es necesario aplicar técnicas de muestreo [51, 52] con el fin de seleccionar una población de estudio que sea representativa de la población total objetivo.

Como se verá en los capítulos posteriores los datos utilizados en esta investigación han sido obtenidos de diversas fuentes pero la mayoría de ellos son de la zona geográfica de levante –sureste de España-, principalmente de las provincias de Alicante, Murcia y Valencia.

Otro de los factores importantes es determinar el tiempo y la frecuencia de muestreo necesarios. En la investigación se han tenido en cuenta intervalos de 15 minutos o de 1 hora para definir la curva de carga diaria de un consumidor. En el caso de tratar con una base de datos, como la que pudiera poseer una compañía eléctrica, la tarea consistiría en identificar conjuntos de perfiles representativos y adecuados de los clientes para ser utilizados posteriormente. Esta tarea se enmarca dentro del denominado “Data Mining” [53], que traduciremos como explotación de datos. Con este término se describe las muchas formas en las cuales los estadistas y los ingenieros de datos intentan automatizar el proceso con el cual el conocimiento inteligible puede ser obtenido a partir de grandes bases de datos, a menudo con atención particular a la presencia de datos anómalos, incompletos, muy heterogéneos, recogidos aleatoriamente y que constituyen un gran volumen a tratar. Muchas de las técnicas empleadas en estos procesos, tal como recoge bibliografía, se distinguen por su naturaleza adaptativa, modificando su estrategia de acuerdo a la naturaleza local de la carga. Es común el uso de una estrategia “divide y conquistarás” o partición jerárquica recursiva. Su principio es dividir de forma recurrente la población en subgrupos (particiones) exclusivos, de modo que los datos de una misma partición son similares en función de algún criterio y los

pertenecientes a particiones diferentes son relativamente distintos. El proceso se repite para cada partición, pero se permite que el tipo y/o detalles del algoritmo empleados para agrupar los datos varíen de acuerdo a criterios locales (esto es, basándose en los datos de la partición que está siendo procesada). La decisión de detener el proceso se toma también en base a criterios locales. Esto nos lleva a la diferente resolución con la que se presentan los datos en los grupos formados: cuando los datos son relativamente dispersos y uniformes tienen lugar un menor número de divisiones, mientras que cuando son más densos o más homogéneos tiene lugar una mayor subdivisión del dominio. Como consecuencia, cada una de las particiones finales en un modelo como éste muestra un grado similar de heterogeneidad dentro de la partición (cada grupo o cluster presenta un porcentaje similar de la variabilidad total de los datos en él incluidos). Una vez se dispone de los datos necesarios, el proceso de gestión [1] de los mismos será el siguiente, ver Figura 13.

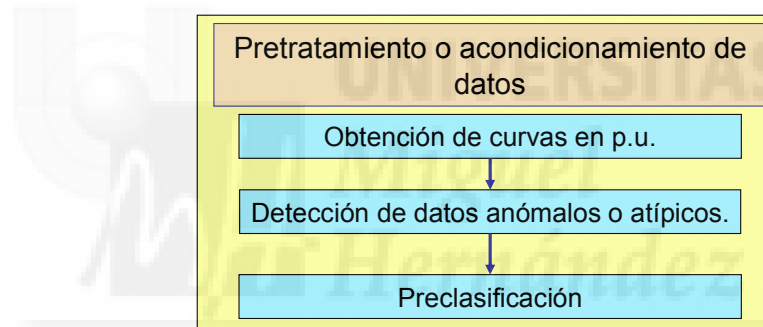


Figura 13. Proceso de gestión de los datos a utilizar

#### 3.1.3.1. PRETRATAMIENTO O ACONDICIONAMIENTO DE DATOS.

Antes de utilizar los datos como espectro de entrada a la técnica de clasificación o agrupación que se vaya a utilizar, estos han de recibir un tratamiento previo. Este tratamiento se divide en tres fases:

- Conversión de los datos a valores p.u.
- Filtrado de comportamientos anómalos comparados con el resto de las muestras.
- Preclasificación de los datos que formarán el espectro final de entrada.

### 3.1.3.2. OBTENCIÓN DE CURVAS EN P.U.

Tras obtener por medición los perfiles o curvas de consumo diario, en caso de que no estar disponibles de forma directa en la base de datos, es conveniente convertir estas mediciones en valores normalizados, curvas en valores por unidad (p.u.).

Entre los valores encontrados en la bibliografía [12], [50], para la obtención de curvas normalizadas encontramos:

- Potencia media; ya sea diaria, semanal, mensual o anual, utilizadas aunque el período de estudio sea el día.
- Potencia pico; valor no recomendado al estar afectado por las posibles condiciones anómalas que pueden hacerla variar de forma importante. Es decir, pueden existir entre los datos picos anómalos de consumo que no corresponden con las máximas habituales alcanzadas por el consumidor.
- Potencia pico del diagrama representativo; al estar promediada suaviza el efecto de posibles factores anómalos o inusuales.
- Media de las potencias pico de todos los perfiles de la clase de cliente o de todos los perfiles de un cliente.
- Potencia media total de todos los clientes.
- Potencia media total del diagrama representativo de la clase.
- Diferencia de las potencias máxima y mínima en un período determinado, por ejemplo, el año.

Dependiendo de la magnitud empleada, se puede trabajar o no desde un inicio con curvas normalizadas. De este modo, el diagrama de carga representativo puede obtenerse en p.u. desde el principio (a partir de curvas en p.u.) o trabajar con valores reales, normalizando la curva a posteriori. La Figura 14 representa un conjunto de curvas de carga diaria de un solo consumidor. En este caso, los datos obtenidos y utilizados también en pruebas a lo largo de la investigación, han sido medidos cada 15 minutos.



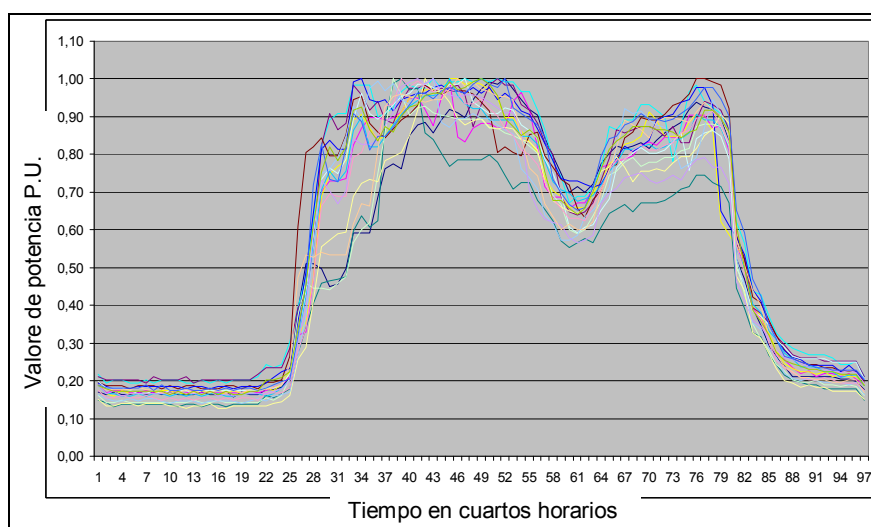


Figura 14. Ejemplo de Curvas de Carga normalizadas a valor 1

### 3.1.3.3. DETECCIÓN DE DATOS ANÓMALOS O ATÍPICOS.

Es necesaria la detección de datos erróneos o anómalos, entendiéndose por éstos los debidos tanto a condiciones anormales que pudieran darse en la red o el propio suministro, como a días que por sus características provoquen patrones de comportamiento en el consumo no habituales – por ejemplo, períodos vacacionales, días del fin de semana, día festivos intercalados entre períodos vacacionales o viceversa, día de calor en períodos de frío y viceversa, o días en los que por un mantenimiento de las máquinas que forman un determinado proceso productivo, la curva de carga se vea considerablemente modificada –. Se emplean para tal fin datos como la magnitud de tensión y la potencia activa y reactiva, medidos con frecuencia de muestreo superior a la utilizada en la toma de datos empleados en las curvas de carga. Por ejemplo tomas de datos cada minuto, de los que se obtiene medias horarias o  $\frac{1}{4}$  horaria.

De nuevo estamos ante una tarea típica del “*Data Mining*” [53], [8]. Las técnicas empleadas en esta disciplina tratan de obtener información salvando los problemas que presentan las grandes bases de datos, a saber: gran tamaño, presencia de ruido, abundancia de grabaciones incompletas, dispersión local con redundancia en determinada variedad de datos (mismo tipo de cliente) y falta de otros tipos, así como una gran heterogeneidad. Para la detección se emplean algoritmos de clasificación de diversos tipos. Entre ellos, los métodos

clásicos de agrupación de datos, los mapas Autoorganizados, SOM (en delante de Mapas de Kohonen o Self-Organizing Maps) [54], [55], algoritmos que imitan leyes básicas como la atracción gravitacional, el empleo de redes neuronales y lógica difusa en conjunto o por separado, así como técnicas estadísticas tradicionales. Posteriormente se discutirán los distintos métodos de clasificación existentes.

Finalmente, una vez identificados, los datos anómalos pueden eliminarse en caso de tratarse de muestras de tamaño importante, por ejemplo correspondientes a un período de interrupción, o en caso contrario podrían ser utilizados para la obtención de perfiles.

#### 3.1.3.4. PRECLASIFICACION

A la vez que se identifican los datos erróneos y anómalos, entre ellos los días festivos, fines de semana, etc..., se aprovecha para hacer una clasificación previa de los perfiles de carga. Esta clasificación trata de agrupar los datos disponibles en grupos definidos por determinadas condiciones de carga. Para tal fin se emplean las técnicas anteriormente nombradas, utilizando como parámetros rasgos de tipo externo, expuestos en el presente capítulo. Por ejemplo, se agrupan todos los perfiles pertenecientes a días laborales de invierno. Una vez formados los grupos se obtendrán una serie de perfiles de días representativos para cada cliente bajo cada una de las condiciones de carga empleadas en la definición de grupos previa. Esto se puede conseguir promediando los diagramas de carga del cliente que pertenezcan a dicho grupo.

## 3.2. TÉCNICAS DE CLASIFICACIÓN

### 3.2.1. MÉTODOS DE ANÁLISIS DE PERFILES DE CARGA

Dentro de los métodos de estudio de perfiles de carga asociados a diferentes tipos de consumidores eléctricos, podemos hablar de diferentes tipos de análisis [56]:

- Análisis intrínseco; la agrupación y clasificación de los clientes se realiza empleando parámetros obtenidos a partir del diagrama de carga, rasgos de forma, perfil de consumo en durante los días laborales, etc....
- Análisis extrínseco; en este caso se emplean rasgos externos, también denominados atributos de predicción, tales como atributos temporales y propios del cliente. Por ejemplo a que sector pertenece, en que región vive, la temperatura ambiente del lugar según diferentes épocas del año, etc..
- Análisis híbrido; es un enfoque bastante típico y consiste en aplicar una primera división de los datos mediante el uso de rasgos externos, análisis extrínseco, para aplicar posteriormente una técnica de agrupación mediante el empleo de factores de forma directos o indirectos, análisis intrínseco. De este modo, con el primer paso del análisis se consiguen formar grupos de perfiles más manejables y a los que se les puede aplicar algoritmos que, debido al volumen de datos inicial, eran impracticables. Además, considerando rasgos externos se descubren patrones de comportamiento que no pueden observarse empleando únicamente el análisis intrínseco.

Otro modo de ver la tarea de análisis extrínseco es modelarla como un proceso de análisis intrínseco a través de lo que se denomina atomización [56]. Este proceso se basa en la construcción de un conjunto mucho menor de perfiles atómicos partiendo del total de los perfiles de una gran base de datos. Cada perfil atómico será la media de todos los perfiles que presentan conjuntos idénticos de determinados rasgos externos, tipo de día, código comercial, etc. Agrupando el conjunto mucho menor de perfiles atómicos en lugar del conjunto original de partida, el tamaño del problema se reduce de forma importante. Después de agrupar usando

datos extrínsecos podemos reducir los cientos de miles de perfiles iniciales a unos cientos o miles y pueden llegar a convertirse en viables algoritmos de agrupación que antes no lo eran. Pero existe un inconveniente, que es la pérdida de información que conlleva este proceso.

En cuanto a los métodos de agrupación, una clasificación genérica [57] podría ser la siguiente:

- Jerárquicos; el número de grupos que se forma depende de las similitudes entre las unidades que se agrupan y los criterios de selección. Dentro de los más usados se encuentran los métodos de combinación o unión, siendo típico de los mismos que al principio cada unidad del conjunto de datos (por ejemplo perfiles) se representa un grupo diferente y en cada paso del proceso las unidades se añaden a los grupos formados. Los métodos de agrupación jerárquicos se caracterizan por el desarrollo de una jerarquía o estructura en forma de árbol.
- No jerárquicos, en estos el número de grupos que se pretende obtener se determina o se conoce a priori. Dentro de este método nos podemos encontrar tres tipos, el de umbral secuencial, el de umbral paralelo y la división para la optimización. En el método del umbral secuencial, se selecciona un centro de grupo y se agrupan todos los objetos dentro de un valor de umbral que se especifica previamente a partir del centro. Después, se selecciona un nuevo centro o semilla de grupo y el proceso se repite para los puntos sin agrupar. Una vez que un objeto se agrupa con una semilla, ya no se considera para su conglomerado con semillas subsecuentes. El método del umbral paralelo funciona de manera similar, excepto que varios centros de grupo se seleccionan simultáneamente y los objetos dentro del nivel del umbral se agrupan dentro del centro más próximo. El método de división para la optimización difiere de los otros dos procedimientos de umbral en que los objetos pueden reasignarse posteriormente a otros grupos, a fin de optimizar un criterio general, como la distancia promedio dentro de los grupos para un número determinado de conglomerados, ver Figura 15.

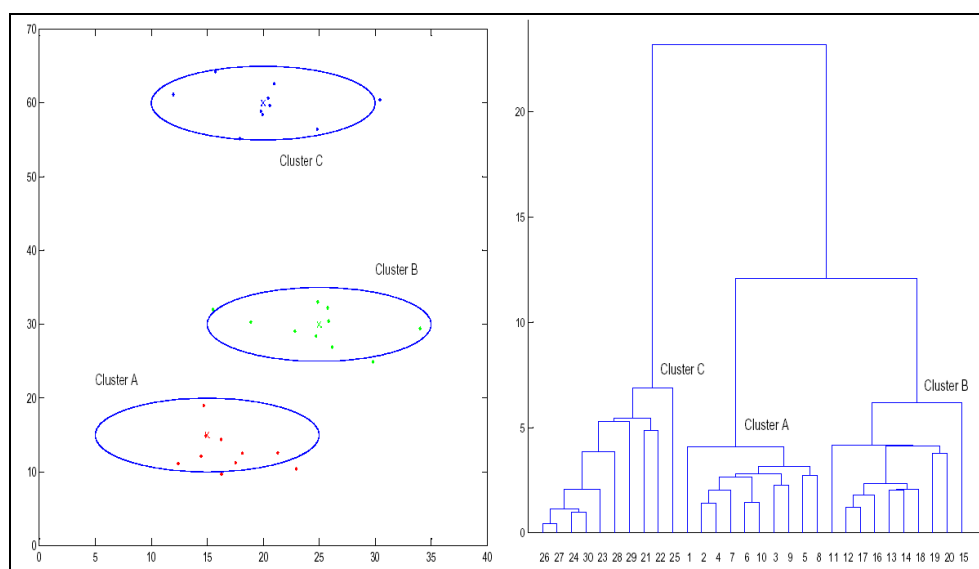


Figura 15. Ejemplo de método de agrupación no jerárquico y jerárquico

### 3.3. TÉCNICAS DE AGRUPACIÓN

Según los rasgos de forma empleados, hablamos de técnicas en el dominio del tiempo o en el dominio de la frecuencia, si bien los métodos de agrupación, aunque trabajen con unos u otros valores son idénticos, poseen la misma base. Las técnicas expuestas a continuación no son empleadas de forma única, sino que pueden utilizarse en conjunto de formas diversas dando lugar a una gran variedad de métodos.

Cuando los rasgos externos del cliente son conocidos y se usan para la clasificación o agrupación de los clientes, es evidente que los algoritmos empleados no son numéricos, sino que se basan en identificar la igualdad de los parámetros. Los árboles jerárquicos obtenidos en los métodos de análisis extrínsecos se obtienen mediante algoritmos de búsqueda binaria exhaustiva y de unión iterativa dos a dos.

En este sentido, las técnicas que se exponen a continuación están enfocadas principalmente a los métodos de agrupación intrínseca. Otro hecho bien distinto, es su utilización para agrupar perfiles de carga de clientes de los cuales se desconocen los rasgos externos. Así, se puede pretender usar técnicas de lógica difusa o técnicas estadísticas, por

considerar un ejemplo [58, 59], para dividir un conjunto de perfiles en dos grupos, uno perteneciente a los consumos de días laborales y otro relativo a los consumos de días festivos, sábados y domingos.

De forma general, podemos distinguir las siguientes técnicas [50] de agrupación:

- Redes neuronales que utilicen algoritmos de agrupación no supervisados o supervisados. Por ejemplo el caso de los Mapas Autoorganizados de Kohonen , son un tipo de red neuronal no supervisada.
- Técnicas estadísticas: Manova, árboles de decisión, escalas multidimensionales, etc..
- Lógica difusa: Anfís, Clustering Sustractivo etc...

### 3.3.1. ALGORITMOS DE AGRUPACIÓN

La diferencia principal entre los algoritmos de clasificación no supervisados y los supervisados [60] es que en los primeros el proceso no está guiado por ningún patrón o mecanismo de enseñanza, sino que los datos presentados se van ordenando en base a propiedades compartidas que el algoritmo va descubriendo, mientras que en los segundos, el proceso está controlado desde el exterior. Este caso está referido principalmente a los procesos de aprendizaje propios de redes neuronales, en cuyo caso se seleccionan perfiles de similares características, para en lugar de emplear un número elevado de datos, introducir solo los que interesan para entrenar la red.

Los algoritmos que se presentan a continuación emplean el cálculo de distancias para la agrupación de los perfiles en clusters. Estas distancias pueden estar basadas en métricas diversas. Los algoritmos básicos de agrupación [61] son los siguientes (donde sea apropiado,  $k$  denota el número de clusters o grupos buscados):

- Agrupación, clustering, de un paso. Los perfiles son procesados en orden arbitrario, asignando cada perfil  $c$  al grupo o cluster existente más cercano  $d$  si

la distancia entre ambos es menor que un límite  $t$  preseleccionado, o estableciendo un nuevo cluster (con un elemento  $c$ ) si ese límite es excedido por todos los clusters existentes (o si  $c$  es el primer perfil procesado), hasta un límite de  $k$  clusters. La complejidad o coste del cálculo no es mayor del orden  $O(kN)$ .

- Algoritmo de división binaria. Este algoritmo es empleado para encontrar agrupaciones binarias. Tenemos de partida dos clusters uno que al inicio contienen todos los perfiles y el otro que está vacío. El perfil seleccionado para ello es el perfil que más aumenta (o menos hace decrecer) la distancia entre los clusters. Esta distancia es grabada en cada paso y el proceso de transferencia es repetido hasta el que está vacío. La pareja que produce la mayor distancia entre clusters es la partición binaria ganadora. Posteriormente se continúa aplicando el método a cada una de las particiones resultante hasta generar un número determinado de clusters.
- Algoritmo iterativo de dos pasos. Inicialmente hay  $N$  clusters, uno por cada perfil. En cada paso se calcula la distancia entre cada pareja de clusters de modo que la pareja de clusters cuya distancia es la menor se funden en un nuevo cluster, se eliminan y se recalculan las distancias (donde sea necesario) entre los  $N+2-1 = N-1$  clusters resultantes y de nuevo las parejas más cercanas son unidas. Se continúa hasta que solo quedan  $k$  clusters.
- Búsqueda binaria exhaustiva. Para cada agrupación binaria posible del total de cluster se calcula la distancia. La agrupación con la mayor distancia calculada gana, separándose el conjunto total de perfiles en esos dos conjuntos. Aplicando el proceso de forma recursiva a cada cluster obtenido de la división binaria se pueden obtener un número  $k$  de clusters mayor que 2.

A medida que se va realizando la división o agrupación de los perfiles es posible obtener un árbol jerárquico [56], [57], que nos permite visualizar como quedan repartidos los datos clasificados.

### 3.3.2. TÉCNICAS ESTADÍSTICAS

Además de constituir por sí solas técnicas de clasificación de consumidores, deben ser vistas como un medio de análisis aplicable a una clasificación previa. Se pueden tratar estadísticamente todos los perfiles de cada clase hallada para determinar un perfil medio representativo de la clase, así como un perfil de desviación estándar dentro de la [4], [52]. Con estos y los perfiles de la clase se puede tratar de ajustar el comportamiento de los clientes a determinadas distribuciones de probabilidad, podemos pensar en una distribución por cada momento de toma de datos (por ejemplo, cada 15 minutos) o una para el perfil diario completo, estableciendo intervalos de confianza para el mismo.

En este sentido debemos recalcar que el uso de la distribución normal no tiene por que dar buenos resultados. De hecho, muchas cargas parecen no tener distribución regular – por ejemplo, las cargas controladas automáticamente, como puedan ser los acumuladores de calor, es más, las distribuciones influenciadas por factores económicos, psicológicos y biológicos (como es el caso del consumo eléctrico) suelen no ser simétricas. Pero como norma general, si la distribución es monótona puede transformarse [62] en una normal de parámetros  $N(0,1)$ .

En otros casos las técnicas estadísticas suponen herramientas empleadas en el propio proceso de agrupación, de tal modo que permiten estudiar valores medios y la variabilidad de los datos. De este modo, se puede estudiar la variabilidad de los perfiles contenidos en un mismo grupo, variabilidad de datos entre clusters [61], o pueden utilizarse para medir la exactitud de predicción de un modelo respecto a los perfiles de carga reales.

Por otro lado, es posible llevar a cabo análisis de correlación entre variables utilizadas en la clasificación con el fin de modelar el comportamiento de los clientes – por ejemplo, de una clase de clientes hallada anteriormente – en función de esas variables. Tomemos como ejemplo el caso más conocido, el de la influencia de la temperatura [52] sobre el consumo. Podríamos tomar los valores de potencia máxima para una clase de cliente o valores de potencia para la misma en un instante de tiempo determinado, por ejemplo a las 12:00 horas, y estudiar la influencia que la temperatura presenta sobre dicho consumo mediante un análisis de correlación. Se determina el coeficiente de correlación y si este es mayor que un determinado valor (0,3 ó 0,5) se aplica el método de regresión de los mínimos cuadrados.



También se emplean métodos de regresión múltiple que consideren la influencia de más de una variable sobre el modelo de consumo, obteniéndose un modelo de regresión lineal, cuadrático, para después comprobar la bondad del ajuste mediante estadísticos como  $R^2$  [63, 64], o mediante algún tipo de test, como el F-test o el T-test [65], [66].

## A) TECNICAS MULTIVARIANTES

### 3.3.2.1. ANALISIS DISCRIMINANTE

El análisis discriminante es una técnica estadística multivariante cuya finalidad es analizar si existen diferencias significativas entre grupos de objetos respecto a un conjunto de variables medidas sobre los mismos para, en el caso de que existan, explicar en qué sentido se dan y proporcionar procedimientos de clasificación sistemática de nuevas observaciones de origen desconocido en uno de los grupos analizados. El análisis discriminante (o análisis factorial discriminante, AFD) [67], [68], [69] es un método que permite modelizar la pertenencia a un grupo de individuos en función de los valores de varias variables, y además, predecir el grupo más probable para un individuo conociendo únicamente los valores de las variables que le caracterizan. Esta técnica se utiliza para clasificar nuevos individuos descritos por varias variables cuantitativas conociendo una muestra de individuos descritos por las mismas variables, cuyos grupos son conocidos. Los objetivos del Análisis Discriminante pueden sintetizarse en dos:

1. Analizar si existen diferencias entre los grupos en cuanto a su comportamiento con respecto a las variables consideradas, y averiguar en qué sentido se dan dichas diferencias.
2. Elaborar procedimientos de clasificación sistemática de individuos de origen desconocido en uno de los grupos analizados.

La discriminación entre los grupos se realiza mediante el cálculo de unas funciones matemáticas denominadas funciones discriminantes. Existen varios procedimientos para calcularlas, siendo el procedimiento de Fisher uno de los más utilizados. Una primera función es calculada, en la cual las medias de los grupos sean tan distintas como sea posible. Una segunda función es entonces calculada correlativamente con la primera, luego se calcula una

tercera función calculada correlativamente con las dos primeras, y así sucesivamente para tantas funciones como sea posible. El número máximo de funciones se calcula como  $g - 1$  (número de grupos dependientes menos 1) o como  $k$  (número de variables independientes). El análisis discriminante lineal, como todo modelo multivariante, supone un gran avance sobre los modelos univariantes al tener capacidad para tratar con un conjunto de variables simultáneamente. Permite conocer qué variables son relevantes para clasificar los datos así como obtener unos  $Z$ , combinaciones lineales de las variables que definen unas regiones. En el caso bidimensional se puede representar gráficamente la clasificación que realiza el análisis discriminante mediante la línea recta que mejor separa dos regiones, ver Figura 16.

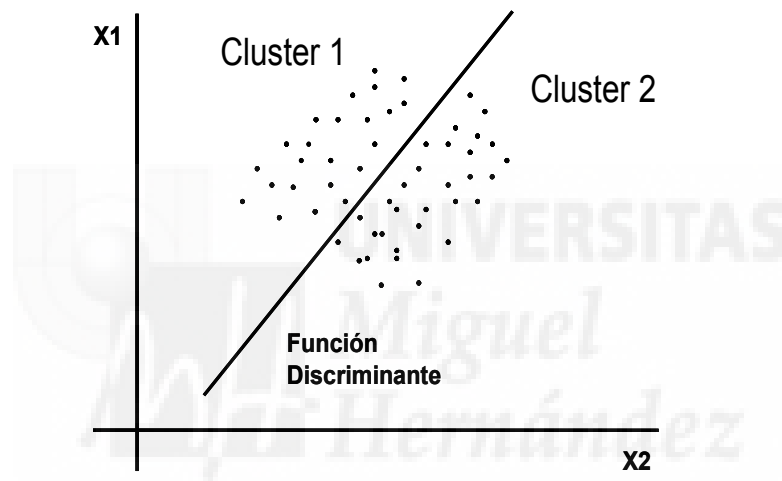


Figura 16. Ejemplo de resultado con función discriminante bidimensional

### 3.3.2.2. ESCALAS MULTIDIMENSIONALES

Medir la distancia que hay entre una serie de puntos en un mapa es sencillo: basta con utilizar una regla. Posteriormente se confecciona una tabla de distancias, como las que usan los automovilistas. Pero, ¿y el proceso contrario?, es decir, a partir de dicha tabla de distancias ¿cómo se podría representar gráficamente los puntos en un mapa?. Ese es el objetivo de las escalas multidimensionales. Una de las metas más importantes en la visualización de los datos es hacerse una idea de lo próximos que están los puntos entre sí. Cuando se dispone de un gran número de variables, es muy difícil visualizar las distancias a menos que los datos puedan ser representados en un número pequeño de dimensiones. El escalado multidimensional [70], permite visualizar lo próximos que los puntos están entre sí y puede

producir una representación de los datos en un pequeño número de dimensiones. Sus aplicaciones son variadas: el tratamiento de encuestas que utilizan análisis de juicios de similitud por parejas, emitidos por un conjunto de sujetos, las percepciones que un grupo de usuarios tienen de productos de diferentes compañías, etc.

En general, se trata de reducir un gran volumen de datos en información útil para la toma de decisiones. Usualmente aparecerán una serie de dimensiones que se plasman en mapas sencillos de interpretar. Para obtener un mapa mediante un escalado multidimensional hay que seguir los siguientes pasos:

1. Si los datos son distancias se trabaja con ellos, en otro caso se estandarizan y es necesario elegir la medida de distancia apropiada. Normalmente, se elige la distancia euclídea.
2. Calcular el número de dimensiones, comprobando que el ajuste sea bueno. Cuantas más dimensiones mejor es el ajuste, pero es más difícil de interpretar. Primero se realiza el escalado multidimensional suponiendo un número de dimensiones igual a dos, y se calcula la función de stress. Si el valor es inferior a un cierto límite, se acepta como válido que ese número de dimensiones es capaz de representar correctamente el problema. Si no, se calcula el escalado multidimensional con tres dimensiones, reiterando el proceso hasta hallar el número óptimo. En el presente trabajo, se ha utilizado el escalado multidimensional con sólo dos dimensiones, ya que además de resultar mucho más sencillo a la hora de interpretar los resultados, la función de stress obtenida para los mismos era suficientemente buena.
3. Por último se obtienen los mapas.

### 3.3.2.3. MANOVA (ANÁLISIS MULTIVARIANTE DE VARIANZA)

La técnica multivariante de análisis de varianza o MANOVA (Multivariate Analysis of Variance) [71], [72], toma un conjunto de datos agrupados y determina si el conjunto de las medias de las variables es diferente entre un grupo y el siguiente. MANOVA calcula los

valores canónicos de las variables. La visualización de las dos primeras variables canónicas muestra más separación entre los grupos que una visualización de cualquier par de las variables originales. De esta forma se puede comprobar mejor las similitudes y diferencias entre los distintos grupos.

#### 3.3.2.4. ALGORITMO K-MEANS

El denominado algoritmo k-means [73], [74], es un caso particular del método de repartición. Este método es conveniente utilizarlo cuando los datos a clasificar son muchos y/o para refinar una clasificación obtenida utilizando otro método. Supone que el número de grupos es conocido a priori. La denominada clusterización k-means, podría denominarse mejor, método de particionamiento. Este algoritmo parte los datos en k clusters, y devuelve un vector de índices indicando a cuál de los k clusters ha sido asignado cada dato. A diferencia del clustering jerárquico, k-means no crea un árbol para describir la agrupación de los datos, sino que crea un único nivel de clusters. Otra diferencia es que k-means usa cada punto de los datos, y no sólo sus proximidades. Estas diferencias hacen que k-means sea una de las técnicas más utilizadas a la hora de clusterizar grandes cantidades de datos.

K-means trata a cada punto de los datos como un objeto con una posición en el espacio. Encuentra una partición en la cual los objetos dentro de cada cluster estén tan cerca entre ellos como sea posible, y tan lejos de los objetos de otros clusters como se pueda. Cada cluster en la partición es definido por su centro. El centro de cada cluster es el punto en el cual la suma de las distancias de todos los objetos en ese cluster es minimizada. Usa un algoritmo iterativo que minimiza la suma de las distancias de cada punto al centro de su cluster. Este algoritmo mueve los objetos entre los clusters hasta que la suma no puede decrecer más. El resultado es un conjunto de clusters que son tan compactos y bien separados como sea posible.

## B) TECNICAS INDUCTIVAS

### 3.3.2.5. ÁRBOLES DE DECISION

Los árboles de decisión es una técnica de aprendizaje inductivo que representan un caso particular entre las técnicas de aprendizaje a partir de ejemplos, siendo su cometido el inducir reglas a partir de los datos históricos disponibles, para lo cual procederá a clasificar en la clase correspondiente diferentes objetos basándose en el valor de las características o atributos que los definen [75]. Un árbol de decisión consiste en una secuencia de preguntas, las cuales pueden ser contestadas como “sí” o como “no”, más un conjunto de valores de respuesta ajustados. Cada pregunta cuestiona si una predicción satisface una condición dada. Dependiendo de la respuesta a una pregunta, bien continúa a la siguiente pregunta, o llega a un valor de respuesta ajustado. Utilizan la técnica “*divide y vencerás*” para clasificar el conjunto de ejemplos en un número finito de clases predefinidas realizando particiones o subconjuntos de forma recursiva. Una vez se obtiene el árbol a partir del conjunto de entrenamiento, los ejemplos del test cuya clase se desconoce son clasificados recorriendo el árbol por aquellas ramas para las cuales los valores de sus atributos se cumplen hasta alcanzar una hoja. Los nodos del árbol quedan etiquetados con nombres de atributos, las ramas con los valores posibles para el atributo, y las hojas con las diferentes clases. Para la aplicación de estos sistemas de aprendizaje inductivo, se parte de un conjunto de ejemplos. Cada ejemplo debe tener la misma estructura consistente en una conclusión (o decisión) y un número de características o atributos que definen esa conclusión o decisión. El sistema construye un árbol de decisión que representa la relación existente entre la conclusión-decisión y sus atributos. Es decir, se produce un proceso de generalización de forma que el árbol de decisión generado clasifica correctamente los ejemplos dados. Este árbol, además, se caracteriza por ser el óptimo en el sentido que minimiza el número de atributos requeridos para alcanzar la conclusión-decisión, siendo esta la explicación de por qué ciertos atributos no aparecen en el árbol. Por ejemplo, suponiendo que se desea saber si un determinado partido de fútbol se va a jugar o se va a suspender dependiendo de la meteorología que haya en el momento del partido, se puede realizar un árbol de decisión como el que se presenta a continuación, donde se tienen en cuenta varios aspectos, ver Figura 17.

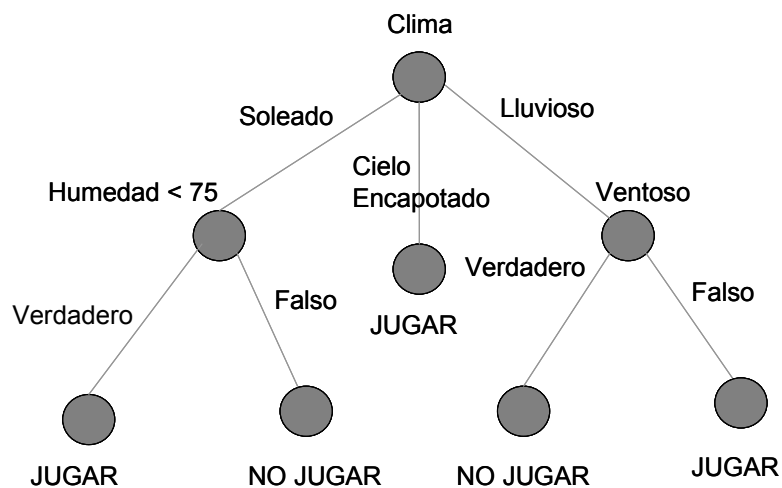


Figura 17. Ejemplo árboles de decisión

### 3.3.3. REDES NEURONALES

Las redes neuronales son sistemas formados por capas de nodos denominados neuronas, las cuales realizan algún tipo de acción u operación con los datos procedentes de la capa anterior para dar lugar a determinadas salidas. Estos sistemas tratan de reproducir el comportamiento de algún sistema real, siendo para ello entrenadas mediante conjuntos de datos conocidos. Estas redes han demostrado tener un buen comportamiento en el modelado de la carga y en la previsión de consumos. Para ello emplean diversos algoritmos de aprendizaje [76] como pueden ser el algoritmo de retro-propagación BM, (Backpropagation-Momentum-Algorithm) y algoritmo de propagación “elástico”, (RP, Resilient-Propagation-Algorithm). A diferencia de otros métodos que predicen picos que en ocasiones no se llegan a producir, las redes neuronales tienen una tendencia a predecir curvas más suaves.

Un tipo de red muy empleada en diversas aplicaciones es el mapa auto-organizado de Kohonen, Self Organizing Map (SOM) [50]. Este tipo de red se ha utilizado en el análisis de datos y monitorización de procesos, como por ejemplo la crisis bancaria española ocurrida en el periodo de 1977 a 1985, [49]. Otras aplicaciones han sido la realización de mapas fonéticos para el reconocimiento de palabras [49], o la organización de grandes cantidades de datos en internet [77]. Se trata de una red no supervisada preservada topológicamente, lo que significa

que mantiene la estructura de los datos del espacio de entrada, que lleva a cabo un análisis de agrupación de los datos iniciales. Estos datos, en este caso los perfiles de carga, se consideran puntos en un espacio n-dimensional, para el caso de perfiles de carga con valores medidos de hora en hora sería un espacio de 24 dimensiones. La red está compuesta de un mallado de unidades predefinido, usualmente bidimensional, que forma una capa competitiva. De este modo, cada unidad está representada por un vector de importancia o peso de n dimensiones y por su posición en la red. Para cada muestra de los datos de entrada, las unidades compiten y solo gana una. Entonces el algoritmo cambia el peso de esa unidad ganadora (la unidad cuyo vector de peso está más cercano), de acuerdo a algún criterio de distancia, a la muestra presentada y el de las unidades vecinas en relación inversamente proporcional de la distancia de éstas a la primera, lo que se denomina función de vecindad. Durante el entrenamiento de la red se van formando áreas, burbujas de actividad, que presentan entre sí unas distancias proporcionales a las distancias que existen entre los datos de entrada en el espacio de entrada de datos original. El rasgo más atractivo del SOM es que, una vez entrenado, el mapa representa la proyección de los datos del conjunto de entrenamiento, pertenecientes a un espacio n – dimensional, en un espacio bidimensional. La preservación topológica de la estructura de entrada y la simulación de las burbujas de actividad distingue este enfoque de los enfoques de agrupación no supervisados tradicionales. Los resultados del SOM deben tratarse posteriormente – post-procesado – para obtener de forma efectiva los clusters y por tanto las clases de consumidores. El número de clusters depende de consideraciones subjetivas y del objetivo de la clasificación ver Figura 18. Puede elegirse inicialmente – lo que permite ajustar los parámetros de los algoritmos – o ser determinado por el método automáticamente.

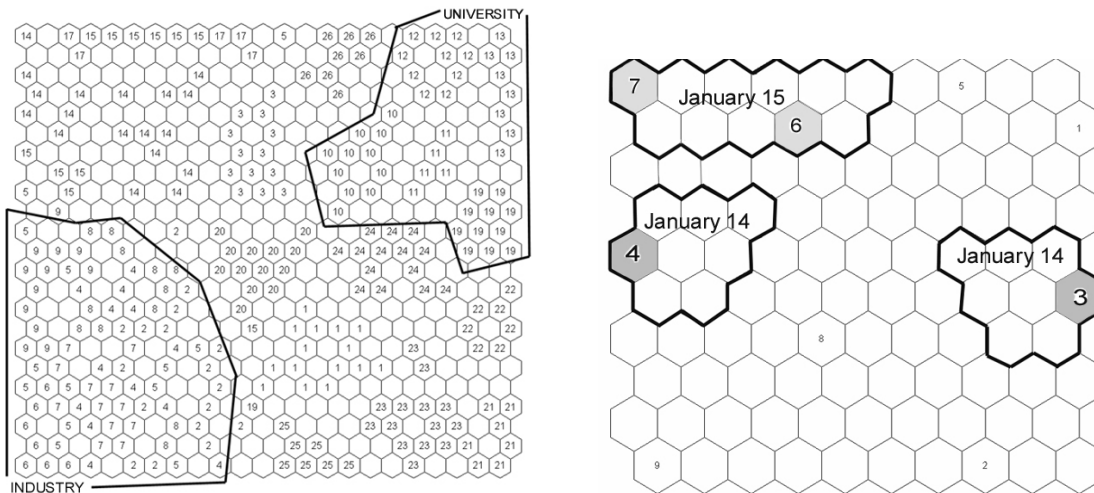


Figura 18. Ejemplo de resultados obtenidos mediante un SOM de Kohonen.

Existen por otro lado, además de las redes neuronales artificiales, otro tipo de redes, las denominadas redes de enlaces funcionales [60] – FLN, Functional Link Network –, propuestas por Pao y Sobajic, y que se emplean junto a las primeras. Se trata de redes sin presencia de capas ocultas, que permiten llevar a cabo procesos de aprendizaje supervisado y no supervisado con la misma estructura de red y de los datos, hecho que no es posible con una misma estructura de una red neuronal. En una red de estas características tendremos enlaces a los que llegan la salida de uno o varios nodos. Cuando esto ocurre cada uno de esos enlaces puede dar lugar a procesos diferentes dependiendo de la entrada, lo que evita tener capas intermedias con una neurona para cada uno de esos procesos.

En [78] por ejemplo se compara la aplicación de el método de lógica difusa c-means con la aplicación de una red neuronal artificial, ANN, Artificial Neural Network, con estructura como la mostrada en la figura

## C) REDES SUPERVISADAS

### 3.3.3.1. ADALINA

La Adalina es un tipo de red considerada como clásica, cuyo nombre proviene de Adaptive Linear Neuron (Neurona Lineal Adaptativa), [79], [80], [81]. Este modelo utiliza una neurona similar a la del perceptrón, pero de respuesta lineal cuyas entradas pueden ser



continuas. Esto permite que sus salidas tomen cualquier valor, mientras que la salida del perceptrón está limitada a 0 o 1. Las redes lineales, al igual que los perceptrones, solamente pueden resolver problemas linealmente separables.

### 3.3.3.2. PERCEPTRÓN MULTICAPA (MLP)

La red neuronal conocida como perceptron multicapa (Multi-Layer Perceptron) [82], [83], [84], nace a partir de añadir capas intermedias (ocultas) a un perceptrón simple. Esta arquitectura suele entrenarse mediante el algoritmo de retropropagación de errores o BP (Back- Propagation), [85], [86], o bien haciendo uso de alguna de sus variantes o derivados, motivo por el que en muchas ocasiones el conjunto arquitectura MLP + aprendizaje BP suele denominarse red de retropropagación, o simplemente BP.

Dentro de la arquitectura de este tipo de redes MLP existen numerosas variantes, que se consiguen por el hecho de incluir neuronas no lineales en la capa de salida, solución que se adopta especialmente en problemas de clasificación, introducir más capas ocultas, emplear diferentes funciones de activación, limitar el número de conexiones entre una neurona y las de la capa siguiente, introducir arquitecturas recurrentes, etc. Uno de los inconvenientes principales en este tipo de redes es la selección del número idóneo de neuronas ocultas para un problema concreto, ya que si se elige un número inferior al óptimo, la clasificación no será apropiada, mientras que si es superior, su aplicación puede ser inabordable en la práctica.

### 3.3.3.3. RED COMPETITIVA LVQ

Existen tres tipos de redes neuronales competitivas: la denominadas redes de capas competitivas, los Mapas Autoorganizados o SOM (Self-Organizing Maps) y las redes LVQ (Learning Vector Quantization), [87], [88]. En el aprendizaje competitivo, las neuronas de la salida compiten por activarse y sólo una de ellas permanece activa ante una determinada información de entrada a la red. Esta neurona se denomina prevalente o vencedora, y en función de ella se ajustan los pesos de las conexiones. Las neuronas de las redes competitivas aprenden a reconocer grupos de vectores de entradas similares. El método LVQ se usa para

entrenar capas competitivas de forma supervisada. Una capa competitiva aprende automáticamente a clasificar vectores de entrada, sin embargo, las clases que encuentra la capa competitiva dependen únicamente de la distancia entre los vectores de entrada. Si dos vectores de entrada son muy similares, la capa competitiva los pondrá probablemente en la misma clase al no proporcionarle etiquetas. Por otra parte, LVQ aprende a clasificar los vectores de entrada en etiquetas elegidas por el usuario. Por tanto, a una capa competitiva se le puede pedir que divida los vectores de entrada en un número determinado de clases, pero no se le puede pedir que clasifique un vector cualquiera en una clase determinada. Las redes LVQ clasifican los vectores de entrada en etiquetas de clases usando una capa competitiva para encontrar subclases en dichos vectores de entrada y combinándolos en las etiquetas de las clases.

#### 3.3.3.4. REDES DE ELMAN

Las Redes de Elman [89], pertenecen a las denominadas redes realimentadas o recurrentes. Es un tipo de red BP (Back-Propagation) de dos capas, la primera de tipo sigmoidea y la segunda lineal, con realimentación desde la salida de la primera capa hasta la entrada de la misma. Esta conexión recurrente permite a la red de Elman detectar y generar patrones variables en el tiempo [90]. El único requerimiento de este tipo de redes es que la capa oculta posea suficientes neuronas.

### **D) REDES NO SUPERVISADAS**

#### 3.3.3.5. RED DE CAPA COMPETITIVA

Este tipo de red es no supervisada y unidireccional [91]. Esto quiere decir que las neuronas de la capa competitiva se distribuyen por sí solas para reconocer los vectores de entrada presentados de forma frecuente. La función de transferencia competitiva acepta un vector de entrada a la red para la capa y devuelve en la neurona una salida 0 para todas las neuronas excepto la ganadora, que será aquella asociada con el elemento más positivo de dicha entrada. La salida para la neurona ganadora será 1. Dentro de este grupo de redes no

supervisadas competitivas nos podemos encontrar las denominadas redes autoorganizadas, en las que solamente una neurona (o grupo de vecinas) pueden quedar finalmente activadas. La base de la operación de estos modelos es la competición entre las neuronas, materializada en forma de inhibiciones laterales, a través de las cuales cada una trata de inhibir a las demás [92]. Dentro de este proceso competitivo la neurona más activada será la que consiga inhibir a las demás, por lo que será la única que permanezca activada, este tipo de redes también se denominan redes WTA (winner-take-all).

En los modelos competitivos, durante la fase de aprendizaje las neuronas vencedoras obtienen como premio el refuerzo de sus conexiones sinápticas. La competición es un comportamiento básico en muchos de los modelos neuronales autoorganizados más conocidos, como el ART [93], Neocognitrón [94], o los Mapas Autoorganizados de Kohonen [55], [54], comentados en el punto 3.3.3 de este capítulo.

## **E) REDES HIBRIDAS**

### **3.3.3.6. REDES DE BASE RADIAL**

El modelo de redes de base radial, RBF (Radial Basis Function) [95], [96], aunque de reciente introducción, cada vez cuenta con más aplicaciones prácticas gracias a su simplicidad, generalidad y rapidez de aprendizaje. Se trata de un modelo que a menudo se estudia junto al MLP por ser una red unidireccional para aproximación funcional, pero que puede considerarse de tipo híbrido por incorporar aprendizaje tanto supervisado como no supervisado. Usa neuronas de la capa oculta con funciones de respuesta Gaussiana. Las redes de base radial pueden requerir más neuronas que las redes de tipo BP (Back-Propagation). Este tipo de red trabaja mejor cuando se dispone de muchos vectores de entrenamiento. Estos modelos de Redes Neuronales utilizan como base de partida técnicas estadísticas conocidas y utilizadas en problemas de clasificación y reconocimiento de patrones o aproximación. En algún caso, el modelo es una representación de una técnica estadística conocida en forma de red neuronal, y también en algún caso, el aprendizaje es directo y no hay ajuste iterativo de los pesos involucrados, lo que aumenta considerablemente la velocidad de entrenamiento. Las redes RBF están compuestas por dos capas de neuronas: una capa oculta de base radial y una

capa de salida lineal. Si una neurona tiene salida 1, sus pesos de salida en la segunda capa pasan sus valores a las neuronas lineales de la segunda capa. Cada entrada a las neuronas es la distancia entre el vector de entrada y sus pesos.

### 3.3.3.7. RED DE REGRESION GENERALIZADA

Las redes GRNN (Generalized Regression Neural Networks) [97] se usan normalmente para aproximación de funciones. Poseen una capa de neuronas de base radial y una capa lineal especial. Cuanto más grande sea el valor del parámetro “*spread*” en estas redes, más suave será la aproximación de la función. Para seguir a los datos de cerca debe usarse un valor de *spread* más pequeño que la distancia típica entre los vectores de entrada. Para seguir a los datos de forma más flexible, se debe usar un valor más grande de este parámetro.

Las redes neuronales de regresión generalizada tienen la propiedad de no requerir entrenamiento iterativo. Aproximan cualquier función arbitraria entre las entradas y las salidas y estiman la función directamente de los datos de entrenamiento. Además, son consistentes; es decir, cuando el conjunto de entrenamiento es de gran tamaño, el error de estimación se acerca a cero.

### 3.3.3.8. RED NEURONAL PROBABILISTICA

Las redes neuronales probabilísticas o PNN (Probabilistic Neural Networks) [98], pueden ser usadas en problemas de clasificación. Cuando se presenta una entrada, la primera capa computa las distancias desde el vector de entrada a los vectores de entrada del entrenamiento y produce un vector cuyos elementos indican lo próxima que está la entrada respecto a la entrada de entrenamiento. La segunda capa suma estas contribuciones para cada clase de entradas para producir como salida de la red un vector de probabilidades. Finalmente, una función de transferencia competitiva en la salida de la segunda capa recoge el máximo de estas probabilidades y produce un 1 para esa clase y un 0 para las demás. Por tanto, la red clasifica un vector de entrada en una de las clases porque esa clase tiene la máxima probabilidad de ser correcta. Si el valor del parámetro “*spread*” es cercano a cero, la red actuará como un clasificador cercano de vecindad. Conforme “*spread*” se hace más grande, la red diseñada tendrá en consideración varios vectores cercanos y no será tan exacta. Las redes

PNN son una representación en formato de red neuronal del bien conocido método “Clasificador Bayesiano” (CB), usado con éxito para resolver muchos problemas de clasificación. Son un método prometedor para muchos problemas de clasificación difíciles, y tienen como ventaja importante respecto del CB, su fácil manejo e interpretación, para lo que el usuario no necesita tener grandes conocimientos estadísticos.

### 3.3.4. LÓGICA DIFUSA

La lógica difusa, metodología fuzzy, trata del trabajo con grupos de datos que presentan límites que no están claramente definidos. Las técnicas de clasificación mediante lógica difusa se basan en el empleo de las denominadas funciones de afinidad [58], [59], [78]. Tales funciones, en ocasiones modeladas mediante técnicas estadísticas y en otras basadas en el cálculo de distancias, tratan de medir la afinidad que un dato de la muestra, un perfil por ejemplo, presenta respecto a un grupo o cluster de datos, de modo que dicho dato pueda ser incorporado al grupo más afín según las características buscadas.

Dentro de las técnicas de lógica difusa es bastante empleado el método c-means o k-means [73], [99]. Este método, del cual existen dos variables, se basa en la minimización de una función objetivo que representa la distancia de cualquier punto de datos dado al centro de un cluster, estando esta distancia ponderada mediante una función de afinidad. Por último, citar que es común el empleo de técnicas combinadas como las redes neuronales y las la lógica difusa.

#### A) NEURO-FUZZY

##### 3.3.4.1. ANFIS (SISTEMA DE INFERENCIA NEURO-FUZZY ADAPTATIVO)

La estructura básica del tipo de sistema de inferencia borrosa [100] es un modelo que ajusta características de entrada a funciones de pertenencia de entrada, funciones de pertenencia de entrada a reglas, reglas a un conjunto de características de salida, características de salida a funciones de pertenencia de salida, y por último las funciones de pertenencia de salida a un único valor de salida. ANFIS es una herramienta para aplicar

técnicas de inferencia borrosa al modelado de datos. El contorno de las funciones de pertenencia depende de los parámetros, y cambiando estos parámetros cambiará el contorno de la función de pertenencia. En vez de elegir los parámetros de la función de pertenencia mirando solamente los datos, con ANFIS estos parámetros pueden ser elegidos automáticamente. Esto es lo que se denomina técnicas de aprendizaje neuro-adaptativas. La idea básica de las técnicas de aprendizaje neuro-adaptativas es proveer, para el proceso de modelado borroso, un método de aprendizaje de información sobre un conjunto de datos para computar los parámetros de la función de pertenencia que mejor permitan al sistema de inferencia borrosa asociado ajustar los datos proporcionados. El método de entrenamiento funciona de forma similar al de las redes neuronales artificiales. Luego utilizando un conjunto de datos, ANFIS construye un sistema de inferencia borrosa o FIS (Fuzzy Inference System) cuyos parámetros de la función de pertenencia son ajustados usando un algoritmo BP (Back-Propagation) sólo, o en combinación con un método de mínimos cuadrados. Esto permite que el sistema borroso aprenda a partir de los datos que modela.

## **B) CLUSTERIZACION BORROSA**

### **3.3.4.2. CLUSTERING SUBSTRACTIVO**

Cuando no se tiene una idea clara de cuántos clusters hay para un determinado conjunto de datos, el clustering substractivo [101], [102], [103] mediante un rápido algoritmo permite estimar el número de clusters y sus centros. Este método puede ser utilizado para inicializar métodos iterativos de clusterización basados en la optimización como el c-means clustering o para modelar métodos de identificación como ANFIS. También es posible implementarlo sólo para proveer un método rápido de una sola pasada tomando los datos con sus etiquetas y generando un sistema de inferencia que modele el comportamiento de los datos sin necesidad de entrenamiento, al contrario que sucedía con ANFIS, lo que hace que sea mucho más rápido y requiera mucha menos memoria.

Mediante el clustering substractivo se genera un FIS (*Fuzzy Inference Systems*) de tipo Sugeno, el cual puede ser usado sólo, o para generar un FIS inicial para el posterior entrenamiento con ANFIS. Se basa en extraer un conjunto de reglas que modelen el

comportamiento de los datos. Primero obtiene el número de reglas y las funciones de pertenencia de los antecedentes y después usa mínimos cuadrados lineales para determinar las ecuaciones de los consecuentes de cada una de las reglas. Para diseñar un FIS usando clustering substractivo, hay que proporcionar un valor para el parámetro denominado “RADII”. Este parámetro especifica el rango de influencia del centro del cluster para todas las dimensiones de entrada y de salida, asumiendo que los datos están dentro de un hipercubo unitario (rango [0 1]). Especificando un radio de cluster pequeño, normalmente se producen clusters más pequeños en los datos y como consecuencia más reglas.

#### 3.3.4.3. FUZZY C-MEANS CLUSTERING

Fuzzy c-means (FCM) [104], [105], [106] es una técnica de clusterización de datos, donde cada uno de los datos pertenece a un cluster (grupo) en un grado determinado que es especificado mediante un grado de pertenencia. Provee un método que muestra cómo agrupar datos que llenan algún espacio multidimensional en un número específico de clusters. FCM empieza asignando una posición inicial a los centros de los clusters, los cuales deben marcar la posición principal de cada cluster. La posición inicial de los centros de los clusters suele ser obviamente incorrecta. Adicionalmente, FCM asigna a cada dato un grado de pertenencia para cada cluster. Después, actualizando de forma iterativa los centros de los clusters y los grados de pertenencia de cada dato, FCM mueve iterativamente los centros de los clusters hacia la posición correcta entre el conjunto de datos. Esta iteración se basa en minimizar una función objetivo que representa la distancia entre cada dato y el centro del cluster debida al grado de pertenencia de ese dato. La salida del método FCM es una lista de los centros de los clusters y los grados de pertenencia para cada dato.

#### 3.3.5. LA BONDAD DE LA AGRUPACIÓN REALIZADA

Con el fin de determinar si la clasificación obtenida mediante alguno de los métodos anteriores es adecuada, o para comparar distintas clasificaciones realizadas, se emplean los denominados índices de idoneidad [107], [50]. Para definir los mismos, es necesario hacer uso de una serie de distancias como son:

- Distancia entre dos diagramas de carga
- Distancia entre el diagrama de carga representativo de una clase y el diagrama representativo de cada uno de los clientes que pertenecen a dicha clase.
- Distancia media interna de una clase, definida como la media geométrica de las distancias sobre todos los miembros de esa clase.
- Distancia entre un diagrama de carga determinado y un diagrama de carga plano con un nivel constante.

A partir de las mismas se pueden calcular los índices de idoneidad. A continuación se exponen los distintos tipos de índices empleados:

- Indicadores que emplean la noción de distancia Euclídea de forma directa. El principal de ellos es el MIA, Mean Index Adequacy, Índice de Idoneidad media.
- Indicadores que emplean información acerca de lo compactos que son los grupos en los que se da cabida a los diagramas de carga pertenecientes al mismo cluster así como información de la interdistancia entre los diagramas representativos de carga. Entre ellos:
  - CDI, Cluster Dispersión Indicador. Es el indicador de dispersión del cluster
    - La proporción de dispersión
    - Varias formas del índice R de Davies-Bouldin [108].
  - Matriz y factores de similitud entre los diagramas de carga representativos de cada clase obtenida.

El índice de Davies-Bouldin está basado en la función k-means. Valores pequeños del índice de DB garantizan una correcta clusterización o agrupación, ya que en este caso los clusters o agrupaciones se muestran muy compactas y los centros entre clusters están bastante alejados unos de otros.

A continuación se describe brevemente como funciona el índice de D.B. Dada una partición de clusters  $(C_1, C_2, \dots, C_k)$ , se define la similaridad relativa entre dos clusters  $C_i$  y  $C_j$  como:



$$RS_{i,j} = \frac{E_i + E_j}{d(m_i, m_j)}$$

Donde  $d(m_i, m_j)$  son las distancias medias entre los clusters  $i$  y  $j$ .  $E_i$  es la distancia promedio cuadrática desde los puntos del  $i$ -ésimo cluster considerado a la media del mismo cluster.

Con  $RS_{i,j}$  podemos conseguir la máxima similaridad relativa ( $MSR_i$ ) entre el cluster  $i$  (cluster considerado) y cualquier otro cluster.

$MRS_i$ , se calcula como:

$$MRS_i = \max_{i*j} \{RS_{i,j}\}$$

El índice de Davies-Bouldin (DB) para una partición de clusters dada ( $C_1, C_2, \dots, C_k$ ) es el valor promedio de  $MRS_i$  ( $i=1, 2, \dots, K$ ) y es denominado como  $DB(k)$ .

El  $DB(k)$  se calcula por tanto como:

$$DB(k) = \frac{1}{k} \sum_{i=1}^k MRS_i$$

La partición en un determinado número de clusters que logra minimizar el índice de DB es elegida como la mejor partición o agrupación.

Existen además indicadores específicos para determinar el número óptimo de clases en el caso del método de lógica difusa c-means [109]. Son las denominadas funciones de validez.

Por último indicar, que determinar el mejor número de clusters a usar, lo que a veces se denomina número de clusters natural, para un conjunto de datos particulares es altamente subjetivo y depende de aquello para lo que van a ser usados los grupos.

### 3.4. CONCLUSIONES

Tras revisar un abanico representativo de técnicas existentes y empleadas en la agrupación y clasificación de datos, el autor se ha encontrado con una gran variedad de herramientas con suficiente capacidad y utilidad en este campo. La investigación bibliográfica ha permitido al autor la realización de un esquema de técnicas clasificadas, ver Figura 11, según su área de conocimiento base.

Quizás las más comúnmente utilizadas desde hace tiempo han sido las técnicas estadísticas. Sin embargo, dentro del área de la Inteligencia Artificial, las redes neuronales y las técnicas de lógica Fuzzy han abierto en los últimos años nuevas líneas de aplicación en el tratamiento de datos. En el caso de de las redes neuronales, el desarrollo de los ordenadores personales ha permitido la utilización de estas técnicas en problemas de cierta envergadura.

En el capítulo 4 el autor realizará un análisis de cada una de las técnicas más importantes con el fin de seleccionar la más ventajosa y representativa para la segmentación de consumidores eléctricos.



# CAPITULO 4

---

Análisis de técnicas de segmentación





## 4.1. INTRODUCCIÓN

El presente capítulo recoge de forma sintetizada la investigación y análisis que se ha llevado a cabo para comprobar y comparar la capacidad de clasificación y segmentación de consumos eléctricos de diferentes técnicas y metodologías, la mayoría de ellas expuestas en el capítulo 3.

## 4.2. TÉCNICAS ANALIZADAS

Las técnicas analizadas han sido las mostradas en el esquema de la Figura 19 y ya comentadas en el capítulo 3.

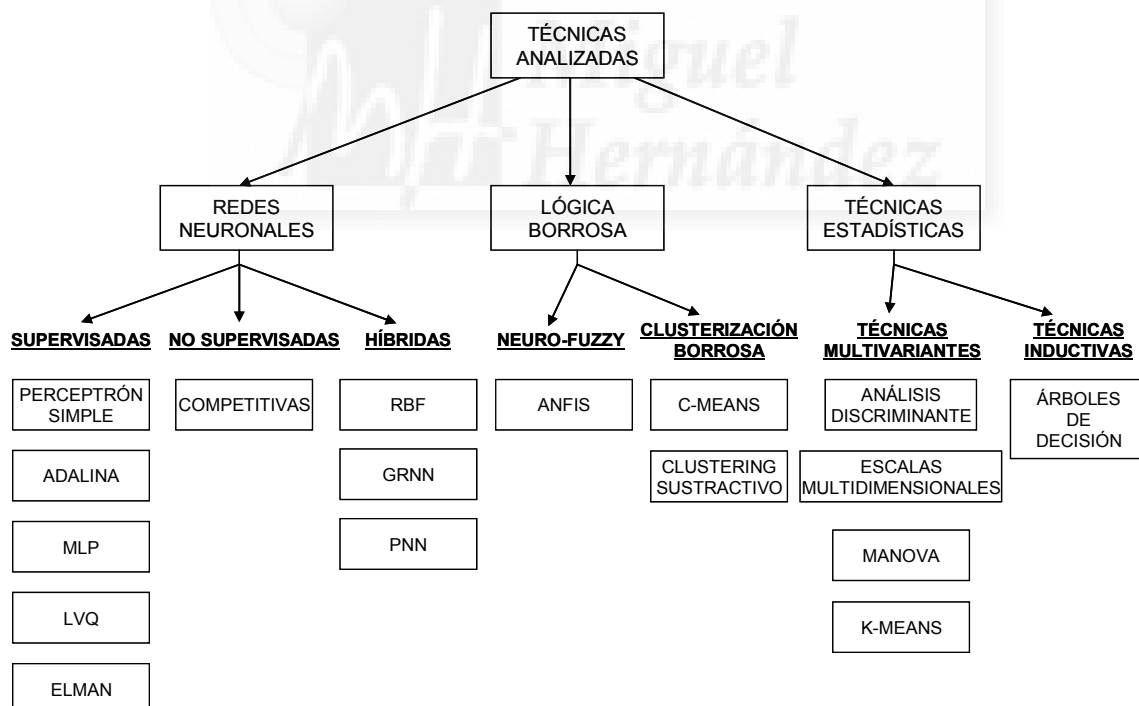


Figura 19. Clasificación de las técnicas analizadas

No obstante, este capítulo recoge un resumen de los resultados obtenidos del análisis de las técnicas, debido principalmente a la extensión de la investigación que se llevó a cabo la cual proporcionó una cantidad notable de gráficos y tablas para cada una de las técnicas analizadas.

### **4.3. HERRAMIENTAS Y SOFTWARE DISPONIBLE**

En este punto se exponen los aspectos relacionados con la puesta en práctica de sistemas neuronales artificiales, sistemas borrosos y técnicas estadísticas.

#### **4.3.1. IMPLEMENTACION DE REDES NEURONALES ARTIFICIALES**

La simulación mediante software de una Red Neuronal es su realización en forma de programa ejecutable sobre un ordenador de propósito general. Por ser el procedimiento más simple, rápido y económico, constituye siempre la primera etapa en el desarrollo de una Red Neuronal. Además, es una solución muy versátil que permite el ensayo con cualquier tipo de arquitectura. El proceso de simulación de una Red comienza con su modelado mediante programas de ordenador escritos en lenguajes de alto nivel como C o Pascal, que se ejecutarán en ordenadores convencionales, PC compatibles. Aunque de esta manera se pierde su capacidad de cálculo en paralelo, las prestaciones que ofrecen los ordenadores actuales resultan suficientes para resolver numerosos problemas, al permitir simular redes de tamaño medio a una velocidad razonable. Los principales problemas de la simulación en PCs convencionales surge del carácter esencialmente secuencial de éstos, en contrapartida al inherente paralelismo de las Redes. Si se tiene en cuenta que muchos modelos neuronales son computacionalmente muy costosos (especialmente en la fase de aprendizaje, donde para modelos como Back-Propagation aplicado sobre redes de gran tamaño se puede hablar de tiempos de cálculo de muchas horas o días), es fácil comprender la dificultad de su aplicación en tiempo real en ciertos casos.

### 4.3.2. PROGRAMAS DE SIMULACION DE REDES NEURONALES

En el desarrollo de una Red Neuronal se hace uso de programas de simulación, que se pueden clasificar en tres tipos:

- i) Comerciales, desarrollados por empresas y puestos a la venta.
- ii) De libre distribución, realizados por grupos de investigadores de Universidades o por particulares, que se pueden obtener de forma gratuita a través de Internet (salvo si se trata de programas shareware, en cuyo caso hay que pagar una pequeña cuota).
- iii) De producción propia, que cada cual puede confeccionar.

#### 4.3.2.1. PROGRAMAS COMERCIALES

Las ventajas de los programas comerciales son los entornos gráficos intuitivos y fáciles de usar que presentan, el elevado nivel de prestaciones que ofrecen, la disponibilidad de productos para aplicaciones especializadas (como reconocimiento de imágenes, predicción de series temporales, etc.), la posibilidad de cursos de formación a usuarios y apoyo por parte de la empresa. El principal inconveniente es el elevado precio y, con frecuencia, ser de productos cerrados en los que no se pueden manipular y adaptar modelos según el interés del usuario. Los siguientes son algunos de los paquetes software de uso más extendido disponibles hoy en día en el mercado:

- NeuralWorks Professional II Plus (NeuralWare, [www.neuralware.com](http://www.neuralware.com)). Es uno de los programas clásicos y más complejos, muy empleado desde hace años. Permite el trabajo con más de treinta modelos de Redes Neuronal con diferentes reglas de aprendizaje (BP, ART, SOM, GRNN, PNN, LVQ, RBF,...) y contempla muchos tipos de visualizaciones gráficas.
- MATLAB Neural Network Toolbox (The MathWorks Inc., [www.mathworks.com](http://www.mathworks.com)). MATLAB es un entorno matemático que se ha



convertido casi en un estándar en el campo de la ingeniería; en particular, son bien conocidas sus Toolboxes (cajas de herramientas) que simplifican enormemente el trabajo en campos como el procesamiento de señal e imagen, identificación y control de sistemas, etc... Las toolboxes para redes neuronales han tenido una excelente acogida, estando bastante extendido su uso tanto en la enseñanza como en la investigación. Contempla un variado número de Redes neuronales (BP, RBF, Elman, asociativas, SOM, LVQ, etc.), y al tratarse de rutinas en código fuente, el usuario puede adaptarlas a sus necesidades, su gran ventaja es la posibilidad de reprogramar y seleccionar parámetros); por el contrario, al trabajar en modo interpretado se pierde algo de velocidad.

- NeuroSolutions (NeuroDimensions Inc. [www.nd.com](http://www.nd.com)). Se trata de un producto de gran calidad, modular visual y fácil de usar. Posee una interfaz gráfica muy vistosa orientada a objetos (bloques funcionales), lo cual facilita la creación de nuevos tipos de Redes. Incluye MLP, Elman, SOFM, RBF, etc...
- Brain Maker (California Scientific Software, [www.brainmaker.com](http://www.brainmaker.com)). Es uno de los programas más famosos, aunque se basa tan sólo en el algoritmo BP, pero incluyendo variantes (como recurrencia) y contemplando técnicas diversas, como validación cruzada, análisis de sensibilidad, etc..

Existen muchos otros simuladores comerciales que no han sido expuestos aquí y que servirían perfectamente para realizar los ensayos objetos de esta investigación.

#### 4.3.2.2. PROGRAMAS DE LIBRE DISTRIBUCION

La principal ventaja de estos programas es que están disponibles de forma gratuita a través de Internet, proporcionando en la mayoría de casos un material de mucho valor, como códigos fuente, manuales, ejemplos, y en ocasiones incluso atención por correo electrónico para las dudas que puedan surgir sobre su instalación o manejo. A

continuación se comentan dos ejemplos representativos de simuladores, uno general y otro específico:

- SNNS (Stuttgart Neural Network Simulator). Se trata de un simulador de carácter general desarrollado en la Universidad de Stuttgart (en la actualidad es mantenido en la Universidad de Tübingen), que ha tenido una elevada aceptación y repercusión. Contempla muchos modelos: BP (con infinidad de variantes), RBF, ART1, ART2, ARTMAP, LVQ, Hopfield, Jordan/Elman, SOFM, etc. Es ampliable por el usuario (funciones de activación, algoritmos de aprendizaje). Inicialmente se desarrolló para estaciones UNIX, pero existe versión para PC. Viene acompañado por unos manuales de implementación y de usuario muy extensos y claros.

<http://www-ra.informatik.uni-tuebingen.de/SNNS/>

<http://www.informatik.uni-stuttgart.de:80/ipvr/bv/projekte/snns/snns.html>

- SOM\_PACK y LVQ\_PACK. Se trata de dos simuladores de carácter específicos desarrollados por el grupo del profesor Teuvo Kohonen en la Universidad Tecnológica de Helsinki. Son dos simuladores sobrios, pero muy rigurosos y completos, con una excelente documentación de consulta.

<http://www.cis.hut.fi/research/som-research/>

### 4.3.3. IMPLEMENTACION DE SISTEMAS BORROSOS

Como en el caso de las redes neuronales, un sistema borroso puede implementarse como programa ejecutable por un microprocesador convencional (o microcontrolador), o puede realizarse en hardware específico. Para el campo de los sistemas basados en lógica borrosa, se han introducido diversos sistemas de desarrollo de propósito específico. Resulta difícil realizar una lista completa de todos los disponibles, ya que de forma continua aparecen nuevos. En el punto siguiente se describe algunos de los programas realizados expresamente para lógica borrosa.

#### 4.3.3.1. ENTORNOS DE LOGICA BORROSA

A continuación se describen algunos de los entornos más difundidos orientados específicamente a la lógica borrosa.

- FuzzyTECH ([www.fuzzytech.com](http://www.fuzzytech.com)). El entorno FuzzyTECH es uno de los más difundidos y completos para el desarrollo de sistemas basados en lógica borrosa. Se pueden utilizar gran variedad de funciones y posee un editor gráfico de reglas, con un formato similar al de una hoja de cálculo. Se pueden asociar a las reglas diversos métodos de inferencia estándar, y también pueden asignarse pesos individuales a cada una de las reglas. El entorno admite la simulación “*off-line*” y en tiempo real. Como cualidad importante a destacar de este entorno es la disponibilidad de un gran número de versiones con una interfaz de usuario común, adaptadas a un gran número de implementaciones.
- FIDE ([www.aptronix.com](http://www.aptronix.com)). El entorno FIDE (Fuzzy Inference Development Environment) ha sido desarrollado por la compañía americana Apronix, en colaboración con Motorola. Es un entorno potente pero posee algunas dificultades en su uso. Por un lado, la diversidad de formatos a utilizar en los diferentes ficheros complica el aprendizaje y utilización del entorno; por otro lado, al permitir solamente el empleo de una familia de microcontroladores (Motorola) su aplicación resulta limitada.
- TILShell ([www.ortech-engr.com/fuzzy/togai.html](http://www.ortech-engr.com/fuzzy/togai.html)). El entorno TILShell, de la compañía Togai InfraLogic, es uno de los entornos para desarrollo de sistemas basados en lógica borrosa más completos. Dispone de una interfaz de usuario que permite la definición gráfica del sistema en base a bloques funcionales conectados. También dispone de una herramienta para la generación automática de las reglas, utilizando la definición de las entradas y salidas del sistema.
- MATLAB Fuzzy Logic toolbox. Esta herramienta permiten diseñar sistemas basados en lógica borrosa con potentes técnicas de entrenamiento, así como su

integración en sistemas de control complejo simulables en SIMULINK. Las funciones gráficas de interfaces usuario (GUIs) son una guía de cómo diseñar paso a paso un sistema de inferencia borrosa (FIS- Fuzzy Inference System)

#### 4.3.4. IMPLEMENTACION DE TECNICAS ESTADISTICAS

Hoy en día cualquier ordenador es capaz de ejecutar los más avanzados programas estadísticos para el tratamiento de datos. SPSS ([www.spss.com](http://www.spss.com)), BMDP ([www.statsol.ie/html/bmdp/bmdp\\_home.html](http://www.statsol.ie/html/bmdp/bmdp_home.html)), SAS ([www.sas.com](http://www.sas.com)) son algunos de los programas estadísticos de análisis de datos más avanzados.

Por otro lado existen programas de propósito más general, como Mathematica ([www.wolfram.com](http://www.wolfram.com)), Maple ([www.maplesoft.com](http://www.maplesoft.com)), MATLAB ([www.mathworks.com](http://www.mathworks.com)) pero no menos potentes. En el caso de la toolbox de Matlab esta incluye funciones y herramientas interactivas para analizar datos históricos, modelización de datos, simulación de sistemas, así como, desarrollo de algoritmos estadísticos.

#### 4.3.5. SOFTWARE SELECCIONADO

Para realizar la investigación se decidió finalmente utilizar el programa Matlab y las *Toolboxes* disponibles para analizar cada una de las técnicas de clasificación. En el paquete de *Toolboxes* de la versión de Matlab 6.5 nos encontramos la *Fuzzy Logic Toolbox*, la *Statistics Toolbox* y la *Neural Network Toolbox* a la que habría que añadir la conocida SOM Toolbox ([www.cis.hut.fi/projects/somtoolbox/](http://www.cis.hut.fi/projects/somtoolbox/)) para la simulación de redes competitivas del tipo autoorganizadas (Self-Organizing Maps) y las denominadas LVQ (Learning vector quantization).

La capacidad de cálculo y visualización del programa, unida a la disponibilidad de paquetes de herramientas específicas donde es posible encontrar funciones y algoritmos

predefinidos de cada una de las técnicas, inclinaron la balanza hacia la utilización de este entorno de programación.

#### **4.4. ESPECTRO DE DATOS DE ESTUDIO**

Se han utilizado 2 conjuntos de datos de entrada formados por un número determinado de consumidores. De cada consumidor se han tomado muchos perfiles (curvas de carga), correspondientes a distintos días de la semana a lo largo de varias semanas.

El primer conjunto de datos está formado por tan solo 2 consumidores, uno de tipo industrial y otro de tipo institucional, en este caso una Universidad. Este primer conjunto ha permitido realizar unas primeras pruebas para conocer la capacidad de separación y clasificación de algunas técnicas sencillas como es el caso de la red neuronal denominada perceptrón simple.

El segundo conjunto de datos está formado por 26 consumidores pero con datos de curvas de carga tomados cada 15 minutos. Los consumidores pertenecen a sectores de tipo industrial, institucional, comercial y residencial. En investigaciones realizadas por otros autores nos encontramos con periodos de medida similares, es decir cada 15 minutos [50, 110].

##### **4.4.1. SELECCIÓN DE LOS RASGOS ADECUADOS PARA LA CARACTERIZACIÓN DE LOS CLIENTES**

El proceso de extracción de información a partir de una base de datos necesita un tratamiento previo de dichos datos para darles un formato uniforme que permita trabajar con ellos. Las curvas de carga diarias presentan picos y diferentes valores para cada cliente, por lo que resulta necesario normalizar los datos para tener un formato común para todos los clientes y poder trabajar con ellos [14].

Estudios previos permitieron a sus autores llegar a la conclusión de que el formato más apropiado para los datos de consumo eléctrico era el de curvas que presentaran 24 o 96 valores diarios de consumo. Basándose en estos estudios, se optó por utilizar 24 valores en el primer conjunto de datos descritos anteriormente, y de 96 valores en el segundo de conjunto de datos, para permitir evaluar la influencia del número de datos/día en la selección de una técnica, ya que se trata de obtener las características más relevantes de los clientes (con el menor coste de tiempo y trabajo) considerando la disponibilidad de las fuentes o recursos (datos tomados normalmente de compañías eléctricas). Es decir, si se tiene una curva de demanda eléctrica diaria, lo que hay que hacer es dividir dicha curva en 24 o 96 puntos, cada uno correspondiente a una hora del día o  $\frac{1}{4}$  de hora, así se tendrá una curva con 24 o 96 valores, esto es, cada vector correspondiente a un consumidor será un vector de 24 o 96 dimensiones. A su vez, los datos son normalizados entre 0 y 1. Para el segundo conjunto de datos también se ha tenido en cuenta a efectos de entrenamiento y análisis de las técnicas el añadir un valor más, en total 97. Este último corresponde a las 0 horas del día siguiente.

#### **4.4.2. PRIMER CONJUNTO DE DATOS (DOS CONSUMIDORES)**

El primer conjunto de datos corresponde a dos tipos de consumidores diferentes; uno correspondiente a una industria de tamaño medio y otro a una universidad, ambos situados en España, dentro del arco mediterráneo. Estos mismos datos fueron utilizados en [13], y corresponden a curvas de demanda eléctrica tomadas en el 2003. El número total de vectores (días) utilizados en el diseño o entrenamiento es de 285, mientras que para el testeo se han utilizado dos vectores de cada tipo de consumidor. Estos cuatro vectores no fueron utilizados en el diseño o entrenamiento previo. La finalidad de testeo es comprobar la capacidad de identificar y asignar vectores nuevos (curvas de carga) a uno de los dos consumidores con los que ha sido entrenada la técnica. En la Tabla 2 se puede observar el número de consumidores utilizados de cada tipo, la actividad del cliente y sus correspondientes etiquetas.

Tabla 2. Primer conjunto de datos

Cliente	Etiqueta	Número de vectores de entrada	Actividad del cliente
I	1	123	Industria
U	2	162	Universidad

Los vectores utilizados a lo largo de la investigación corresponden a curvas de demanda eléctrica para un día determinado, expresadas en Potencia Activa (kW) respecto a Tiempo (1 día) y normalizadas entre 0 y 1. En el caso de este primer conjunto de datos se han tomado valores de la potencia activa cada hora, así se tienen 24 valores de la misma a lo largo de un día completo. Se pretende que las técnicas vistas en el capítulo 3, mediante su correcto diseño y/o entrenamiento, representen de forma correcta el mapping subyacente de los datos de entrada, es decir, que segmenten adecuadamente los datos distinguiendo entre los dos tipos de consumidores que se les presentan. Una vez aprendido adecuadamente el mapping, las técnicas se testearán mediante nuevos datos no presentados previamente en el diseño o entrenamiento para comprobar si también generalizan correctamente, es decir, para ver si no sólo han memorizado los datos de entrenamiento, sino que han reconocido su significado y son capaces de clasificar correctamente nuevos vectores de entrada.

En las siguientes figuras se muestran las gráficas de los datos utilizados, tanto para el diseño de la técnica, donde se han seleccionado 50 vectores de cada tipo de consumidor de los 285 disponibles, como para el testeo posterior.

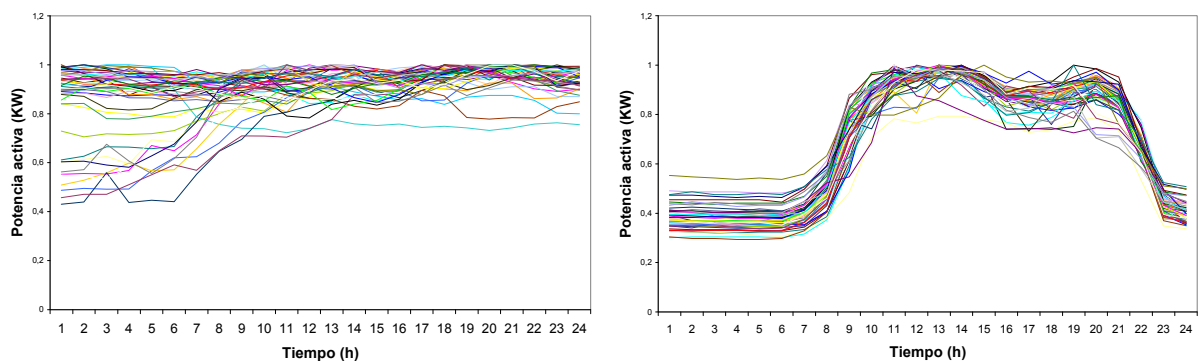


Figura 20. Perfiles de carga del consumidor Industria y del consumidor Universidad

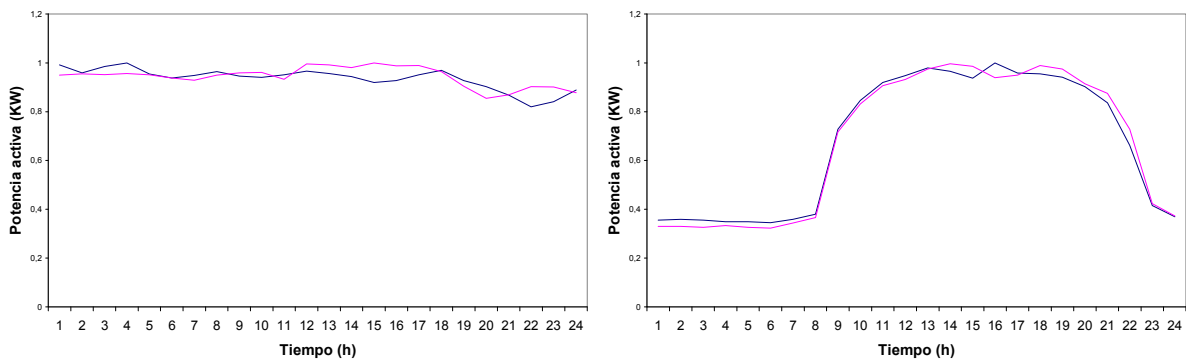


Figura 21. Perfiles Utilizados para el Test de capacidad de identificación de nuevos consumidores

#### 4.4.3. SEGUNDO CONJUNTO DE DATOS (VEINTISÉIS CONSUMIDORES)

El segundo conjunto de datos corresponde a veintiséis consumidores distintos, correspondientes a sectores españoles de tipo institucional, industrial, comercial y residencial. Las curvas de carga diarias fueron tomadas en periodos de invierno y verano de los años 2002 y 2003. El número total de vectores (días) utilizados en el diseño o entrenamiento fue de 451. En la Tabla 3 y Tabla 4 se muestra el número de vectores pertenecientes a cada tipo de consumidor y la etiqueta que corresponde a cada uno de ellos. Para el testeo se utilizaron cuatro conjuntos de datos no usados en el diseño o entrenamiento, concretamente, 25 vectores de tipo industria, 13 vectores de tipo institucional, 29 vectores de tipo comercial y 2 vectores de tipo residencial.

Tabla 3. Segundo conjunto de Datos (Distribución de Etiquetas)

Sector	Actividad	Etiqueta
Industria	Industrias Medianas	1,2
	Almacén	3
	Industrias Grandes	4,5,6,7,8,9
Institucional	Universidades Medianas	10,12,13
	Universidades Grandes	11,19
Comercial	Hoteles Pequeños	14,15,16,18
	Hotel mediano	17
	Hospitales y Centros Médicos	20,21,22,23
	Comercios Pequeños	24,25
Residencial	Pequeños consumidores	26



Tabla 4. Espectro de consumidores con sus etiquetas asociadas y la actividad a la que pertenecen.

Cliente	Etiqueta	Número de vectores de entrada	Actividad del cliente
I1	1	18	Industria
I2	2	18	Industria
I3	3	17	Industria
I4	4	16	Industria
I5	5	19	Industria
I6	6	18	Industria
I7	7	17	Industria
I8	8	16	Industria
I9	9	18	Industria
T1	10	16	Institucional
T2	11	11	Institucional
T3	12	16	Institucional
T4	13	16	Institucional
C1	14	17	Comercial
C2	15	18	Comercial
C3	16	19	Comercial
C4	17	18	Comercial
C5	18	17	Comercial
T5	19	16	Institucional
C6	20	20	Comercial
C7	21	19	Comercial
C8	22	18	Comercial
C9	23	19	Comercial
C10	24	17	Comercial
C11	25	20	Comercial
R1	26	17	Residencial

En las siguientes figuras se muestran las gráficas de los vectores (curvas de carga) utilizados para el segundo conjunto de datos.

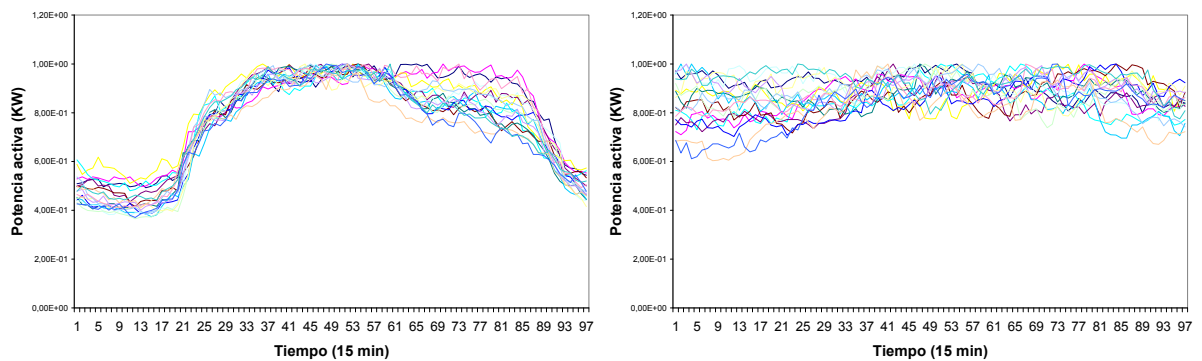


Figura 22. Curvas de Carga de los consumidores Industria I1 e I2

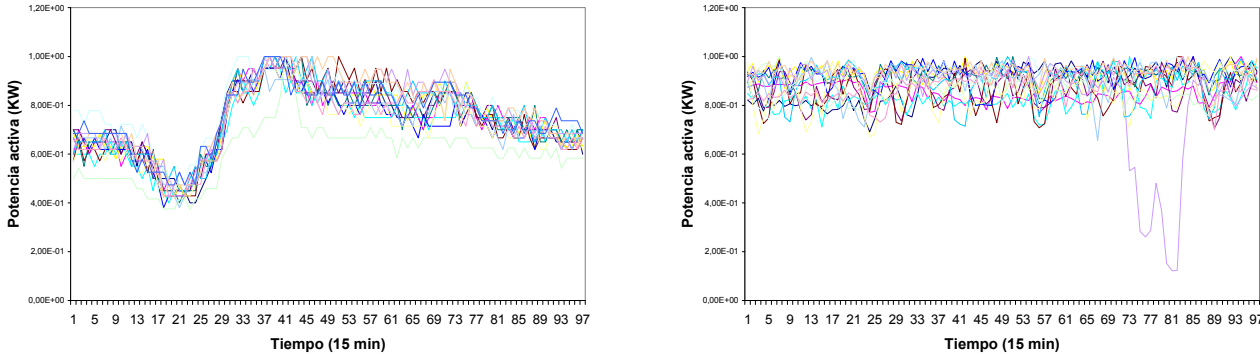


Figura 23. Curvas de carga de los consumidores Industrias I3 e I4

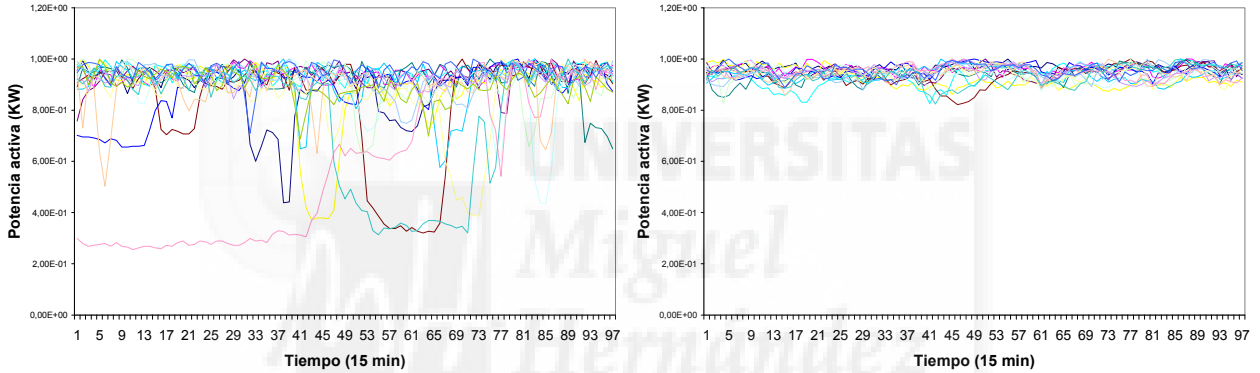


Figura 24. Curvas de carga de los consumidores Industria I5 e I6

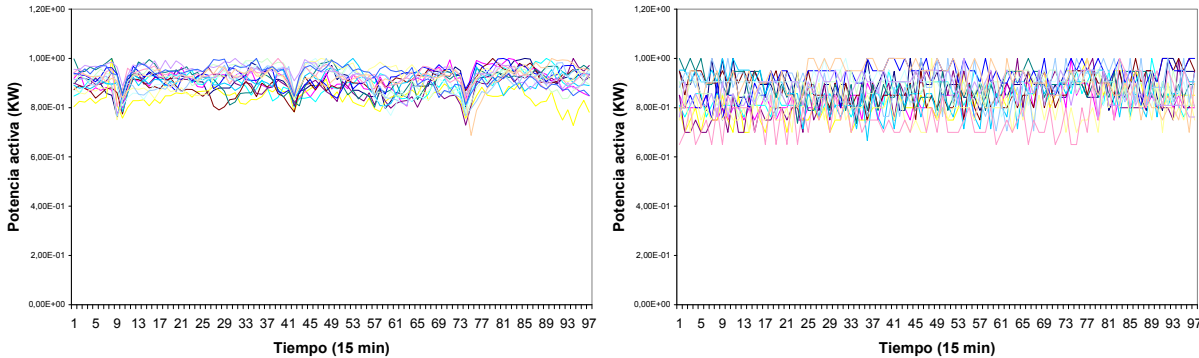


Figura 25. Curvas de carga de los consumidores Industria I7 e I8

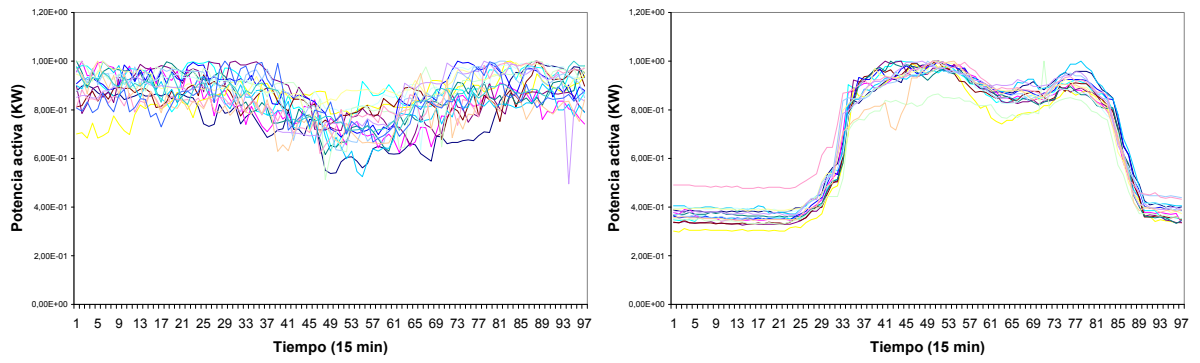


Figura 26. Curvas de carga de los consumidores Industria 19 e Institucional T1

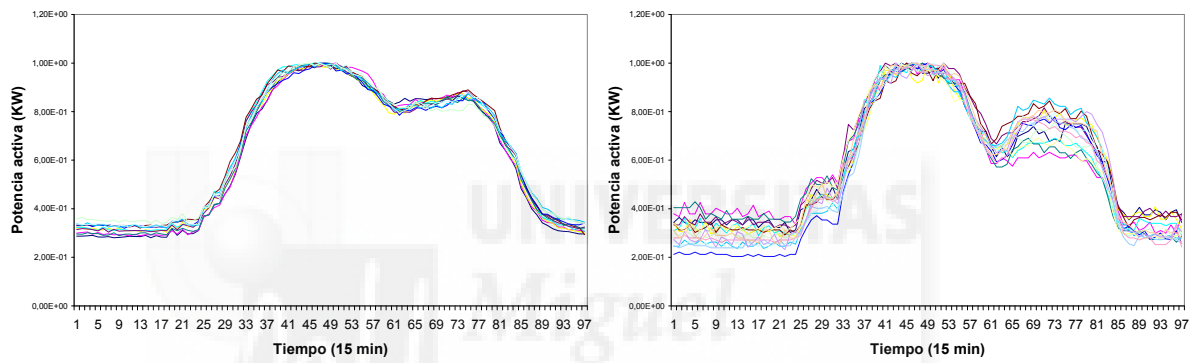


Figura 27. Curvas de carga de los consumidores Institucional T2 e Institucional T3

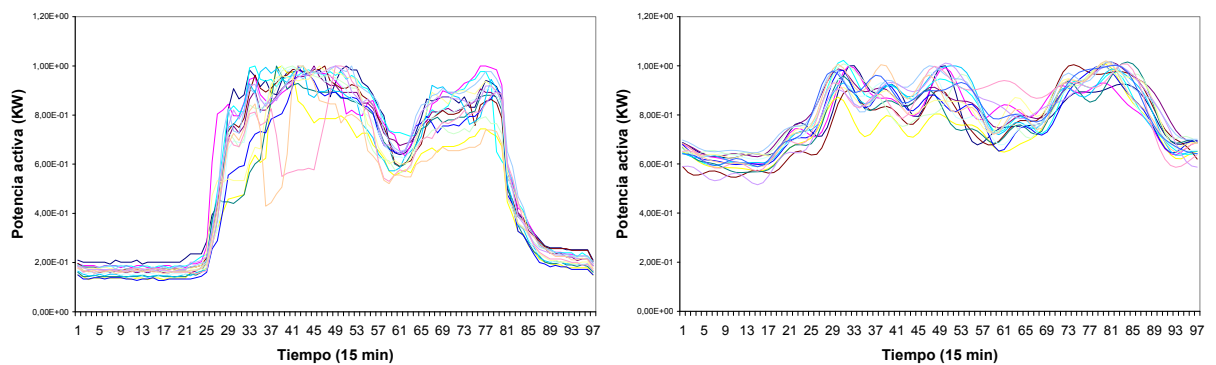


Figura 28. Curvas de carga de los consumidores Institucional T4 y Comercial C1

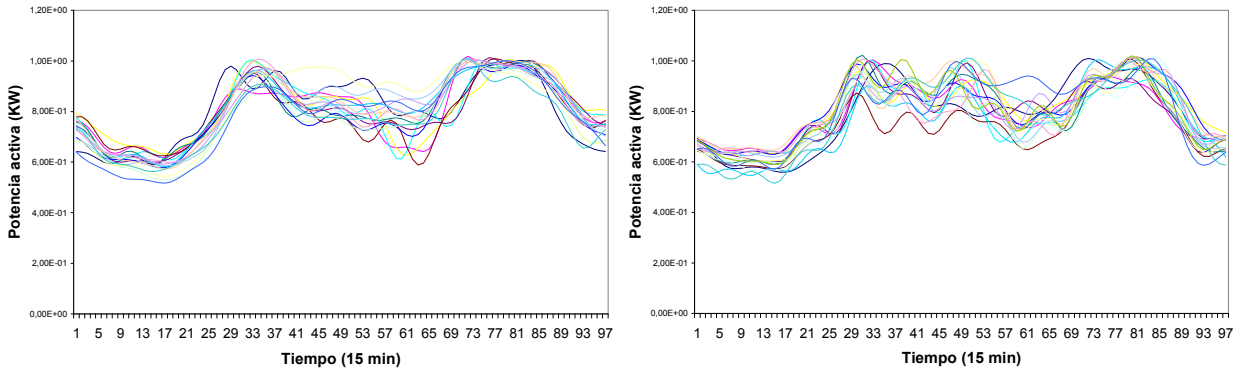


Figura 29. Curvas de Carga de los consumidores Comercial C2 y Comercial C3

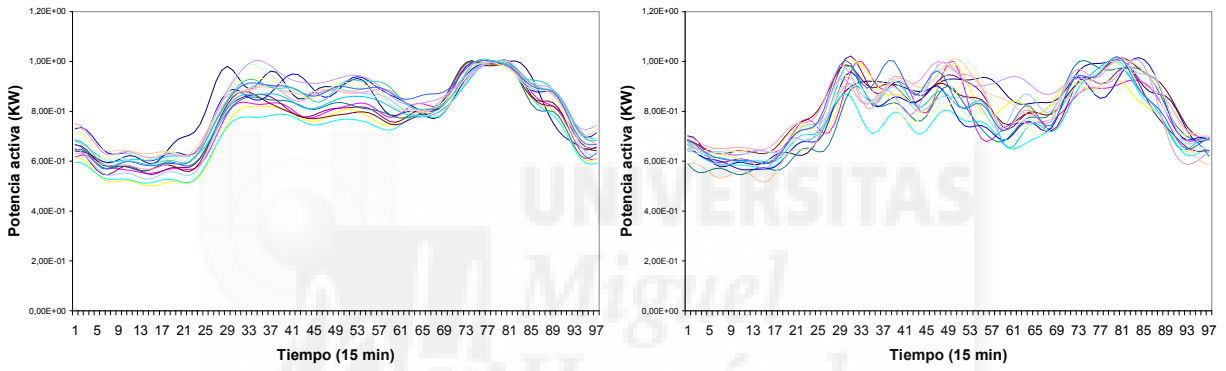


Figura 30. Curvas de Carga de los consumidores Comercial C4 y C5

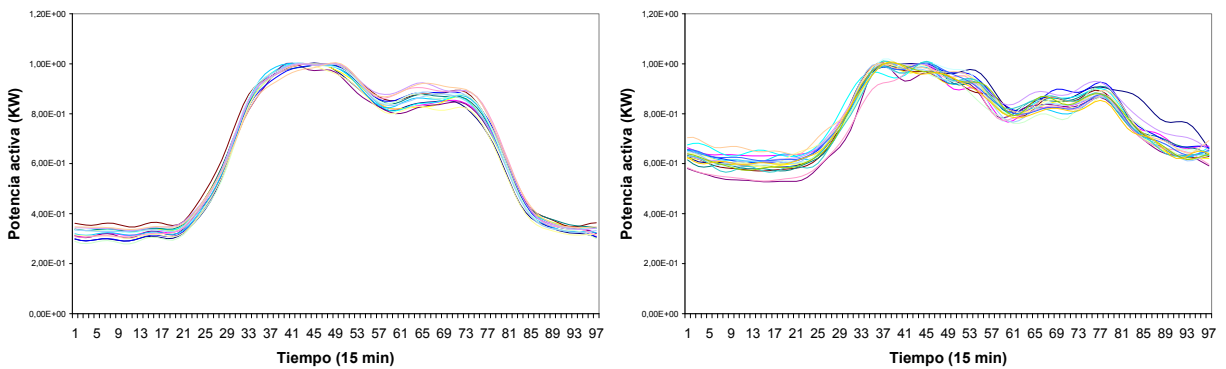


Figura 31. Curvas de Carga de los consumidores Institucional T5 y Comercial C6

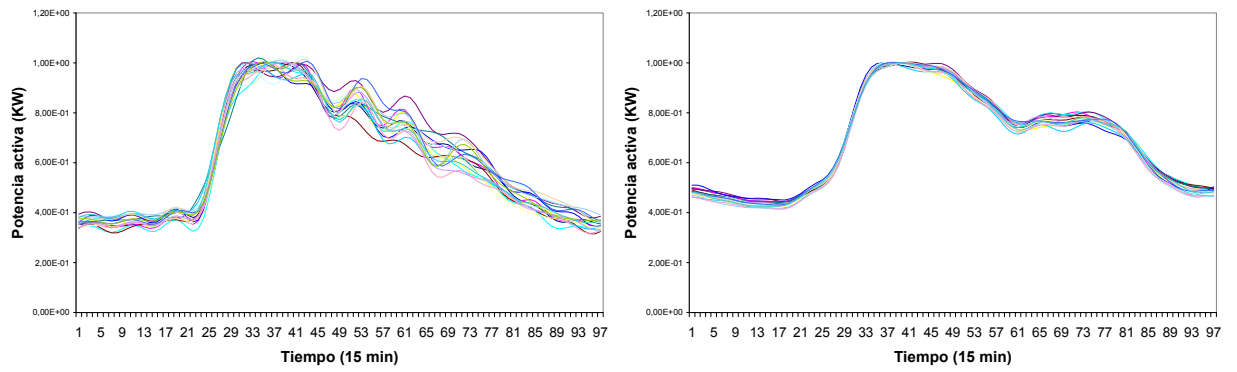


Figura 32. Curvas de Carga de los consumidores Comercial C7 y Comercial C8

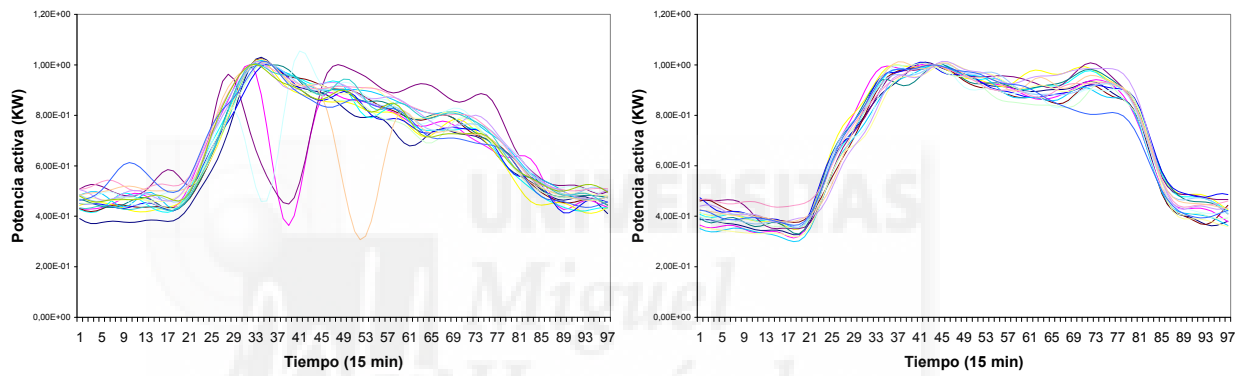


Figura 33. Curvas de Carga de los consumidores Comercial C9 y Comercial C10

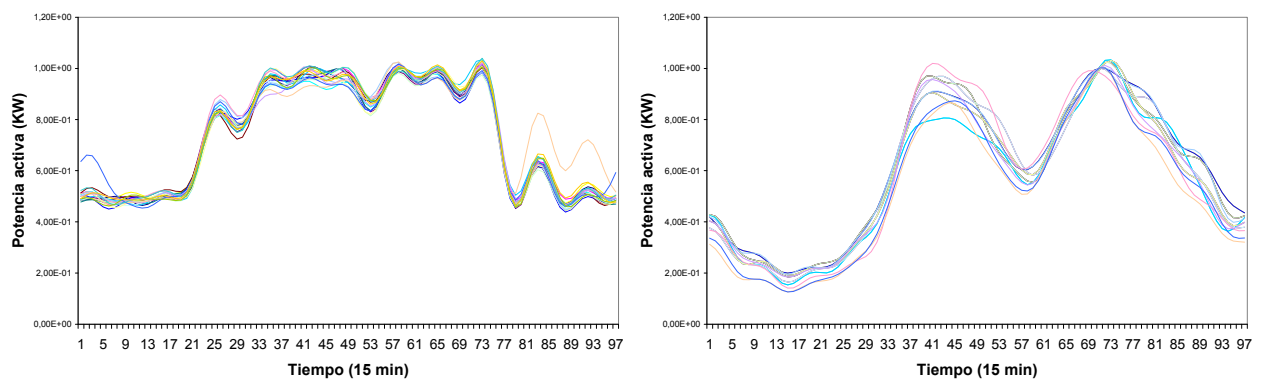


Figura 34. Curvas de Carga del consumidor Comercial C11 y Residencial R1

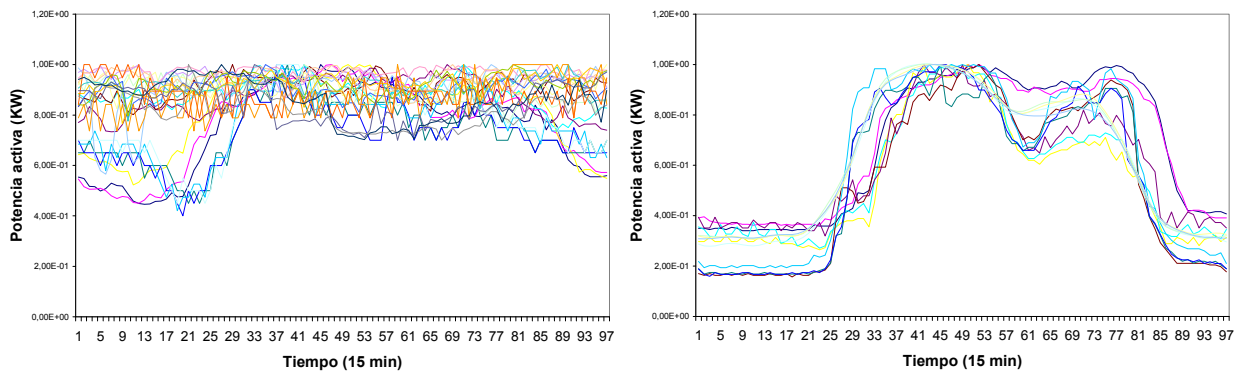


Figura 35. Curvas de carga utilizadas posteriormente en los Testeos para comprobar la capacidad de identificación de nuevos consumidores de cada una de las técnicas. Curvas de un consumidor Industrial y otro Institucional.

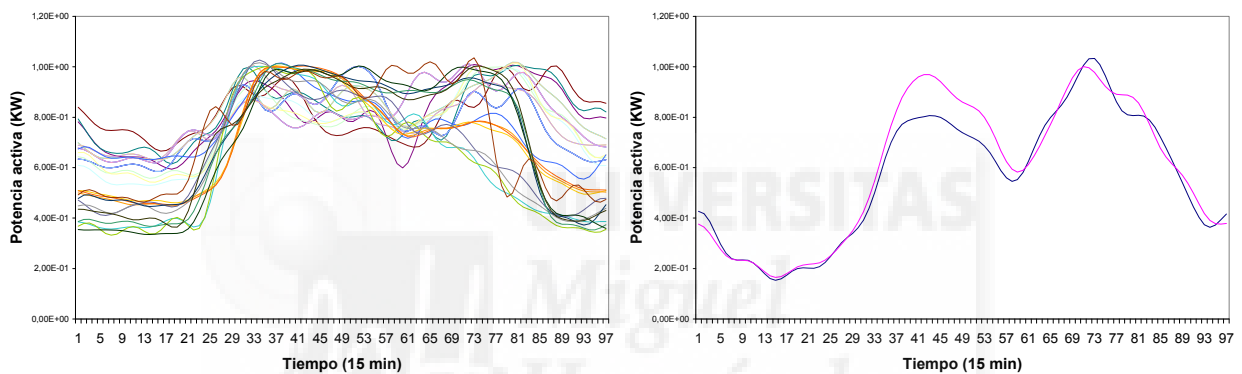


Figura 36. Curvas de carga utilizadas posteriormente en los Testeos para comprobar la capacidad de identificación de nuevos consumidores de cada una de las técnicas. Curvas de un consumidor Comercial y otro Residencial.

Tal y como se puede apreciar en las gráficas, los vectores utilizados para el segundo conjunto de datos corresponden a valores de la potencia activa tomados cada 15 minutos a lo largo de todo un día, obteniéndose curvas representadas por un total de 97 puntos y no de 24 como en el primer conjunto, donde el valor número 97 corresponde a las 0 horas del día siguiente. Los valores de la potencia activa también fueron normalizados entre 0 y 1, es decir el valor 1 corresponde al valor más alto de potencia alcanzado en la curva de carga que estamos normalizando.

Una vez analizadas y ensayadas las distintas técnicas y metodologías para el segundo conjunto de datos, se ha procedido a realizar una serie de tests utilizando sólo aquellas técnicas que han obtenido mejores resultados. Los tests han sido realizados con curvas de

carga de los mismos consumidores pero de épocas diferentes a las empleadas en los ensayos más algunas curvas de consumidores nuevos.

#### **4.4.4. ANÁLISIS DE LAS TÉCNICAS DE CLASIFICACIÓN**

El análisis y aplicación de cada técnica no es un proceso cerrado, ni sigue una serie de pasos estrictos, sino que depende de la técnica empleada en si misma. Existe además la posibilidad de combinar diferentes técnicas y diferentes rasgos de clasificación en el proceso, con lo cual existe una amplia variedad de formas de obtener la segmentación de los datos de entrada.

De forma general podemos distinguir tres pasos en el proceso de entrenamiento:

1. Reunir y tratar los datos de entrada.
2. Entrenar con los datos las distintas técnicas y diseñar la metodología a nivel de selección de parámetros adecuados.
3. Testear nuevas entradas (Testeo o Generalización).

Cabe destacar que antes del análisis se ha realizado un filtrado de datos para eliminar errores de medida o días anómalos de consumo. Sobre este tema se entrará con detalle en el capítulo 5.

#### **4.4.5. IDONEIDAD DE LAS TÉCNICAS**

Con el fin de evaluar la bondad de las técnicas, se emplearán índices que representen lo bien que aprende o generaliza una técnica en particular para comprobar de forma sencilla cuales son las que proporcionan mejores resultados.

Las técnicas que solamente clasifican no se les pueden aplicar ningún tipo de índice o medida de idoneidad, sino que debe ser el usuario el que compruebe si la clasificación es

o no aceptable, bien mediante gráficas o con algún otro método. Sin embargo, a las técnicas que además de clasificar los datos en grupos o clusters permiten la identificación de nuevos datos en los grupos realizados, sí que les es posible aplicar índices que representen la bondad de la agrupación realizada. Por tanto haremos en los puntos siguientes de este capítulo dos distinciones: Ensayos y simulaciones con técnicas que solo han permitido clasificar o segmentar datos y ensayos y simulaciones con técnicas que han permitido clasificar e identificar nuevos vectores (curvas de carga).

En esta investigación se han utilizado dos tipos distintos de índices, a los que se ha denominado *Error de aprendizaje* e *Índice de acierto*. El error de aprendizaje es el número de clientes utilizados en el diseño o entrenamiento de la técnica correspondiente que han sido identificados en un sector distinto del que proceden, es decir, si por ejemplo, un cliente o consumidor de tipo comercial es identificado como consumidor de tipo institucional después del diseño o entrenamiento, el error de aprendizaje aumentará en una unidad y así sucesivamente. Por tanto, aquellas técnicas cuyo error de aprendizaje sea cero, habrán identificado correctamente a todos los consumidores utilizados en el diseño o entrenamiento.

Por otra parte, el índice de acierto es un parámetro que varía entre cero y uno, correspondiendo el valor de uno a una clasificación perfecta. Este índice se aplica al testeo de clientes o consumidores eléctricos no empleados en el diseño o entrenamiento de las técnicas, y proporciona una medida de lo bien que la técnica ha logrado reconocer el mapping subyacente de los datos de entrada. Luego es posible que una técnica cuyo error de aprendizaje sea muy bajo, tenga un índice de acierto lejano a uno, lo que significaría que aunque ha logrado identificar los vectores de entrenamiento correctamente, ha fallado a la hora de identificar nuevos datos no utilizados en el entrenamiento o diseño, presentando una falta de lo que se denomina generalización, que es en realidad el propósito más importante a la hora de realizar una buena clasificación.



A continuación se presenta un resumen en forma de esquema de los experimentos realizados para aclarar los pasos seguidos:

1. Técnicas de clasificación. En este punto se probarán las distintas técnicas con los dos conjuntos de datos para comprobar si realizan una correcta segmentación.

2. Técnicas de clasificación e identificación

- Primer conjunto de datos (dos consumidores) con curvas de carga diaria de 24 valores, es decir, con un valor de potencia por hora.
  - Entrenamiento
  - Testeo
  - Selección de las metodologías o técnicas que obtengan mejores resultados
- Segundo conjunto de datos (veintiséis consumidores) con curvas de carga diaria de 96 valores, es decir, con un valor de potencia cada 15 minutos.
  - Entrenamiento
  - Testeo
  - Selección de las metodologías o técnicas que obtengan mejores resultados

## 4.5. RESULTADOS Y DISCUSIÓN

En este punto se exponen los resultados obtenidos, mostrando los mejores resultados para cada tipo de metodología o técnica analizada junto a una discusión de los mismos. Antes de continuar es necesario recordar la diferencia que existe entre las técnicas que únicamente segmentan (clusterizan o agrupan) los datos de entrada proporcionados y las técnicas que, además de segmentar los datos, permiten su utilización para identificar y clasificar (testear) nuevos datos no utilizados en el entrenamiento o diseño de la metodología. Por lo tanto, este punto se dividirá, inicialmente, en dos partes. La primera

parte será dedicada a aquellas técnicas que sólo permiten obtener un conjunto de segmentos (clusters) de los datos de entrada, esto es, agrupan los datos de acuerdo a las similitudes que éstos presentan, clasificándolos en un determinado grupo pero no permiten que les sean presentados datos no utilizados en esta clusterización previa para clasificarlos en los grupos realizados. Estas técnicas se han denominado únicamente y a modo de diferenciarlas de las otras como “*Técnicas de Clasificación*”.

La segunda parte se dedicará a las técnicas que además de segmentar los datos de entrada, definiendo su pertenencia a un determinado grupo (cluster) de datos, permiten el testeado con nuevos datos con el fin de identificarlos en uno de estos grupos. Estas técnicas se han denominado “*Técnicas de Clasificación e Identificación*”.

#### **4.5.1. TÉCNICAS DE CLASIFICACIÓN**

Tal y como se ha dicho anteriormente, las técnicas de clasificación simplemente permiten definir grupos de consumidores con características semejantes que pueden ser agrupados independientemente del sector o área a la que pertenezcan, con el fin de segmentar el mercado y así ayudar tanto a los proveedores de electricidad como a los clientes a la hora de seleccionar un tipo de política. En este sentido, como pueden ser el desarrollo de tarifas específicas, programas de eficiencia energética, puesta en marcha de generación distribuida, etc...

A continuación se muestran los resultados obtenidos mediante estas técnicas, visualizándolos gráficamente para una mejor comprensión.

##### **4.5.1.1. TÉCNICAS ESTADÍSTICAS**

###### ***ESCALAS MULTIDIMENSIONALES***

El escalado multidimensional permite visualizar lo próximos que están los puntos entre sí representando los datos en un pequeño número de dimensiones. Es una técnica

estadística multivariante que sirve para hacerse una idea de cómo es el mapping de los datos de entrada. Esta técnica no realiza ningún tipo de clustering, sino que simplemente reduce los datos a dos dimensiones para que éstos puedan ser visualizados y así poder apreciar cómo de juntos o distantes están unos datos de otros. En las siguientes figuras se representa un escalado multidimensional de los dos conjuntos de datos descritos en al comienzo de este Capítulo. El escalado multidimensional calcula la distancia Euclídea entre cada par de observaciones de la matriz de datos de entrada. Las filas corresponden a las observaciones y las columnas a las variables.

***Primer conjunto de datos (dos consumidores)***

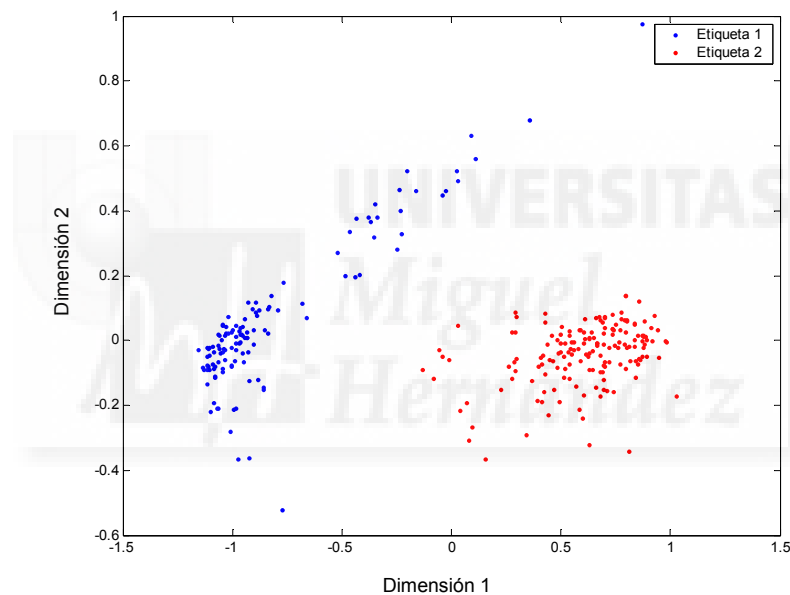


Figura 37. Escalado Multidimensional aplicado al primer conjunto de datos

Observando la Figura 37, se aprecia una clara separación de los dos tipos de consumidores, incluso se podría decir que son linealmente separables. La etiqueta 1 corresponde al consumidor de tipo industria, mientras que la etiqueta 2 corresponde al consumidor de tipo universidad. En este caso y debido a que el espectro de datos de entrada está formado por solo 2 consumidores la agrupación en clusters de ambos es claramente identificable.

### *Segundo conjunto de datos (veintiséis consumidores)*

Tras aplicar la técnica de escalado multidimensional al segundo conjunto de datos utilizado para la investigación, se obtiene el gráfico de la Figura 38.

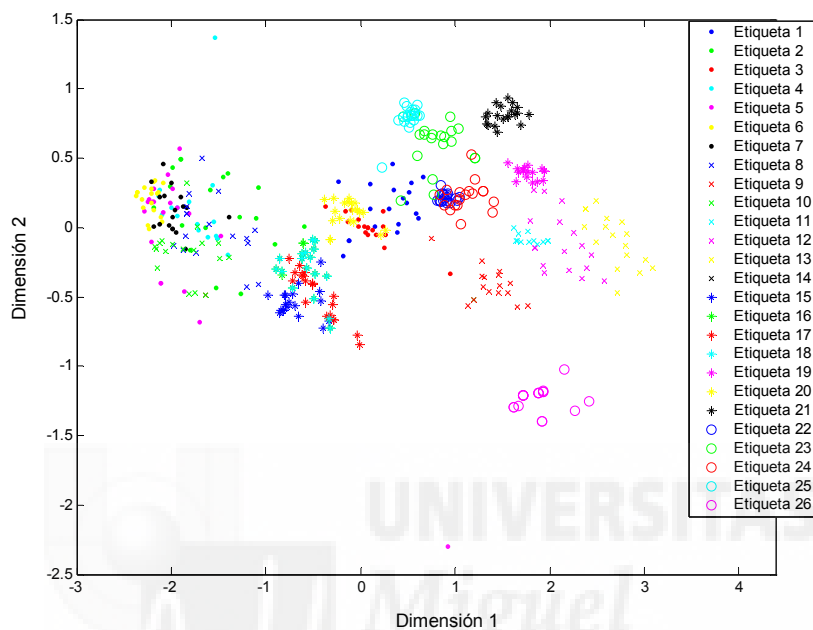


Figura 38. Aplicación de la Técnica de escalado multidimensional al segundo conjunto de datos

Aquí también se observa un agrupamiento claro en sectores, por ejemplo, los de tipo industrial (etiquetas 2-8, parte izquierda de la figura), consumidores tipo institucional (etiquetas 11-13, 19, parte derecha de la figura), consumidores tipo comercial (etiquetas 14-17, 20-23, zona central de la figura). También se observa que al consumidor tipo residencial con etiqueta 26 el mapa lo sitúa sólo en la zona inferior derecha de la figura, y que al consumidor tipo industria con etiqueta 1 lo sitúa más cerca de los de tipo comercial que de los de su propio sector.

### **MANOVA**

La técnica estadística multivariante de análisis de varianza o MANOVA [72] (Multivariate Analysis of Variance), en su implementación en MATLAB, toma un conjunto de datos de entrada dividido en grupos y permite determinar el grado de similitud o disimilitud entre los mismos.

Como la técnica MANOVA calcula los valores canónicos de las variables, la representación de las dos primeras variables canónicas mostrará más separación entre los grupos que la representación de cualquier par de variables originales. Además de poder visualizar las diferencias entre grupos representando las dos primeras variables canónicas, MANOVA también puede realizar una representación de los clusters en forma de dendograma para que resulte más sencillo comprobar la clusterización realizada mediante esta técnica. En un dendograma se muestra la separación entre grupos uniéndolos mediante un árbol jerárquico, de forma que dos grupos están más próximos entre sí cuanto más baja sea su unión, es decir, cuanto más cerca esté su nivel jerárquico. A continuación se mostrarán los resultados obtenidos al aplicar la técnica MANOVA a los dos conjuntos de datos disponibles.

***Primer conjunto de datos (dos consumidores)***

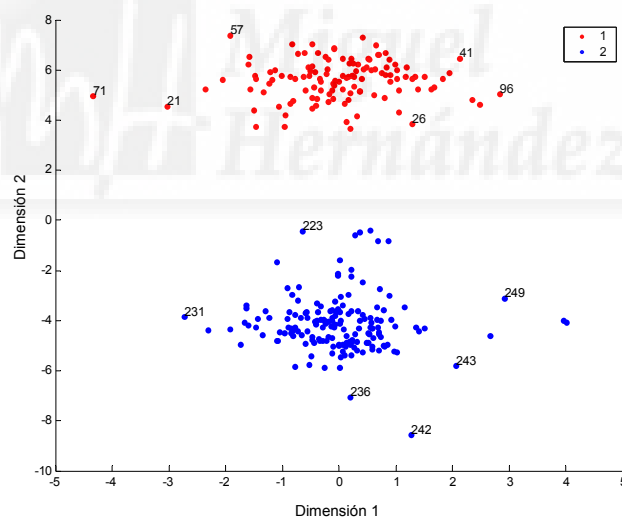


Figura 39. Mapa obtenido aplicando la técnica de MANOVA al primer conjunto de datos

Para el primer conjunto de datos, en el que sólo hay dos tipos de consumidores distintos, quizá no sea muy útil ver la distancia entre ambos. Lo único que puede apreciarse es que MANOVA clasifica muy bien ambos clusters.

Tal y como se puede observar en la Figura 39, al representar con MANOVA un conjunto de datos, es posible identificar cada punto o vector mediante el número de fila que ocupa en el archivo utilizado en la implementación de la técnica. Se han seleccionado algunos puntos para mostrarlo. En el caso del primer conjunto de datos no se ha realizado una representación en forma de dendograma debido a que no presenta interés, al existir un único nivel de jerarquía que une los dos clusters.

### *Segundo conjunto de datos (veintiséis consumidores)*

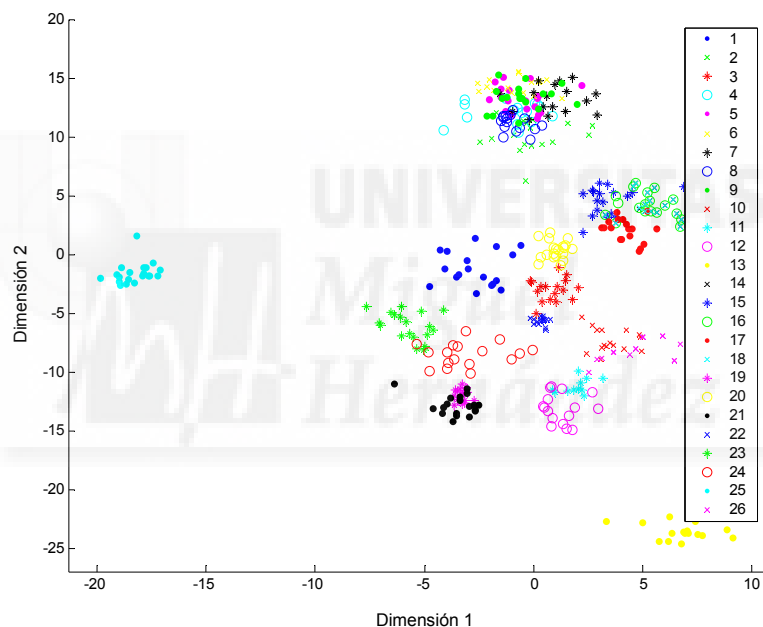


Figura 40. Mapa obtenido aplicando la técnica de MANOVA al segundo conjunto de datos

Esta clasificación, ver Figura 40, también resulta muy satisfactoria, ya que se puede apreciar cómo se agrupan los consumidores tipo industria en la parte superior de la figura, y los de tipo institucional en la parte inferior derecha. Los consumidores de tipo comercial aparecen más esparcidos, pero su mayor acumulación se presenta en la parte central y derecha de la figura, mientras que el consumidor de tipo residencial (etiqueta 26) se encuentra a la derecha de la figura junto a algunos consumidores de tipo institucional.

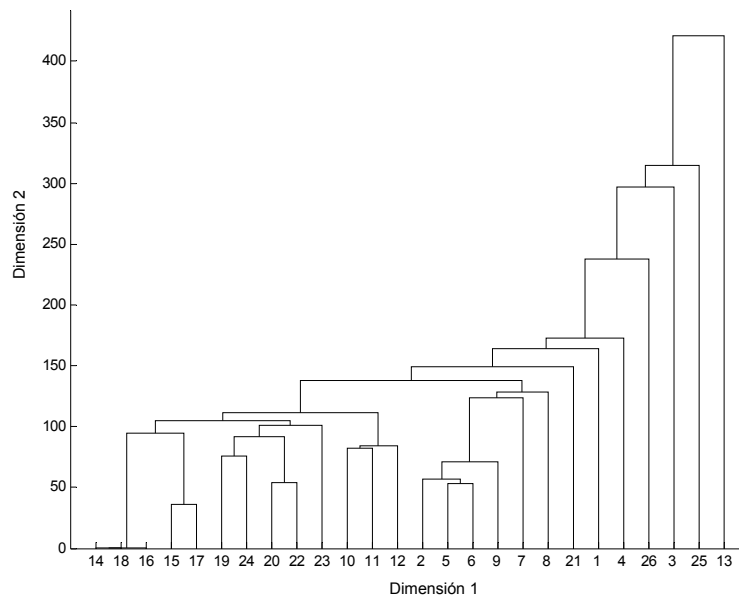


Figura 41. MANOVA, segundo conjunto de datos (dendograma)

En el dendograma resultante para el segundo conjunto de datos, ver Figura 41, se puede comprobar la clasificación realizada por la técnica MANOVA.

### ***K-MEANS***

La función denominada K-means, utilizada bajo programas desarrollados en Matlab, divide y agrupa los datos de entrada en K clusters, y devuelve un vector de índices indicando a cual de los K clusters ha sido asignado cada dato. Encuentra una partición en la cual los objetos dentro de cada cluster estén tan cerca entre ellos como sea posible, y tan lejos de los objetos de otros clusters como se pueda. Cada cluster en la partición es definido por su centro. El centro de cada cluster es el punto en el cual la suma de las distancias de todos los objetos en ese cluster es minimizada. Esta técnica estadística no reconoce etiquetas, tal y como sucede con las redes competitivas y fuzzy c-means, por lo tanto, los resultados obtenidos se mostrarán en los siguientes apartados haciendo uso de un escalado multidimensional de los mismos para facilitar su visualización y comprensión.

K-means usa un algoritmo iterativo que minimiza la suma de las distancias de cada punto al centro de su cluster. Este algoritmo mueve los objetos entre los clusters hasta que la suma no puede decrecer más. El resultado es un conjunto de clusters que son tan compactos y bien separados como sea posible.

Para hacerse una idea de los bien separados que están los clusters resultantes, se puede hacer una representación en forma de silueta de los índices de salida procedentes de k-means. Esta representación muestra una medida de lo cerca que un punto en un cluster está de los puntos de los clusters vecinos. Esta medida va de +1, indicando puntos que están muy distantes a los clusters vecinos, pasando por 0, indicando puntos que podrían estar indistintamente en un cluster o en otro, y a -1, indicando puntos que probablemente hayan sido asignados a un cluster equivocado. Por tanto, después de probar los tres conjuntos de datos con distintos números de clusters, se han seleccionado los mejores, los cuales se muestran a continuación.

### ***Primer conjunto de datos (dos consumidores)***

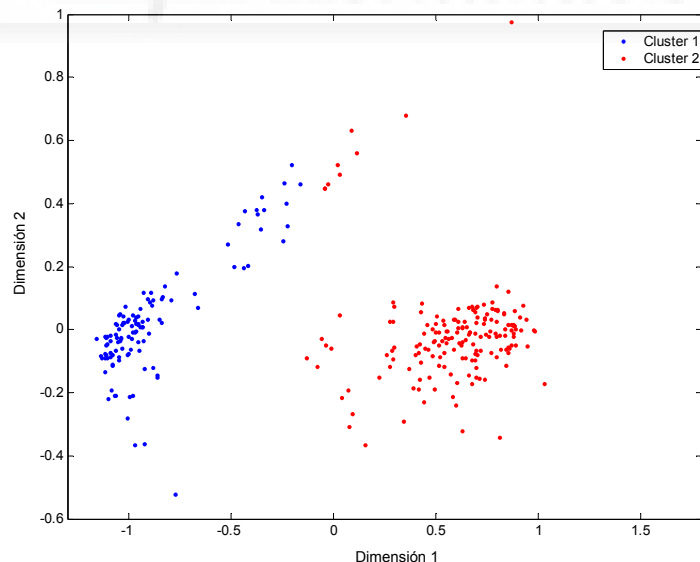


Figura 42. K-means, primer conjunto de datos (dos clusters)



En este primer experimento, Figura 42, k-means realiza una clasificación bastante buena. Prácticamente divide la figura en dos clusters distintos, asignando la pertenencia al cluster 2 (consumidor universidad) a algunos puntos situados en la parte superior, que en principio pertenecían al cluster 1 (consumidor industria).

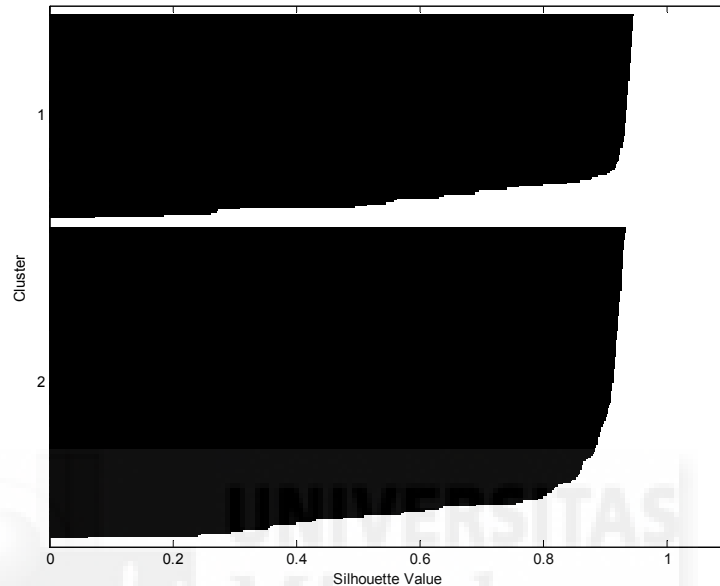


Figura 43. K-means, primer conjunto de datos (silueta, 2 clusters)

A partir de la representación en forma de silueta, Figura 43, se puede observar que la mayoría de puntos en los dos clusters tienen un valor grande de silueta, mayor que 0.8. Esto indica que dichos puntos están muy bien separados desde el cluster donde pertenecen con respecto al cluster vecino. Es decir, el valor de silueta de cada punto es una medida de lo similar que es este punto a otros dentro del mismo cluster y en comparación con otros puntos de otros clusters.

*Segundo conjunto de datos (veintiséis consumidores)*

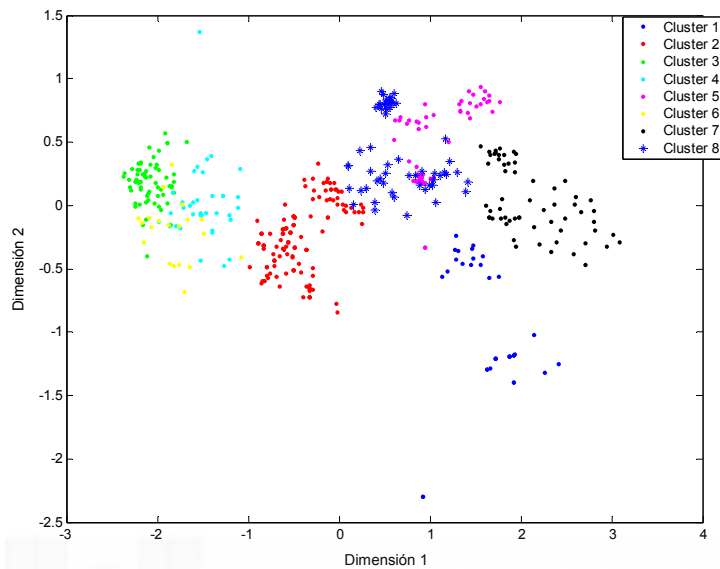


Figura 44. K-means, segundo conjunto de datos (8 clusters)

En esta ocasión ha sido posible obtener una clasificación óptima para el segundo conjunto de datos, que es aquella en la que queda dividido en 8 clusters distintos. En la Figura 44 se aprecia que el consumidor tipo residencial queda aislado en el cluster 6, los consumidores tipo comercial se encuentran en los clusters 1, 2, 4 y 7, los consumidores tipo institucional aparecen en los clusters 3, 7 y 8, mientras que los de tipo industria están en los clusters 2, 5, 7 y 8.

#### 4.5.1.2. REDES NEURONALES

##### ***REDES NEURONALES COMPETITIVAS***

Las redes neuronales competitivas son un tipo de redes no supervisadas y unidireccionales normalmente utilizadas en problemas de clasificación. Este tipo de redes no reconoce etiquetas a priori ya que es de tipo no supervisado, por lo que los datos se agrupan en clusters nombrados aleatoriamente. Esto supone una dificultad añadida para

comprobar si la red reconoce adecuadamente el mapping de los datos de entrada, así como para identificar nuevos datos no entrenados (esto no ocurre con la variante de red competitiva denominada Self-Organizing Map). Luego estas redes sí que permiten el testeado de nuevos datos, pero los asignan a grupos de nombres aleatorios, al no reconocer etiquetas. El empleo de capas competitivas que reconocen etiquetas será utilizado en las redes LVQ [54].

Han sido empleadas las capas competitivas para clusterizar los dos conjuntos de datos, probando a agruparlos en diferentes números de clusters para encontrar la manera en que esta técnica reconozca mejor el mapping de los datos de entrada. Así, para cada conjunto de datos se probará a dividirlo de forma que los sectores que se tienen a priori queden encuadrados en un mismo grupo o lo más cerca posible del mismo, si fuera posible. De no ser así, se tendría que un determinado tipo de consumidor no presenta características similares a los de su sector.

En las técnicas de clasificación o segmentación, con el fin de posibilitar una mejor comprensión de la clasificación realizada, se representarán los resultados utilizando un escalado multidimensional de los mismos para poder comprobar qué etiqueta se le ha asignado a cada vector, comparándolas con las etiquetas originales.

### ***Primer conjunto de datos (dos consumidores)***

En la Figura 45 se muestra un mapa con 2 clusters tras aplicar un escalado multidimensional a los resultados obtenidos mediante redes competitivas, se observa que tiende a agrupar o acercar a un mismo cluster datos que a priori no pertenecían al mismo grupo o sector. El cluster 1, formado por puntos de color azul pertenece al consumidor tipo industria y el cluster 2 formado por puntos de color rojo pertenece al consumidor tipo universidad.

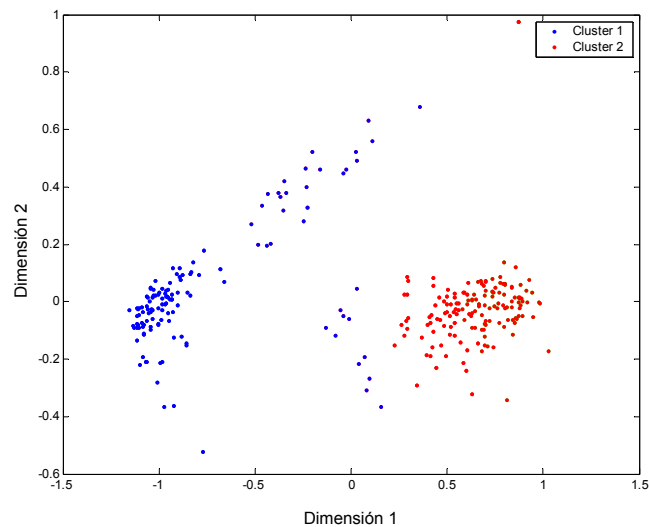


Figura 45. Aplicación de la Técnica de escalado multidimensional a los resultados obtenidos de entrenar el primer conjunto de datos con un red neuronal competitiva

En la Figura 46 se puede observar la clasificación que la red competitiva realiza utilizando 3 clusters distintos en vez de los dos que en principio existen para el primer conjunto de datos. En la figura se ve cómo ha mezclado datos de los dos grupos para realizar otro cluster (cluster 2). Esto significa que existen curvas de carga de ambos consumidores con perfiles algo similares que hacen que la red competitiva los agrupe en un tercer cluster al entrenar al red indicándole que encuentre más de 2 clusters.

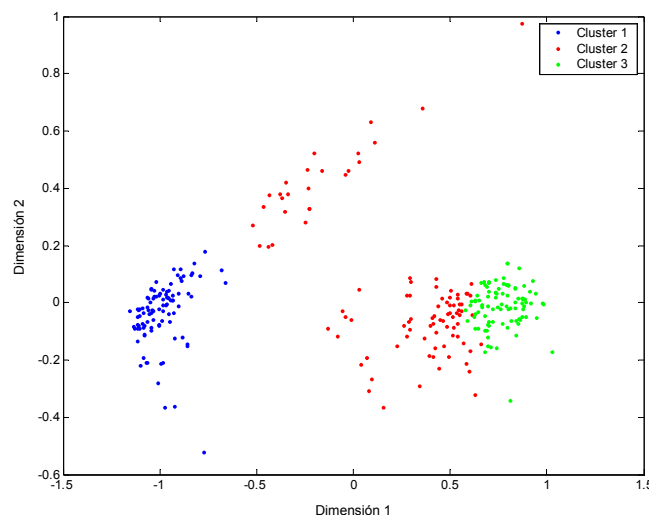


Figura 46. Aplicación de la Técnica de escalado multidimensional a los resultados obtenidos de entrenar el primer conjunto de datos con un red neuronal competitiva

***Segundo conjunto de datos (veintiséis consumidores)***

Tras entrenar el segundo conjunto de datos se obtienen 5 clusters. La Figura 47 representa las agrupaciones al aplicar la técnica de escalado multidimensional. En este caso quedan separados muy bien los consumidores del tipo residencial, comercial y algunos de tipo institucional (parte derecha de la figura, cluster 1, 4 y 5), pero el resto quedan mezclados en la parte izquierda de la Figura 47.

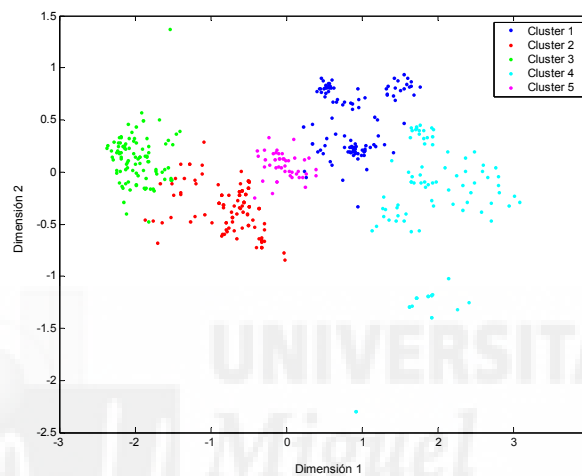


Figura 47. Aplicación de la técnica de escalado multidimensional a los resultados obtenidos de entrenar el segundo conjunto de datos con un red neuronal competitiva. Resultado tras solicitar 5 clusters.

***MAPAS AUTOORGANIZADOS DE KOHONEN (SELF ORGANIZING MAPS-SOM)***

En 1982 Teuvo Kohonen [55] presentó un modelo de red denominado mapas autoorganizados o SOM (*Self-Organizing Maps*), basado en ciertas evidencias descubiertas a nivel cerebral y con un gran potencial de aplicabilidad práctica. Este tipo de red se caracteriza por poseer un aprendizaje no supervisado competitivo. Los Mapas SOM son una variedad de las redes competitivas.

A diferencia de lo que sucede en el aprendizaje supervisado, en el no supervisado (o autoorganizado) no existe ningún maestro externo que indique si la red neuronal está

operando correcta o incorrectamente, pues no se dispone de ninguna salida objetivo hacia la cual la red neuronal deba tender. Así, durante el proceso de aprendizaje la red autoorganizada debe descubrir por sí misma rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e incorporarlos a su estructura interna de conexiones. Se dice, por tanto, que las neuronas deben autoorganizarse en función de los estímulos (datos) procedentes del exterior. Dentro del aprendizaje no supervisado existe un grupo de modelos de red caracterizados por poseer un aprendizaje competitivo. En el aprendizaje competitivo las neuronas compiten unas con otras con el fin de llevar a cabo una tarea dada. Con este tipo de aprendizaje, se pretende que cuando se presente a la red un patrón de entrada, sólo una de las neuronas de salida (o un grupo de vecinas) se active. Por tanto, las neuronas compiten por activarse, quedando finalmente una neurona como vencedora y anuladas el resto, que son forzadas a sus valores de respuesta mínimos. El objetivo de este aprendizaje es categorizar (clusterizar) los datos que se introducen en la red. De esta forma, las informaciones similares son clasificadas formando parte de la misma categoría y, por tanto, deben activar la misma neurona de salida. Las clases o categorías deben ser creadas por la propia red, puesto que se trata de un aprendizaje no supervisado, a través de las correlaciones entre los datos de entrada.

Un modelo SOM está compuesto por dos capas de neuronas (entrada - salida). La capa de entrada (formada por  $N$  neuronas, una por cada variable de entrada) se encarga de recibir y transmitir a la capa de salida la información procedente del exterior. La capa de salida (formada por  $M$  neuronas) es la encargada de procesar la información y formar el mapa de rasgos. Normalmente, las neuronas de la capa de salida se organizan en forma de mapa bidimensional como se muestra en la Figura 48, aunque a veces también se utilizan capas de una sola dimensión (cadena lineal de neuronas) o de tres dimensiones (paralelepípedo).

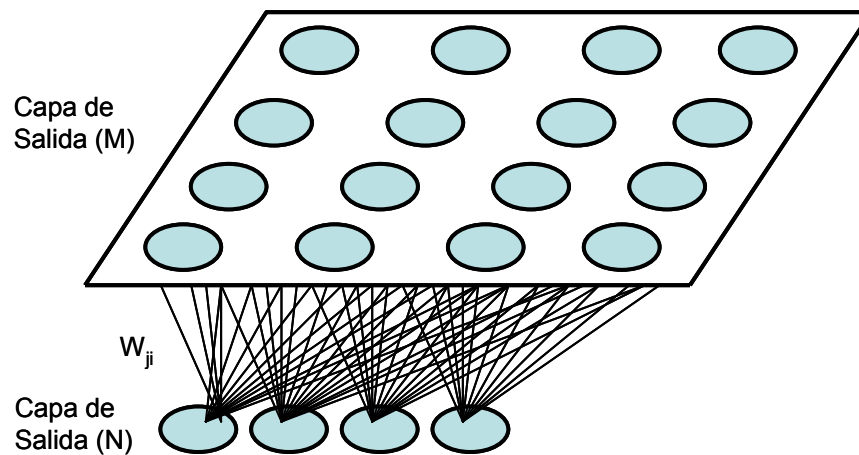


Figura 48. Arquitectura del SOM.

Las conexiones entre las dos capas que forman la red son siempre hacia delante, es decir, la información se propaga desde la capa de entrada hacia la capa de salida. Cada neurona de entrada  $i$  está conectada con cada una de las neuronas de salida  $j$  mediante un peso  $w_{ji}$ . De esta forma, las neuronas de salida tienen asociado un vector de pesos  $W_j$  llamado vector de referencia (o *codebook*), debido a que constituye el vector prototipo (o promedio) de la categoría representada por la neurona de salida  $j$ . Entre las neuronas de la capa de salida, puede decirse que existen conexiones laterales de excitación e inhibición implícitas, pues aunque no estén conectadas, cada una de estas neuronas va a tener una cierta influencia sobre sus vecinas. Esto se consigue a través de un proceso de competición entre las neuronas y de la aplicación de una función denominada de vecindad como se verá más adelante.

***Primer conjunto de datos (dos consumidores)***

En la Figura 49 se muestra el mapa obtenido entrenándolo con el espectro de datos formado por 2 consumidores.

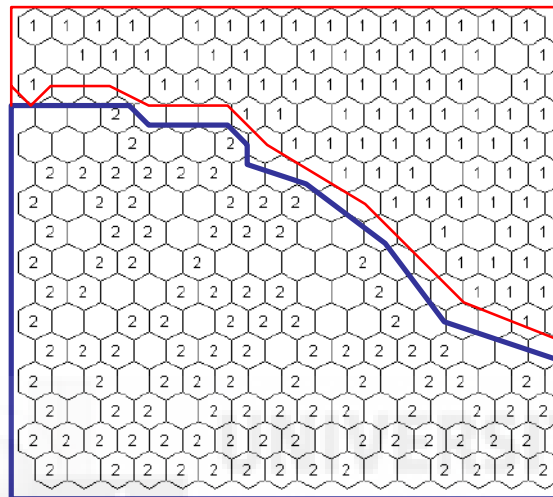


Figura 49. Mapa SOM de etiquetas obtenido tras un entrenamiento con el primer conjunto de datos.

Observando el mapa entrenado es fácil identificar los dos clusters de datos. Por un lado aparecen agrupadas en la parte superior del mapa las etiquetas número 1 pertenecientes a las curvas de carga del consumidor tipo industrial, y por otro lado aparecen agrupadas en la parte media e inferior del mapa las etiquetas número 2 pertenecientes al consumidor universidad. El mapa SOM ha logrado separar perfectamente ambos consumidores.

***Segundo conjunto de datos (veintiséis consumidores)***

La Figura 50 muestra el mapa obtenido entrenando una red autoorganizada de 23 x 23 neuronas con el segundo espectro de datos de entrada.



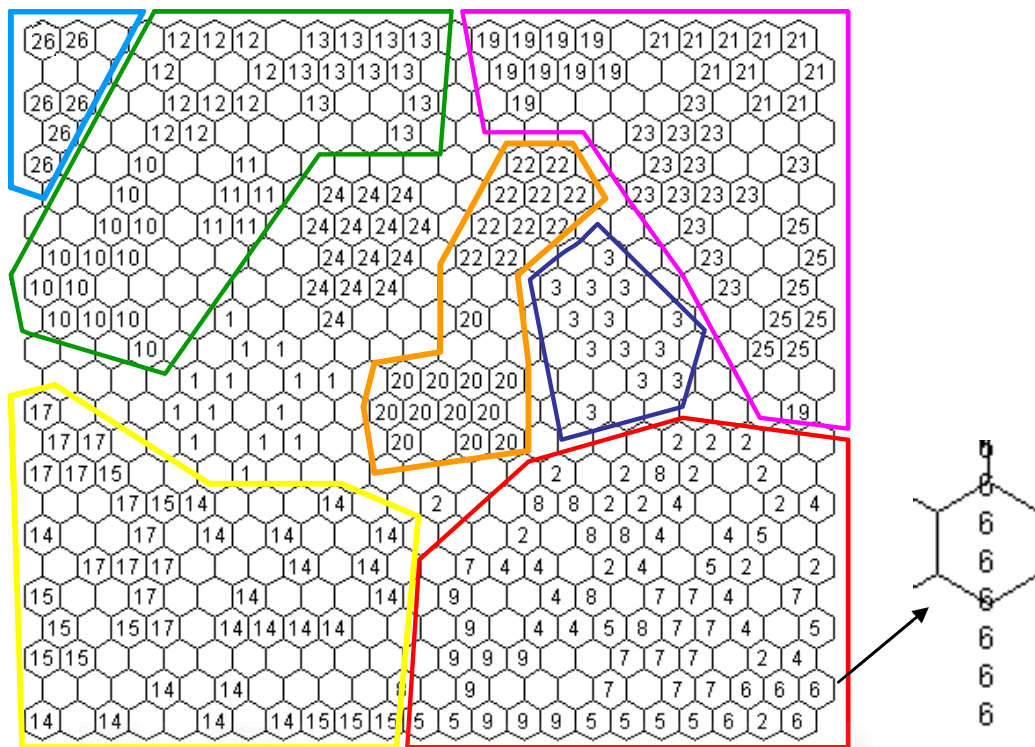


Figura 50. Mapa SOM de etiquetas obtenido tras un entrenamiento con el segundo conjunto de datos.

Observando el mapa entrenado es posible identificar clusters de consumidores con perfiles o características similares. Los 26 consumidores quedan clasificados o agrupados en siete zonas diferentes. Los hoteles han sido agrupados en la esquina inferior izquierda del mapa (etiquetas 14-17), las industrias grandes y una mediana (etiquetas 2 y de la 4 a la 9) también han sido agrupadas en un único cluster en la parte inferior derecha del mapa. En la parte superior derecha han sido agrupados los consumidores correspondientes a Hospitales y centros médicos (etiquetas 21 y 23) junto con un comercio pequeño (etiqueta 25) y una Universidad Grande (etiqueta 19). Por último el consumidor de tipo residencial ha sido situado en la esquina superior izquierda muy próximo al cluster formado por las universidades (etiquetas 10-13). Ver Figura 50.

## 4.5.1.3. LOGICA FUZZY

***FUZZY C-MEANS CLUSTERING***

Fuzzy C-means [111] es una técnica de clusterización de datos donde cada uno de los datos pertenece a un cluster (grupo) en un grado determinado que es especificado mediante un grado de pertenencia. Provee un método que muestra cómo agrupar datos que llenan algún espacio multidimensional en un número específico de clusters. Esta técnica, al igual que la de capas competitivas, no reconoce etiquetas de patrones, por lo que resulta muy complicado comprobar los resultados de su clasificación. A pesar de ello, se ha creído conveniente utilizarla para ver cómo agrupa a los consumidores en clusters y después visualizar el resultado mediante un escalado multidimensional.

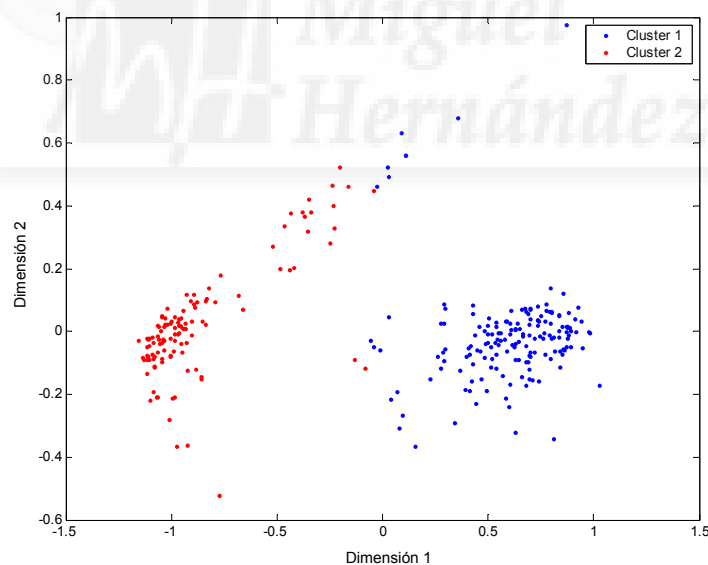
***Primer conjunto de datos (dos consumidores)***

Figura 51. Fuzzy c-means, primer conjunto de datos (2 clusters)

Se puede observar en la Figura 51 que algunos vectores han cambiado de grupo, esto es debido a que Fuzzy c-means los ha agrupado de esta forma, sin embargo la separación obtenida entre clusters es muy buena.

***Segundo conjunto de datos (veintiséis consumidores)***

En este caso se optó por solicitar al programa 4 y 12 divisiones máximas de clusters. A continuación se muestra el mapa obtenido solicitando 4 clusters, ver Figura 52.

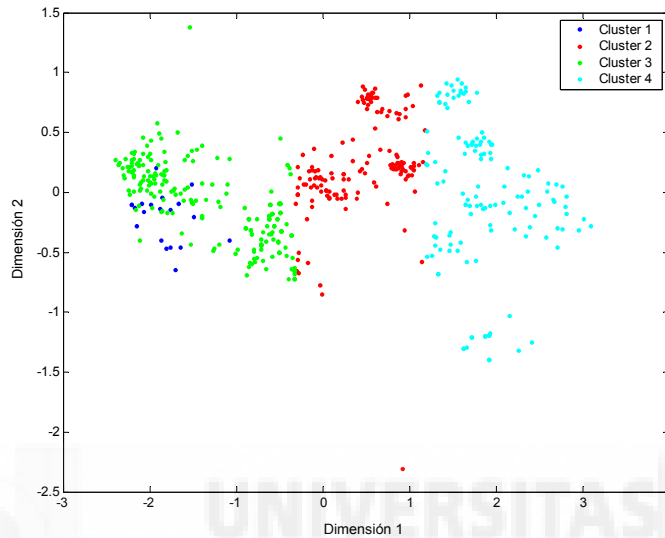


Figura 52. Fuzzy c-means, segundo conjunto de datos (4 clusters)

Esta clasificación podría catalogarse como muy buena, ya que c-means parece haber reconocido de forma aceptable el mapping de los datos. Se observa una división de consumidores de tipo institucional y comercial (cluster 4), otros consumidores tipo comercial (cluster 2), consumidores tipo industria (cluster 3) y otros consumidores tipo institucional (cluster 1).

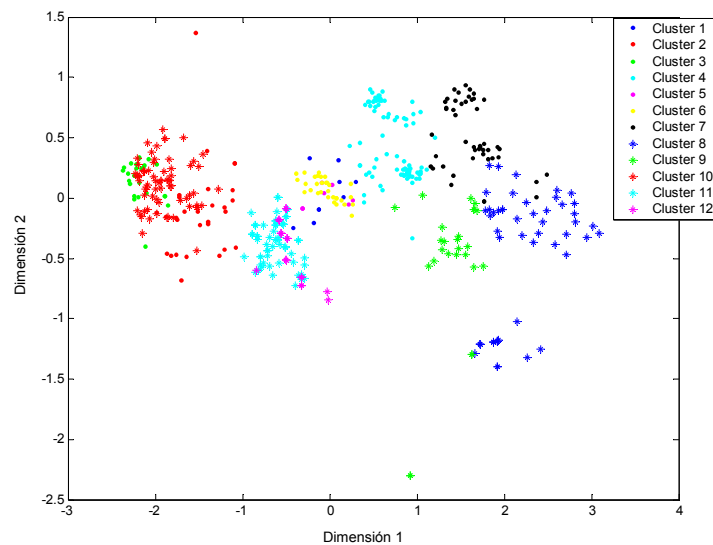


Figura 53. Fuzzy c-means, segundo conjunto de datos (12 clusters)

Si se clasifica el segundo conjunto de datos en 12 clusters, ver Figura 53, se puede distinguir bastante bien los distintos sectores. Por ejemplo, los consumidores tipo institucional se sitúan en los clusters 1 y 6, los de tipo residencial en el cluster 9 y los de tipo comercial en los clusters 3, 4, 5, 7, 10 y 11. Los consumidores del tipo institucional aparecen más mezclados.

#### 4.5.1.4. CONCLUSIONES DE LAS TECNICAS DE CLASIFICACION

A modo resumen, las técnicas analizadas para la clasificación (agrupación) de datos de entrada han demostrado ser herramientas muy útiles para la segmentación de consumidores eléctricos con el fin de diferenciar clusters o grupos distintos en los datos y así poder colocar juntos consumidores semejantes, diferenciándolos del resto, a pesar de no poder identificar nuevos consumidores en los clusters realizados. Así, las técnicas estadísticas de Escalas Multidimensionales y MANOVA, son técnicas que permiten la visualización de datos n-dimensionales (24 y 97 dimensiones en el caso del presente estudio), reduciéndolos a sólo dos dimensiones para permitir su visualización y comprobar así qué datos o clusters se encuentran más cercanos o más alejados. Además, la técnica MANOVA realiza un clustering de los datos, cuya visualización en forma de

dendograma facilita en gran medida su comprensión. Por otra parte, las otras tres técnicas de segmentación utilizadas: Redes Competitivas, Fuzzy C-Means y la técnica estadística K-Means, han resultado ser muy adecuadas para el propósito que se les encomendaba, que no era otro que descubrir la clusterización más óptima de los datos de entrada, utilizando para ello diferentes números de clusters para elegir la mejor división posible. Entre estas tres técnicas cabe resaltar los buenos resultados obtenidos mediante K-Means, técnica que además proporciona una forma muy interesante de medir la bondad de la agrupación realizada mediante la visualización de las siluetas resultantes de la clasificación. Los Mapas SOM han proporcionado no solo una importante agrupación sino que también han proporcionado una clara y legible representación gráfica utilizando el software Matlab [112] y las librerías disponibles de Self-Organizing Maps.

#### **4.5.2. TECNICAS DE CLASIFICACION E IDENTIFICACION**

En este apartado se muestran los resultados obtenidos con las técnicas que además de segmentar los datos de entrada permiten identificar nuevos consumidores en los segmentos o clusters realizados. Para ello se realizarán experimentos mediante los dos conjuntos de datos definidos anteriormente con las distintas técnicas analizadas en el estudio. Para mostrar los resultados se realizará una tabla para cada test realizado dentro de un conjunto de datos determinado.

Las tablas se componen del tipo de técnica empleada, el error de aprendizaje (ERA), que es el error cometido por la técnica en el proceso de entrenamiento, y el resultado del testeo o identificación de nuevos consumidores, índice de acierto (IA), que es el resultado obtenido al introducir a la red entrenada o técnica diseñada los datos de cada test, los cuales no fueron empleados en el entrenamiento o diseño previo.

El ERA, por tanto, es el número de vectores (curvas de carga de un consumidor) que siendo utilizados como datos de entrada a la técnica para su entrenamiento, se vuelven a utilizar para conocer si el mapa los asocia o no con las neuronas etiquetadas. Es decir, se testea la técnica con los mismos datos de entrada, para conocer la capacidad de

aprendizaje a la hora de asignarles su misma etiqueta. Un valor de ERA alto significa que no ha logrado identificar correctamente los vectores. Un valor de ERA igual a cero significa que el error de aprendizaje es nulo, la técnica analizada es capaz de identificar y asociar correctamente los vectores con sus etiquetas.

El resultado del testeo se compone a su vez de tres campos: etiquetas identificadas, donde se muestra la etiqueta asignada por la técnica a los datos de test, el tipo de consumidor identificado, es decir, el porcentaje de vectores que la técnica ha identificado en una etiqueta concreta, y el índice de acierto (IA).

El IA proporciona un valor para medir la bondad de la técnica. En este sentido el valor 1 del IA significa identificación perfecta y el valor 0 del IA todo lo contrario. Se alcanza el valor 1 cuando todos los nuevos vectores presentados al mapa entrenado (vectores de testeo) son identificados como pertenecientes a un cluster y tipo de consumidor similar, por tanto, al que realmente pertenece este nuevo vector (curva de carga). En caso contrario, es decir, cuando los nuevos vectores no son asociados a neuronas (clusters) de tipos de consumidores similares al tipo de consumidor al que pertenecen estas nuevas curvas, se obtiene un valor de 0.

#### 4.5.2.1. SEGUNDO CONJUNTO DE DATOS (VEINTISÉIS CONSUMIDORES)

Para la segmentación del segundo conjunto de datos, se ha omitido, además de las redes Perceptrón simple y Adalina por su simplicidad, dos algoritmos MLP (Levenberg-Marquardt y BFGS) y la técnica borrosa ANFIS. Estas tres técnicas se han omitido debido a los tiempos de ejecución que han sido muy altos.

**Test 1 (Consumidor tipo Industria)**

En este caso, cada una de las técnicas analizadas ha sido testeada con curvas de carga de un consumidor tipo Industrial. En total fueron testeados 25 nuevos vectores. Ver Tabla 5.

Tabla 5. Resultados, segundo conjunto de datos, test 1

Técnica Empleada	ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA	
PERCEPTRÓN MULTICAPA (MLP)	Gradiente Descendiente	243	24 vectores con etiqueta 6 1 vector con etiqueta 20	I6 96% C6 4%	0.96
	Gradiente Descendiente con Momento	274	6 vectores con etiqueta 2 16 vectores con etiqueta 8 2 vectores con etiqueta 11 1 vector con etiqueta 17	I2 24% I8 64% T2 8% C4 4%	0.88
	Ratio de Aprendizaje Adaptativo	166	4 vectores con etiqueta 2 3 vectores con etiqueta 4 12 vectores con etiqueta 8 3 vectores con etiqueta 9 1 vector con etiqueta 16 1 vector con etiqueta 17 1 vector con etiqueta 25	I2 16% I4 12% I8 48% I9 12% C3 4% C4 4% C11 4%	0.88
	Ratio de Aprendizaje Adaptativo + Momento	60	2 vectores con etiqueta 1 3 vectores con etiqueta 2 4 vectores con etiqueta 3 1 vector con etiqueta 5 9 vectores con etiqueta 6 1 vector con etiqueta 7 1 vector con etiqueta 8 1 vector con etiqueta 9 1 vector con etiqueta 15 1 vector con etiqueta 16 1 vector con etiqueta 23	I1 8% I2 12% I3 16% I5 4% I7 36% I8 4% I9 4% C2 4% C3 4% C9 4%	0.88
	Resistente	3	3 vectores con etiqueta 1 4 vectores con etiqueta 3 2 vectores con etiqueta 4 8 vectores con etiqueta 6 2 vectores con etiqueta 7 6 vectores con etiqueta 9	I1 12% I3 16% I4 8% I6 32% I7 8% I9 24%	1
	Fletcher-Reeves	10	3 vectores con etiqueta 1 4 vectores con etiqueta 3 13 vectores con etiqueta 6 5 vectores con etiqueta 9	I1 12% I3 16% I6 52% I9 20%	1
	Polak-Ribière	5	3 vectores con etiqueta 1 4 vectores con etiqueta 3 1 vector con etiqueta 5 12 vectores con etiqueta 6 5 vectores con etiqueta 9	I1 12% I3 16% I5 4% I6 48% I9 20%	1

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
	Powell-Beale	6	3 vectores con etiqueta 1 1 vector con etiqueta 2 4 vectores con etiqueta 3 3 vectores con etiqueta 5 9 vectores con etiqueta 6 1 vector con etiqueta 7 3 vectores con etiqueta 9	I1 12% I2 4% I3 16% I5 12% I6 36% I7 4% I9 12%	1
	Gradiente conjugado escalado	4	3 vectores con etiqueta 1 1 vector con etiqueta 2 4 vectores con etiqueta 3 1 vector con etiqueta 4 9 vectores con etiqueta 6 3 vectores con etiqueta 7 4 vectores con etiqueta 9	I1 12% I2 4% I3 16% I4 4% I6 36% I7 12% I9 16%	1
	Algoritmo secante de un paso	6	3 vectores con etiqueta 1 6 vectores con etiqueta 3 10 vectores con etiqueta 6 1 vector con etiqueta 7 1 vector con etiqueta 8 4 vectores con etiqueta 9	I1 12% I3 24% I6 40% I7 4% I8 4% I9 16%	1
LVQ		76	3 vectores con etiqueta 1 6 vectores con etiqueta 3 9 vectores con etiqueta 9 2 vectores con etiqueta 22 5 vectores con etiqueta 25	I1 12% I3 24% I9 36% C8 8% C11 20%	0.72
SOM		0	14 vectores con etiqueta 1 9 vectores con etiqueta 5 2 vectores con etiqueta 8	I1 56% I5 36% I8 8%	1
ELMAN		7	3 vectores con etiqueta 1 3 vectores con etiqueta 2 4 vectores con etiqueta 3 1 vector con etiqueta 4 7 vectores con etiqueta 6 4 vectores con etiqueta 7 3 vectores con etiqueta 9	I1 12% I2 12% I3 16% I4 4% I6 28% I7 16% I9 12%	1
RBF		0	15 vectores con etiqueta 14 6 vectores con etiqueta 16 4 vectores con etiqueta 18	C1 60% C3 24% C5 16%	0
GRNN		0	3 vectores con etiqueta 1 4 vectores con etiqueta 2 4 vectores con etiqueta 3 1 vector con etiqueta 5 4 vectores con etiqueta 6 5 vectores con etiqueta 7 1 vector con etiqueta 8 3 vectores con etiqueta 9	I1 12% I2 16% I3 16% I5 4% I6 16% I7 20% I8 4% I9 12%	1
PNN		0	3 vectores con etiqueta 1 4 vectores con etiqueta 2 4 vectores con etiqueta 3 1 vector con etiqueta 5 4 vectores con etiqueta 6 5 vectores con etiqueta 7 1 vector con etiqueta 8 3 vectores con etiqueta 9	I1 12% I2 16% I3 16% I5 4% I6 16% I7 20% I8 4% I9 12%	1



Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
CLUSTERING SUSTRATIVO		0	3 vectores con etiqueta 1	I1 12%	1
			3 vectores con etiqueta 2	I2 12%	
			4 vectores con etiqueta 3	I3 16%	
			1 vector con etiqueta 4	I4 4%	
			1 vector con etiqueta 5	I5 4%	
			4 vectores con etiqueta 6	I6 16%	
			5 vectores con etiqueta 7	I7 20%	
			1 vector con etiqueta 8	I8 4%	
			3 vectores con etiqueta 9	I9 12%	
ANÁLISIS DISCRIMINANTE		0	3 vectores con etiqueta 1	I1 12%	1
			2 vectores con etiqueta 2	I2 8%	
			4 vectores con etiqueta 3	I3 16%	
			7 vectores con etiqueta 6	I6 28%	
			6 vectores con etiqueta 7	I7 24%	
ÁRBOLES DE DECISIÓN		2	2 vectores con etiqueta 1	I1 8%	0.96
			3 vectores con etiqueta 2	I2 12%	
			4 vectores con etiqueta 3	I3 16%	
			2 vectores con etiqueta 4	I4 8%	
			4 vectores con etiqueta 5	I5 16%	
			5 vectores con etiqueta 7	I7 20%	
			1 vector con etiqueta 8	I8 4%	
			3 vectores con etiqueta 9	I9 12%	
			1 vector con etiqueta 20	C6 4%	

### ***Test 2 (Consumidor tipo Institucional)***

En este caso, cada una de las técnicas analizadas ha sido testeada con curvas de carga de un consumidor tipo Institucional. En total fueron testeados 13 nuevos vectores. Ver Tabla 6

Tabla 6. Resultados, segundo conjunto de datos, test 2

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
PERCEPTRÓN MULTICAPA (MLP)	Gradiente Descendiente	243	6 vectores con etiqueta 6 7 vectores con etiqueta 16	I6 46.15% C3 53.85%	0
	Gradiente Descendiente con Momento	274	2 vectores con etiqueta 14 11 vectores con etiqueta 17	C1 15.38% C4 84.61%	0

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
	Ratio de Aprendizaje Adaptativo	166	2 vectores con etiqueta 6 1 vector con etiqueta 9 4 vectores con etiqueta 13 3 vectores con etiqueta 17 1 vector con etiqueta 19 2 vectores con etiqueta 25	I6 15.38% I9 7.7% T4 30.77% C4 23.1% T5 7.7% C11 15.38%	0.38
	Ratio de Aprendizaje Adaptativo + Momento	60	1 vector con etiqueta 10 2 vectores con etiqueta 12 9 vectores con etiqueta 13 1 vector con etiqueta 26	T1 7.7% T3 15.38% T4 69.23% R1 7.7%	0.92
	Resistente	3	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
	Fletcher-Reeves	10	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
	Polak-Ribière	5	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
	Powell-Beale	6	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
	Gradiente conjugado escalado	4	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
	Algoritmo secante de un paso	6	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
LVQ	76	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1	
SOM	0	1 vectores con etiqueta 19 4 vectores con etiqueta 10 3 vectores con etiqueta 11 5 vectores con etiqueta 13	T5 7.69 % T1 30.76 % T2 23.07 % T4 38.46 %	1	
ELMAN	7	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1	
RBF	0	10 vectores con etiqueta 14 2 vectores con etiqueta 16 1 vector con etiqueta 19	C1 77% C3 15.38% T5 7.7%	0.077	
GRNN	0	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1	

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
PNN		0	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
CLUSTERING SUSTRATIVO		0	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
ANÁLISIS DISCRIMINANTE		0	2 vectores con etiqueta 10 3 vectores con etiqueta 12 4 vectores con etiqueta 13 4 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 30.77% T5 30.77%	1
ÁRBOLES DE DECISIÓN		2	2 vectores con etiqueta 10 3 vectores con etiqueta 12 3 vectores con etiqueta 13 5 vectores con etiqueta 19	T1 15.38% T3 23.1% T4 23.1% T5 38.5%	1

### ***Test 3 (Consumidor tipo Comercial)***

En este caso, cada una de las técnicas analizadas ha sido testada con curvas de carga de un consumidor tipo comercial. En total fueron testados 29 nuevos vectores. Ver Tabla 7.

Tabla 7. Resultados, segundo conjunto de datos, test 3

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
PERCEPTRÓN MULTICAPA (MLP)	Gradiente Descendiente	243	20 vectores con etiqueta 6 5 vectores con etiqueta 16 1 vector con etiqueta 20 3 vectores con etiqueta 26	I6 68.96% C3 17.24% C6 3.45% R1 10.34%	0.69
	Gradiente Descendiente con Momento	274	3 vectores con etiqueta 8 1 vector con etiqueta 10 14 vectores con etiqueta 11 2 vectores con etiqueta 14 5 vectores con etiqueta 17 4 vectores con etiqueta 26	I8 10.34% T1 3.45% T2 48.28% C1 6.9% C4 17.24% R1 13.8%	0.24

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
	Ratio de Aprendizaje Adaptativo	166	1 vector con etiqueta 8 10 vectores con etiqueta 9 1 vectores con etiqueta 12 3 vectores con etiqueta 15 4 vectores con etiqueta 16 5 vectores con etiqueta 17 2 vectores con etiqueta 23 2 vectores con etiqueta 24 1 vector con etiqueta 25	I8 3.45% I9 34.48% T3 3.45% C2 10.34% C3 13.8% C4 17.24% C9 6.9% C10 6.9% C11 3.45%	0.55
	Ratio de Aprendizaje Adaptativo + Momento	60	1 vector con etiqueta 7 3 vectores con etiqueta 15 7 vectores con etiqueta 16 3 vectores con etiqueta 17 4 vectores con etiqueta 21 3 vectores con etiqueta 22 1 vector con etiqueta 23 6 vectores con etiqueta 24 1 vector con etiqueta 25	I7 3.45% C2 10.34% C3 24.14% C4 10.34% C7 13.8% C8 10.34% C9 3.45% C10 20.7% C11 3.45%	0.96
	Resistente	3	1 vector con etiqueta 3 7 vectores con etiqueta 14 2 vectores con etiqueta 15 4 vectores con etiqueta 16 2 vectores con etiqueta 17 1 vector con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	I3 3.45% C1 24.14% C2 6.9% C3 13.8% C4 6.9% C6 3.45% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	0.96
	Fletcher-Reeves	10	3 vectores con etiqueta 9 8 vectores con etiqueta 15 5 vectores con etiqueta 16 1 vector con etiqueta 19 4 vectores con etiqueta 21 4 vectores con etiqueta 22 3 vectores con etiqueta 24 1 vector con etiqueta 25	I9 10.34% C2 27.6% C3 17.24% T5 3.45% C7 13.8% C8 13.8% C10 10.34% C11 3.45%	0.86
	Polak-Ribiére	5	3 vectores con etiqueta 9 5 vectores con etiqueta 15 7 vectores con etiqueta 16 1 vector con etiqueta 17 1 vector con etiqueta 20 4 vectores con etiqueta 21 3 vectores con etiqueta 22 4 vectores con etiqueta 24 1 vector con etiqueta 25	I9 10.34% C2 17.24% C3 24.14% C4 3.45% C6 3.45% C7 13.8% C8 10.34% C10 13.8% C11 3.45%	0.9

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
	Powell-Beale	6	1 vector con etiqueta 9 5 vectores con etiqueta 15 5 vectores con etiqueta 16 3 vectores con etiqueta 17 2 vectores con etiqueta 18 1 vector con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	I9 3.45% C2 17.24% C3 17.24% C4 10.34% C5 6.9% C6 3.45% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	0.96
	Gradiente conjugado escalado	4	6 vectores con etiqueta 15 10 vectores con etiqueta 16 4 vectores con etiqueta 20 2 vectores con etiqueta 21 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	C2 20.7% C3 34.5% C6 13.8% C7 6.9% C9 6.9% C10 13.8% C11 3.45%	1
	Algoritmo secante de un paso	6	9 vectores con etiqueta 15 7 vectores con etiqueta 16 1 vector con etiqueta 20 3 vectores con etiqueta 21 3 vectores con etiqueta 22 1 vector con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	C2 31% C3 24.14% C6 3.45% C7 10.34% C8 10.34% C9 3.45% C10 13.8% C11 3.45%	1
LVQ		76	3 vectores con etiqueta 15 8 vectores con etiqueta 17 4 vectores con etiqueta 21 9 vectores con etiqueta 22 4 vectores con etiqueta 25 1 vector con etiqueta 26	C2 10.34% C4 27.6% C7 13.8% C8 31% C11 13.8% R1 3.45%	0.96
SOM		0	15 vectores con etiqueta 14 10 vectores con etiqueta 16 4 vectores con etiqueta 18	C1 51.7 % C3 34.48 % C5 13.79 %	1
ELMAN		7	3 vectores con etiqueta 9 5 vectores con etiqueta 14 3 vectores con etiqueta 15 3 vectores con etiqueta 17 2 vectores con etiqueta 18 1 vector con etiqueta 20 3 vectores con etiqueta 21 3 vectores con etiqueta 22 1 vector con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	I9 10.34% C1 17.24% C2 10.34% C4 10.34% C5 6.9% C6 3.45% C7 10.34% C8 10.34% C9 3.45% C10 13.8% C11 3.45%	0.9
RBF		0	20 vectores con etiqueta 14 5 vectores con etiqueta 16 4 vectores con etiqueta 18	C1 69% C3 17.24% C5 13.8%	1

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
GRNN		0	6 vectores con etiqueta 15 7 vectores con etiqueta 16 3 vectores con etiqueta 17 1 vector con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	C2 21% C3 24% C4 10.34% C6 3.45% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	1
PNN		0	6 vectores con etiqueta 15 7 vectores con etiqueta 16 3 vectores con etiqueta 17 1 vector con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	C2 20.7% C3 24.14% C4 10.34% C6 3.45% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	1
CLUSTERING SUCTRACTIVO		0	3 vectores con etiqueta 14 6 vectores con etiqueta 15 4 vectores con etiqueta 16 3 vectores con etiqueta 17 1 vector con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	C1 10.34% C2 20.7% C3 13.8% C4 10.34% C6 3.45% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	1
ANÁLISIS DISCRIMINANTE		0	3 vectores con etiqueta 15 10 vectores con etiqueta 16 3 vectores con etiqueta 17 1 vector con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	C2 10.34% C3 34.5% C4 10.34% C6 3.45% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	1
ÁRBOLES DE DECISIÓN		2	3 vectores con etiqueta 8 3 vectores con etiqueta 15 4 vectores con etiqueta 16 3 vectores con etiqueta 17 4 vectores con etiqueta 20 2 vectores con etiqueta 21 3 vectores con etiqueta 22 2 vectores con etiqueta 23 4 vectores con etiqueta 24 1 vector con etiqueta 25	I8 10.34% C2 10.34% C3 13.8% C4 10.34% C6 13.8% C7 6.9% C8 10.34% C9 6.9% C10 13.8% C11 3.45%	0.9

**Test 4 (Consumidor tipo Residencial)**

En este caso, cada una de las técnicas analizadas ha sido testeada con curvas de carga de un consumidor tipo residencial. En total fueron testeados 2. Ver Tabla 8

Tabla 8. Resultados, segundo conjunto de datos, test 4

Técnica Empleada		ERA	Etiquetas Identificadas	Tipo de Consumidor Identificado (%)	IA
PERCEPTRÓN MULTICAPA (MLP)	Gradiente Descendiente	243	2 vectores con etiqueta 6	I6 100%	0
	Gradiente Descendiente con Momento	274	2 vectores con etiqueta 17	C4 100%	0
	Ratio de Aprendizaje Adaptativo	166	2 vectores con etiqueta 26	R1 100%	1
	Ratio de Aprendizaje Adaptativo + Momento	60	2 vectores con etiqueta 26	R1 100%	1
	Resistente	3	2 vectores con etiqueta 26	R1 100%	1
	Fletcher-Reeves	10	2 vectores con etiqueta 26	R1 100%	1
	Polak-Ribière	5	2 vectores con etiqueta 26	R1 100%	1
	Powell-Beale	6	2 vectores con etiqueta 26	R1 100%	1
	Gradiente conjugado escalado	4	2 vectores con etiqueta 26	R1 100%	1
	Algoritmo secante de un paso	6	2 vectores con etiqueta 26	R1 100%	1
LVQ	76	2 vectores con etiqueta 26	R1 100%	1	
SOM	0	2 vectores con etiqueta 26	R1 100%	1	
ELMAN	7	2 vectores con etiqueta 26	R1 100%	1	
RBF	0	2 vectores con etiqueta 26	R1 100%	1	
GRNN	0	2 vectores con etiqueta 26	R1 100%	1	
PNN	0	2 vectores con etiqueta 26	R1 100%	1	
CLUSTERING SUSTRATIVO	0	2 vectores con etiqueta 26	R1 100%	1	
ANÁLISIS DISCRIMINANTE	0	2 vectores con etiqueta 26	R1 100%	1	
ÁRBOLES DE DECISIÓN	2	2 vectores con etiqueta 26	R1 100%	1	

Tras realizar todos los tests con el segundo conjunto de datos se comprueba que muchas de las técnicas utilizadas han clasificado satisfactoriamente los datos de entrada y reconocido el mapping subyacente de los datos para identificar consumidores no

empleados durante el diseño o entrenamiento. Entre estas técnicas cabe destacar las redes MLP con algoritmos Resistente, Fletcher-Reeves, Polak-Ribière, Powell-Beale, Gradiente Conjugado Escalado, las redes recurrentes de Elman, los mapas autoorganizados de Kohoone (SOM), las redes de base radial GRNN y PNN, la técnica borrosa de Clustering Sustractivo y las técnicas estadísticas de Análisis Discriminante y Árboles de Decisión.

## 4.6. CONCLUSIONES

De entre todas las técnicas analizadas, el autor de esta Tesis, decidió seguir la línea de investigación utilizando la red no supervisada denominada “Self Organizing Map”, conocidos también como mapas autoorganizados de kohonen, como herramienta de segmentación, agrupación y clasificación de consumidores eléctricos. La capacidad de estos mapas para extraer patrones de comportamiento y agrupar vectores de datos con perfiles similares, junto con los buenos resultados obtenidos en esta primera investigación tanto en la parte de segmentación (clasificación), como en la parte de identificación de nuevos consumidores, han conseguido que a partir de ahora la investigación se centre en estos mapas y su capacidad para ser utilizados en mayor profundidad en este campo.

No obstante, a continuación se muestra la Tabla 9 con un resumen de las mejores técnicas. En ella aparecen aquellas técnicas que han obtenido un valor de Error de aprendizaje igual a cero y un valor de índice de acierto igual a uno en los cuatro test realizados con vectores de entrada (curvas de carga) de nuevos consumidores.

Tabla 9. Resumen de las mejores técnicas analizadas

<b>Técnica</b>	<b>ERA</b>	<b>IA (Test consumidor Industrial)</b>	<b>IA (Test consumidor Institucional)</b>	<b>IA (Test consumidor Comercial)</b>	<b>IA (Test consumidor Residencial)</b>
SOM	0	1	1	1	1
GRNN	0	1	1	1	1
PNN	0	1	1	1	1
Clustering Substractivo	0	1	1	1	1
Análisis Discriminante	0	1	1	1	1





# CAPITULO 5

---

Filtrado de días anómalos y patrones de consumo mediante Self-Organizing Maps





## 5.1. INTRODUCCIÓN

Este capítulo recoge la investigación realizada con los Mapas Autoorganizados de Kohonen (SOMs) aplicados a la organización, filtrado, clasificación y extracción de patrones de la demanda eléctrica de distintos consumidores. Se han utilizado curvas de carga diaria de los consumidores analizados. Para simplificar el estudio han sido considerados dos usuarios medios: una Universidad pública Española y una Industria de tamaño medio perteneciente a la provincia de Alicante. Los resultados obtenidos tras la investigación ha demostrado claramente la capacidad de esta Técnica para mejorar la gestión de datos y facilitar la segmentación de consumidores eléctricos.

## 5.2. ACONDICIONAMIENTO Y PROCESAMIENTO DE DATOS

### 5.2.1. CONCEPTOS

El proceso de extracción de información desde una base de datos requiere un tratamiento previo de los datos para darles un formato uniforme que permita trabajar con ellos de forma adecuada. Los perfiles de carga diaria presentan picos y valores diferentes para cada consumidor, es necesario, por tanto, normalizar los datos para tener un formato común para todos los usuarios y facilitar su entrada a la red neuronal o a la técnica considerada de estudio para segmentar los niveles de demanda en una primera aproximación.

Los estudios previos que se han llevado a cabo a lo largo de esta investigación han permitido conocer el formato más adecuado para los datos, es decir, convertir las curvas de carga en 24 o en 96 valores diarios de consumo [113]. Este tratamiento se hace necesario para extraer las características más relevantes del usuario con el mínimo coste de tiempo y de trabajo, este formato es el que normalmente obtienen las compañías eléctricas al realizar sus mediciones. Finalmente, la comparación entre un gran número de

consumidores puede ser realizada a través de diversos índices y factores obtenidos a partir de estos perfiles.

### **5.2.2. APLICACIÓN DE LOS SOMS AL DATA MINING**

A partir del momento en que los datos tienen el mismo formato, se pueden aplicar varias técnicas para lograr un filtrado adecuado y la detección de valores puntuales o series no deseables. En este sentido los Mapas Autoorganizados de Kohonen [54] se presentan como una potente herramienta capaz de identificar la presencia de datos anómalos, o debido a la presencia de ruido o a la pérdida de información en las bases de datos, como pueden ser los perfiles correspondientes a días de consumo anormal, periodos de vacaciones, días con problemas de fiabilidad, etc.. Incluso muchas de estas redes, permiten encontrar patrones de comportamientos anómalos propios de cambios puntuales en el hábito de consumo (cargas temporales) y permitir así, la definición de modelos representativos de los consumidores a través de promedios de los perfiles típicos clasificados.

### **5.2.3. CASO DE ESTUDIO**

Los perfiles de carga diaria que componen el conjunto de mediciones para utilizar como datos de entrada para el entrenamiento y posterior análisis de la SOM, corresponden a 2 consumidores usuales de tamaño medio: una Universidad Española y una Industria de tamaño medio, estando ambos situados en España en la provincia de Alicante. Los datos de entrada a la red para el entrenamiento corresponden a curvas de carga de días de entre semana, de lunes a viernes, para ambos y medidos en el año 2003. Para simplificar el estudio no han sido consideradas las curvas de carga de los sábados y de los domingos. En la Tabla 10, se muestra la etiqueta asociada a los perfiles de carga diaria, un número de etiqueta para todos los perfiles de carga para un consumidor dado.

Estas etiquetas se utilizan también a la hora de estudiar la clasificación de los consumidores, realizada en la sección 5.3 de este capítulo ya que permiten identificar los datos asignados a cada neurona con el correspondiente consumidor. La Tabla 10, también muestra el número de curvas de carga consideradas para cada consumidor, incluyendo los días anómalos que serán eliminados posteriormente (tal como se recoge en la explicación del estudio). El número de curvas de carga es un factor importante para determinar, en cierta medida, el tamaño de la red.

Tabla 10. Espectro de Consumidores para el entrenamiento de la Red SOM

<b>Consumidor</b>	<b>Etiqueta</b>	<b>Número inicial de vectores de datos de entrada</b>	<b>Actividad del Consumidor</b>
Usuario 1	1	147	Industria de tamaño medio
Usuario 2	2	195	Universidad

Se ha utilizado también otra alternativa de etiquetado de los vectores de entrada a lo largo de la investigación para realizar adecuadamente la propuesta de filtrado de los datos anómalos. Con este otro etiquetado a cada perfil (curva de carga o vector) se le asigna un número siguiendo el siguiente criterio: los dos últimos dígitos indican el día del mes y los dos primeros dígitos indican el mes en cuestión al que pertenece dicha curva, es decir, mm/dd. Por tanto el mapa obtenido, ver Figura 56, permite identificar, las curvas de carga de cada día de la semana y mes asignados a cada neurona tras el entrenamiento.

Después se seleccionó como formato de los datos de entrada a la red que estuvieran en el Dominio del Tiempo (valores de potencia por cada periodo de tiempo, horas o  $\frac{1}{4}$  de hora). Las curvas que estaban medidas en intervalos de 15 minutos fueron recalculadas a 24 valores diarios de consumo de potencia activa. La razón de elegir este tratamiento fueron los buenos resultados obtenidos en las investigaciones llevadas a cabo anteriormente [113].

### 5.3. CLASIFICACIÓN DE PATRONES DE CONSUMO

Como se ha podido ver en el Capítulo 4 la aplicación de SOM permite la obtención de un mapa entrenado con datos de entrada de consumos eléctricos de diversos usuarios. Este mapa estará segmentado en distintas áreas clasificando consumidores con perfiles similares. Esta primera clasificación debería servir en un futuro para ser utilizada en otras tareas y aplicaciones, como la caracterización de patrones de consumo para la creación de sistemas de seguimiento de cargas [52], o la predicción de curvas de carga de periodos anómalos de consumo en el corto plazo [114] & [115].

A continuación se presenta el experimento de clasificación que se llevó a cabo utilizando los dos consumidores eléctricos presentados en el punto 5.2.3 de este capítulo. Los perfiles de carga típicos de estos dos consumidores son bien conocidos y muestran características que los hacen muy diferentes en sus comportamientos.

En primer lugar se realizó un entrenamiento utilizando toda la información disponible de cada consumidor sin aplicar antes ningún tipo de filtrado de días anómalos. Simplemente se realizó el proceso de acondicionamiento de datos de entrada descrito en los puntos anteriores, es decir, curvas de carga normalizadas y 24 valores de potencia por día. El mapa de etiquetas obtenido utilizando la máscara 1 para la Industria de tamaño medio y la máscara 2 para la Universidad puede verse en la Figura 54.

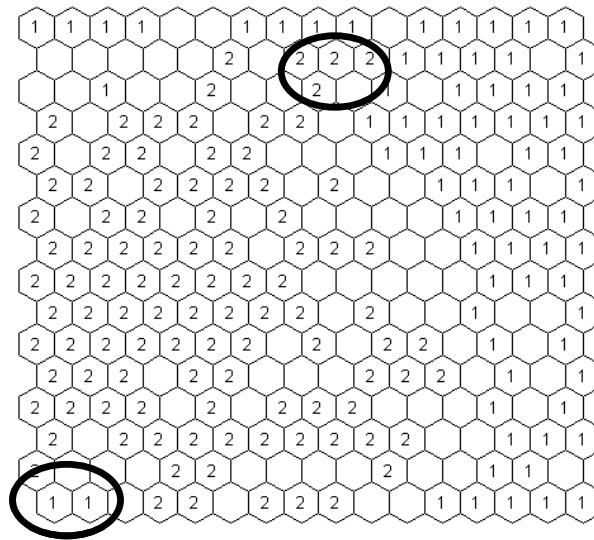


Figura 54. Clasificación de consumidores con datos de entrada sin filtrado

En este mapa se muestra como la red es capaz de agrupar perfiles de consumidores en la misma área del Mapa separando y diferenciando cada uno de ellos. No obstante, es posible observar cómo el mapa presenta zonas dónde la separación no está clara e incluso se mezclan etiquetas de ambos clientes. Este hecho sugiere que algunos perfiles que corresponden a diferentes consumidores no estén bien situados en la zona de perfiles de consumos usuales y por consiguiente la red entrenada no es capaz de distinguirlos correctamente, en consecuencia es necesario filtrar previamente los datos. Por esa razón se realizó un nuevo entrenamiento con datos de entrada donde se habían eliminado los perfiles de demanda anómalos.

## 5.4. APLICACIÓN DE SOM AL FILTRADO DE DATOS ANÓMALOS

A lo largo de las simulaciones que se han llevado a cabo han sido probadas varias configuraciones con varios parámetros de entrenamiento de la red. Finalmente el entrenamiento realizado con un mapa de tamaño 16 x 16 neuronas y un número de pasos de 2000 y de 1000 para el entrenamiento primario y el secundario (finetuning) respectivamente, demostró ser la mejor configuración.



Los resultados obtenidos con el entrenamiento de días de lunes a viernes del Usuario 2, son mostrados en el mapa de la Figura 55.

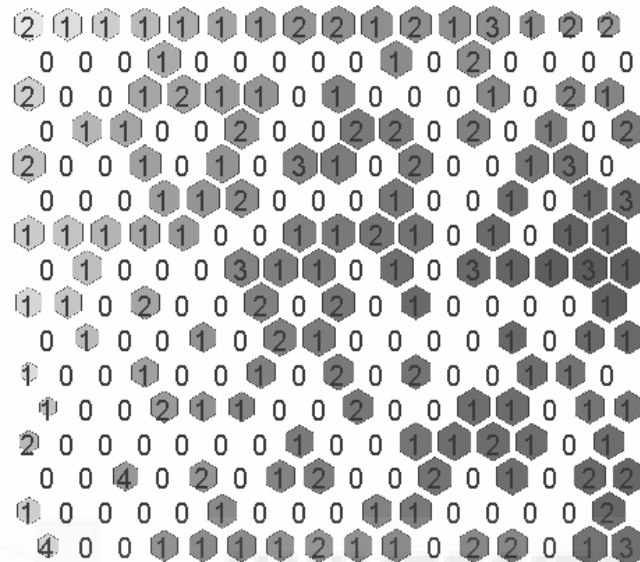


Figura 55. Mapa de Color entrenado con el Usuario tipo 2

En este mapa es posible observar que el número de perfiles representados por el SOM y asignados a la misma neurona (celda) están marcados con un dígito, este número representa los perfiles asignados a esta neurona al final del entrenamiento. Por otra parte el tamaño relativo de cada una de las neuronas indica el nivel de parecido de los perfiles en relación con el resto del conjunto de datos de entrada empleados en el entrenamiento. Además las zonas aparecen dibujadas con diferentes colores (escala de grises) dependiendo de la similitud de los datos y su distancia frente a otros clusters, permitiendo por tanto realizar una primera localización de los perfiles de carga anómalos o poco usuales, localizándolos en la esquina inferior izquierda del mapa.

No obstante, también es posible utilizar el “Mapa de etiquetas” obtenido tras el entrenamiento para identificar los días concretos en los que el mapa de color anterior nos ha fijado como días de interés, ver Figura 55 y Figura 56. Según el criterio que hemos adoptado de que cada etiqueta viene representada por 4 dígitos, los dos primeros son el mes y los otros dos el día de dicho mes. No obstante, en el mapa se pueden observar

fechas con tres dígitos en lugar de cuatro, esto es porque el programa de simulación descarta el primer cero, es decir, la fecha “101” que aparece en la parte inferior izquierda del mapa es el uno de Enero.

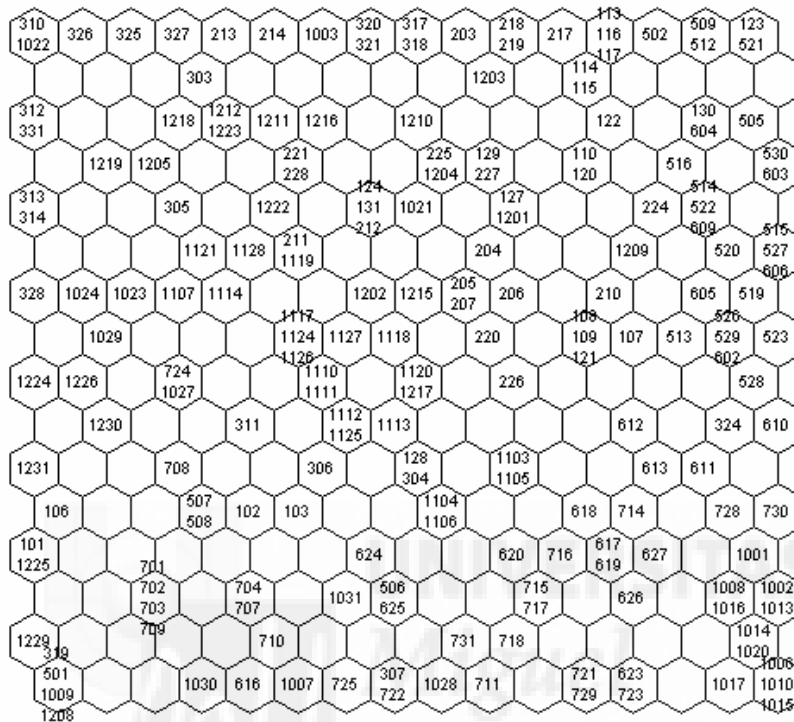


Figura 56. Mapa de etiquetas obtenido tras el entrenamiento (Universidad)

Analizando la Figura 56, las neuronas de la esquina inferior izquierda tienen las etiquetas “501” y “1208”, es decir el 1 de Mayo y el 8 de Diciembre, respectivamente. Ambos corresponden con dos días festivos en España. También aparece en la misma neurona la etiqueta “1009”. Obviamente las etiquetas “1231” y “1225” corresponden a días festivos de las Fiestas Navideñas.

Finalmente, tras ser entrenado el mapa, es posible mediante una función de Matlab basada en el algoritmo k-means [73], forzar al mapa a que busque o se divida en un número de clusters máximo predeterminado por el usuario. En este caso se pidió al mapa que encontrará una división no superior a 2 clusters. Esto es posible mediante la utilización del índice de Davies-Bouldin [108] el cual permite calcular distancias entre clusters y encontrar en un mapa SOM entrenado el número óptimo de divisiones (cuando

el usuario fija un número máximo de clusters). El índice de Davis Bouldin es una función que representa el cociente entre la suma de la distancia promedio de los objetos de un cluster y la distancia o separación entre los centros de los clusters:

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{i \neq j} \left\{ \frac{S_n(Q_i) + S_n(Q_j)}{S(Q_i, Q_j)} \right\}$$

Donde  $n$  es el número de clusters,  $S_n$  la distancia promedio de todos los objetos del cluster hasta su centro,  $S(Q_i, Q_j)$  es la distancia entre los centros de los clusters. Se consiguen valores pequeños de este índice cuando los clusters son compactos y están lejos los unos de los otros. Consecuentemente valores muy bajos de este índice significa que ha conseguido una buena clusterización. El resultado obtenido se puede ver en la Figura 57.

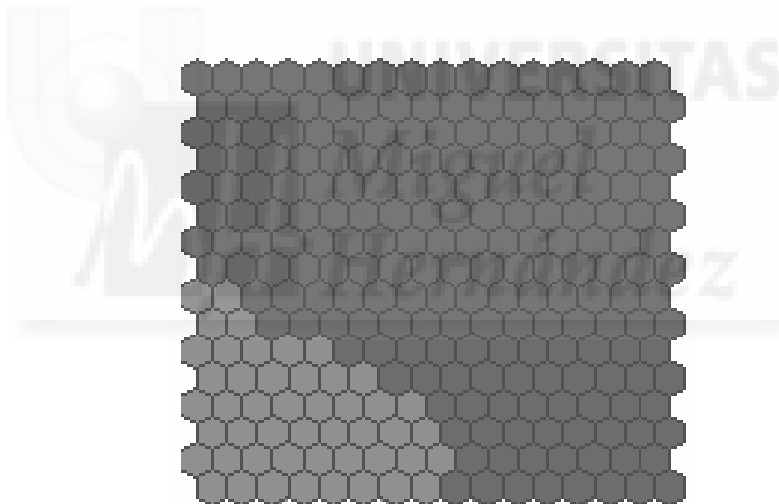


Figura 57. Mapa de clusters de la Universidad con límite máximo de búsqueda de clusters 2

A modo de una primera conclusión en la Figura 57 se puede observar como el Mapa de Clusters ha dividido en dos zonas los datos considerados como diferentes. Comparando el área inferior izquierda con el mapa de etiquetas de la Figura 56 los perfiles asignados al área de menor tamaño corresponden a consumos en periodos de vacaciones, festivos y de perfil diferente al resto de días considerados como perfiles típicos del consumidor Universidad y situados en la parte superior derecha del mapa de

clusters. Tan solo faltaría aumentar el número máximo de búsqueda de clusters para conocer hasta que punto es posible dividir en el mapa con patrones diferentes.

A continuación se pedirá al mapa entrenado que se divida como máximo en un número de 10 clusters. La Figura 58 muestra el *Mapa de Clusters* obtenido.

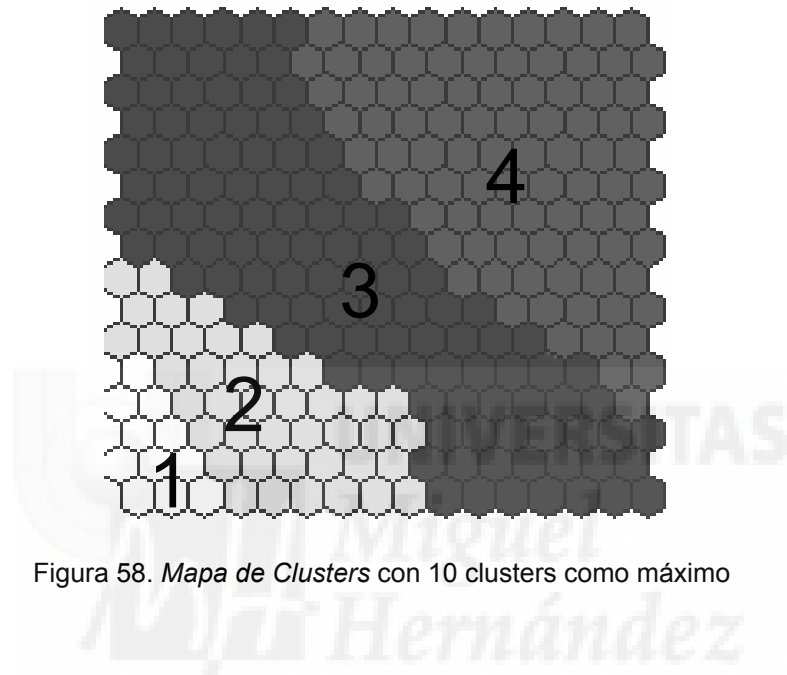


Figura 58. *Mapa de Clusters* con 10 clusters como máximo

En este caso se obtienen 4 zonas claramente diferenciadas. Comparando este mapa con el mapa de etiquetas es posible observar como la red puede diferenciar tres clases de perfiles diferentes.

Los perfiles correspondientes a consumos típicos del usuario Universidad aparecen representados en la Figura 59. Estos están asignados a las regiones 3 y 4 de la Figura 58. Los perfiles de consumo correspondientes a días anómalos, como son los días festivos o periodos vacacionales están representados en la Figura 60. En este caso estos perfiles están localizados en la región 1 de la Figura 58. Los perfiles de consumo de días con un comportamiento algo diferente al perfil típico, ver Figura 61, se encuentran en el área 2 de la Figura 58.

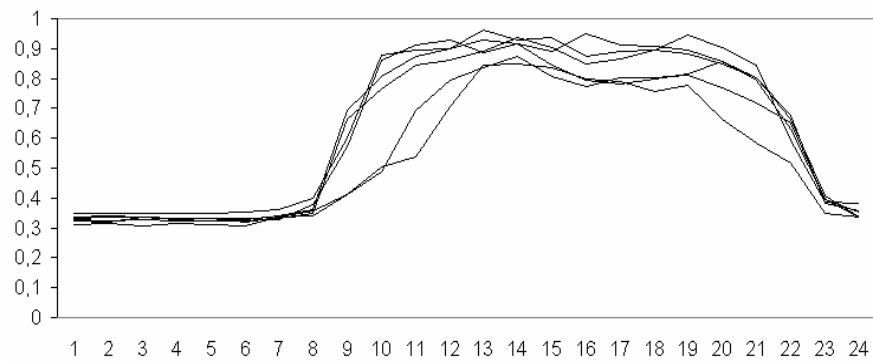


Figura 59. Curvas de carga de los días de consumo normal.

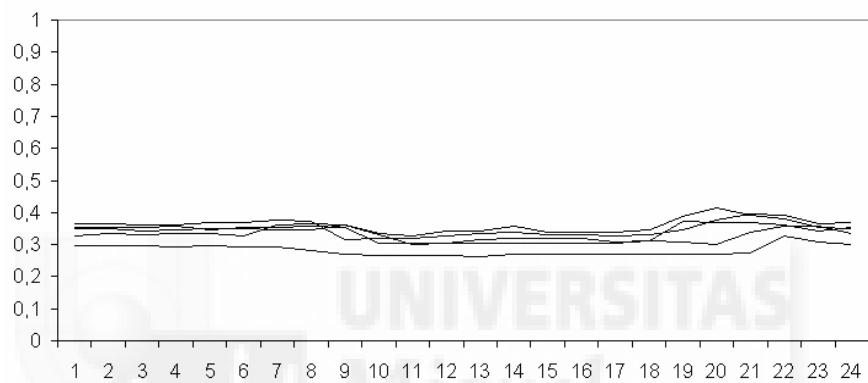


Figura 60. Curvas de carga de los días anómalos.

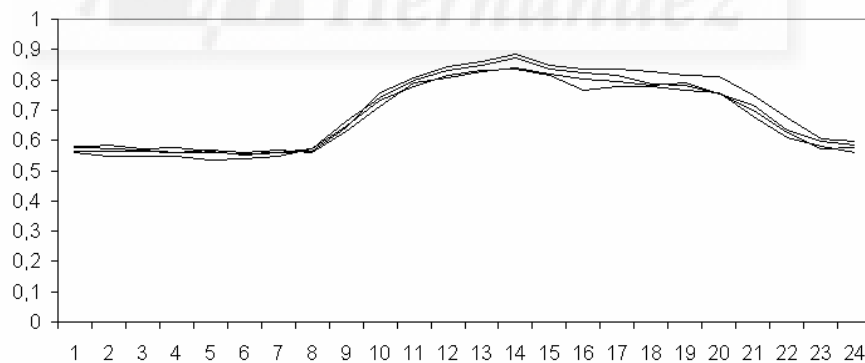


Figura 61. Curvas de carga de los días de consumo algo diferente.

Los resultados previos muestran como la red no es capaz por ella misma de identificar automáticamente los días anómalos con la suficiente certeza como para eliminarlos de las bases de datos, pero presenta otras capacidades muy útiles. Una de estas capacidades consiste en aislar aquellos perfiles de consumo que contengan errores de medición

causados por un fallo en los aparatos de medida utilizados y perfiles anómalos, facilitando a las compañías la gestión o mantenimiento de los mismos, para eliminar los datos erróneos.

Otra capacidad de aplicar los mapas SOM es que permite un estudio particular del comportamiento de un consumidor. Los perfiles de la demanda del consumidor Universidad identificados en la región 2, ver Figura 58, corresponden a días de Julio en los cuales los niveles de consumo durante el periodo nocturno crecen debido al continuo uso de los equipos de Aire Acondicionado. También se puede observar que durante el periodo diurno el pico de la curva de carga queda algo reducido debido a que el mes Julio coincide con el periodo vacacional de los alumnos.

Esta técnica de Inteligencia Artificial se presenta como una herramienta muy útil para el estudio y estimación del uso final del consumo, en este caso aire acondicionado, y permitirá realizar operaciones de gestión como el control de acciones, introducción de nuevas tecnologías y medidas de eficiencia energética.

## **5.5. RESULTADOS TRAS APLICAR EL FILTRADO**

Tras aplicar el filtrado de datos anómalos (descrito en el punto anterior) separadamente a ambos clientes, se ha realizado un nuevo entrenamiento con los datos de entrada ya filtrados. Los resultados son mostrados en la Figura 62. La red puede ahora clasificar correctamente y además separar con gran precisión los datos que corresponden a ambos clientes, tan solo hay que comparar el mapa obtenido (mapa izquierdo) con los datos en bruto con el mapa obtenido con los datos filtrados (mapa derecho). El segundo mapa no mezcla etiquetas de un consumidor en la zona o cluster del otro consumidor.

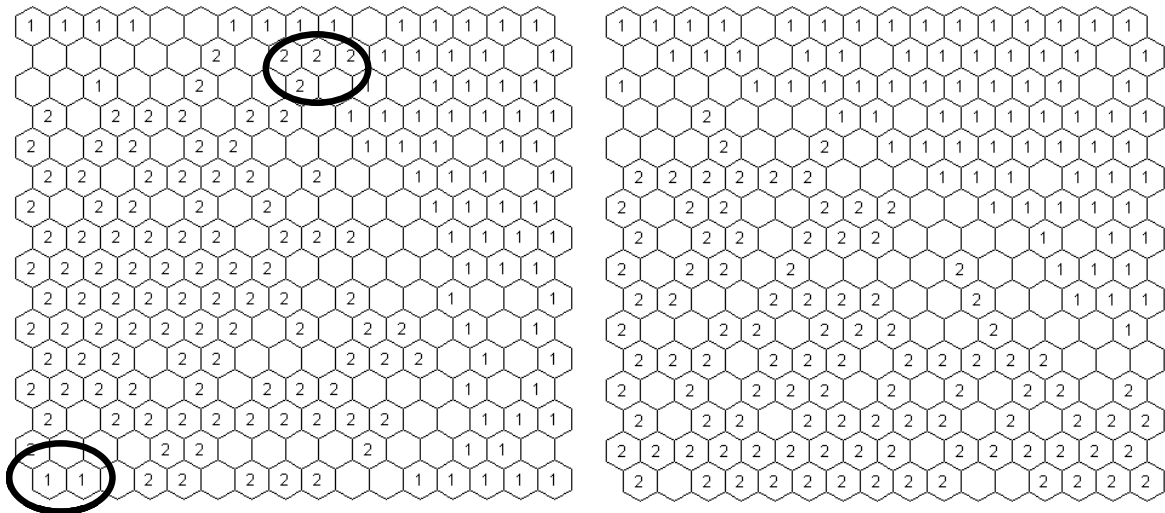


Figura 62. Clasificación de consumidores con datos de entrada sin filtrar y con datos de entrada filtrados.

Los perfiles usuales o típicos de los consumidores clasificados por el Self-Organizing Map están representados en la Figura 63. Es necesario por tanto un filtrado previo de los datos de entrada debido a las similitudes entre algunos perfiles de carga anómalos (festivos, periodos de vacaciones, etc...) de la Industria y de la Universidad.

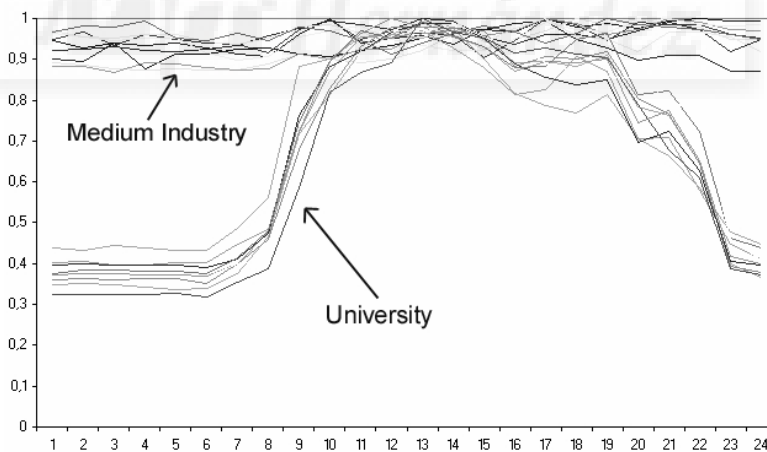


Figura 63. Perfiles de Carga de la Universidad y de la Industria

La mejora obtenida en el segundo entrenamiento, Figura 62, queda patente con la reducción obtenida en el error de cuantización, “ $Q_e$ ” (Quantization error) [112], ver Tabla 11. Este error da información de la exactitud con que el mapa representa las

muestras de datos, mide la resolución del mapa. Por otro lado el error topográfico, “ $T_e$ ” [112] mide el porcentaje de perfiles que no han conseguido obtener la primera y segunda neurona ganadora muy cercanas entre sí, mide la distorsión del mapa. Como se puede observar en la tabla el valor del Error Topográfico no es tan relevante como el valor del Error de Cuantización. Valores pequeños indican que la red no duda en el momento de asignar los perfiles a las neuronas.

El Error de Cuantización es calculado a partir de los datos de entrada mediante la siguiente fórmula:

$$Q_e = \frac{1}{N} \sum_{i=1}^N \|x_i - m_c(x_i)\|$$

Donde “ $N$ ” es el número de vectores que forman el conjunto de entrada “ $\mathbf{x}_i$ ” es cada vector de entrada, “ $m$ ” es el vector de peso y “ $c$ ” indica la neurona o celda mejor casada o asignada (The Best-Matching Unit) BMU para “ $\mathbf{x}$ ”. Tras el entrenamiento, se realiza un proceso de búsqueda de la BMU para cada vector de entrada, devolviendo el valor promedio de los respectivos valores de los errores de cuantización obtenidos para cada vector. Este valor promedio es el “ $Q_e$ ”.

El Error Topográfico es calculado mediante el siguiente algoritmo:

$$T_e = \frac{1}{N} \sum_{k=1}^N u(x_k)$$

Donde  $u(\mathbf{x}_k)$  es 1 si la primera y segunda BMU consecutivas de  $\mathbf{x}_k$  no se encuentran próximas la una de la otra, de otra forma  $u(\mathbf{x}_k)$  es 0.

Tabla 11. Errores de Cuantización y Topográfico

Tipo de entrenamiento	$Q_e$	$T_e$	$Q_e$ Reducción Relativa (%)
Con días anómalos (sin filtrar)	0.1520	0.0114	---
Sin días anómalos (con un tratamiento previo)	0.1303	0.0150	14,3



Por consiguiente se puede concluir que a través del proceso de filtrado, los resultados de la aplicación de SOM a un mayor número de clientes puede llevar a una clasificación excelente con una gran utilidad.

## **5.6. CONCLUSIONES**

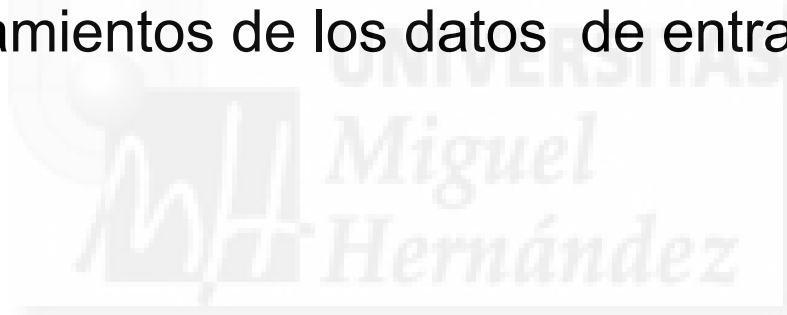
En este capítulo se ha analizado una fase previa de la aplicación de los Mapas Autoorganizados de Kohonen a la gestión y operación de los sistemas de distribución de energía eléctrica para lograr la segmentación de dos consumidores eléctricos típicos y sus modelos de uso de la energía en base a sus curvas de carga diarias. En el caso, bastante habitual donde exista la posibilidad de presencia de datos anómalos y de incertidumbre, como es el caso de la utilización de grandes bases de datos, este tipo de red neuronal proporciona la detección de datos erróneos, pérdidas de información y además posibles desplazamientos de la tendencia normal debido a cambios en los precios, o a factores externos como es el caso del aumento de la temperatura ambiente, que pueden distorsionar la interpretación de resultados posteriores en el proceso de segmentación y clasificación.

El método presentado aquí puede ayudar a comercializadores y distribuidores a la hora de clasificar y segmentar consumidores eléctricos. Este sería el primer paso para evaluar la relación coste-efectividad de muchas políticas necesarias desde el punto de vista del lado de la demanda, por ejemplo el potencial de alternativas de energías más eficientes, la respuesta del consumidor a los precios del mercado o a tarifas, la diversificación de estas últimas, el éxito de aplicaciones de almacenamiento de la energía o las posibilidades de la generación distribuida en usuarios pequeños y medianos.

# **CAPITULO 6**

---

Comportamiento de la SOM con diferentes  
tratamientos de los datos de entrada





## 6.1. CLASIFICACIÓN DE PATRONES DE CONSUMO

### 6.1.1. INTRODUCCIÓN

En los anteriores capítulos se ha demostrado la capacidad de los mapas SOM como herramienta de agrupación, clasificación e identificación de consumidores eléctricos. En este capítulo se recogen los resultados de la investigación que se ha llevado a cabo con la finalidad de mejorar la clasificación de los mapas tras su entrenamiento, a partir del análisis de la influencia del formato del conjunto de datos de entrada a la red para su entrenamiento.

Cada conjunto de datos está formado por vectores que recogen el comportamiento de la curva de carga asociada a un usuario. Este conjunto debería contener la información necesaria para que la red fuera capaz de asignar y clasificar cada vector en un cluster, en función del comportamiento elemental de cada consumidor. Tradicionalmente la clasificación de consumidores eléctricos se ha basado en el tipo de actividad económica declarada o identificada para cada usuario (por ejemplo, a través de los códigos NACE [116] y a partir de los niveles de tensión. Pero esta clasificación ha demostrado no ser lo suficientemente eficiente, debido principalmente a que varios patrones se pueden encontrar asociados dentro de una misma actividad económica, o usuarios con diferentes actividades económicas mostrar similares patrones de demanda o consumo).

Desde el punto de vista técnico y de investigación, dentro del campo de la respuesta de la demanda y la generación distribuida, es necesario encontrar características similares en función de las curvas de carga, que a su vez pueden ser obtenidas a partir de mediciones llevadas a cabo por los mismos consumidores o por los comercializadores para gestionar y reducir costes asociados a los consumos. En nuestro país es usual realizar estas mediciones mediante dispositivos que toman la potencia cada hora o cada quince minutos.

El conjunto de datos de entrada utilizado para los entrenamientos (curvas de carga diaria) [12], [107] ha sido el que ha demostrado hasta ahora los mejores resultados en cuanto a clasificación e identificación con redes neuronales. Sin embargo, hasta este

capítulo se ha investigado la clasificación con curvas de carga correspondientes a días de entre semana (días laborables) pero no se ha tenido en cuenta las curvas de carga correspondientes a los sábados y los domingos. Esta alternativa de no incluir los fines de semana, puede no tener mucho sentido cuando el objetivo es estudiar y desarrollar posibles mejoras de tarifas y almacenamientos de energía térmica [107], donde las curvas de carga de los sábados y domingos pueden ser importantes para conseguir este objetivo, debido principalmente a que muchas industrias, hospitales y hoteles trabajan los fines de semana. El llevar a cabo la investigación justificando la inclusión de los fines de semana en el espectro de datos de entrada a la red puede llevar a que el tamaño del conjunto de datos sea muy grande ( $1 \times N$ , donde  $N$  es 7 días x 96 muestras / día) con la consiguiente dificultad del mapa SOM de entrenar un mayor número de datos y la ralentización de los entrenamientos.

En el siguiente punto de este capítulo se explica el procedimiento que se ha llevado a cabo para incluir esta nueva información a los mapas SOM, pero reduciendo el tamaño de los vectores de entrada.

### 6.1.2. ESPECTRO DE DATOS

El conjunto de datos de entrada a la red que se ha utilizado está formado por curvas de carga de consumidores ya utilizados en las investigaciones anteriores recogidas en los capítulos 4 y 5. En este caso, el espectro está formado por 23 consumidores pertenecientes a diversos sectores del área mediterránea: industrial, comercial, residencial e institucional. El rango de picos de potencia anuales va desde los 100 kW a los 10MW. En la Tabla 12 se muestra las etiquetas asociadas a las curvas de carga de cada uno de los consumidores.

Tabla 12. Espectro de consumidores

Tipo de consumidor eléctrico	Etiqueta asociada
Industrial de Alimentación A	1
Industria de Alimentación B	2
Almacén Industrial	3
Industria Papelera C	4
Industria Papelera D	5
Industria Plástica E	6
Industria Plástica F	7
Industria Química G	8
Industria Química H	9
Universidad de tamaño grande 1	10
Universidad de tamaño grande 2	11
Universidad de tamaño medio 1	12
Universidad de tamaño medio 2	13
Hotel pequeño	14
Hotel mediano	15
Hotel grande	16
Hospital y centro médico 1	17
Hospital y centro médico 2	18
Hospital y centro médico 3	19
Hospital y centro médico 4	20
Comercio minorista 1	21
Comercio minorista 2	22
C.T. Residencial	23

### 6.1.3. TRANSFORMACIÓN DE LOS DATOS DE LA DEMANDA DEL DOMINIO DEL TIEMPO AL DOMINIO DE LA FRECUENCIA.

El objetivo fue utilizar la mayor cantidad de información posible, pero comprimiendo, filtrando y simplificando esta información (los fines de semana). Parece bastante lógico que un simple “array” (vector) de datos de entrada debería incluir toda la información relevante sobre el comportamiento de la demanda de un consumidor, pero también se debería mejorar la proyección topológica de los mapas SOM. Es decir, en esta investigación se intentó una doble transformación desde el dominio del tiempo a un

dominio de salida de la SOM, a través de una transformada de frecuencia, y por tanto obtener algunas mejoras en la clusterización de consumidores. Se han investigado varios métodos para comprimir y transformar la información desde el dominio del tiempo a una señal en el dominio de la frecuencia. Esta metodología es bien conocida para resolver otros problemas típicos de sistemas de potencia eléctrica, tales como la previsión de curvas de carga. Por ejemplo en [117], el análisis con las series de Fourier se aplica al filtrado de datos de carga antes de aplicar un modelo ARIMA. En [118] se propone utilizar la transformada de Wavelet para obtener un modelo de carga a corto plazo. Esta última transformada ha sido ampliamente utilizada para extraer patrones de comportamiento anómalos en los análisis de transitorios de los sistemas de potencia eléctrica [119]. Partiendo del conocimiento bibliográfico de la utilización en multitud de campos y aplicaciones de la transformada de Wavelet, en esta investigación se ha seleccionado la Transformada Rápida de Fourier (FFT, Fast Fourier Transform) como la transformada más interesante y sencilla a efectos de extraer comportamientos estacionarios de la demanda a partir de los perfiles de carga medidos.

A continuación se explican los conceptos básicos de la Transformada de Fourier y su extensión de la FFT.

### Transformada de Fourier

La Transformada de Fourier [120] es una transformada integral que expresa una función en términos de funciones básicas senoidales, es decir, como una suma o integral de funciones senoidales multiplicadas por algunos coeficientes (“amplitudes”). En ingeniería la transformada de Fourier se utiliza para pasar al "dominio de la frecuencia" y obtener información que no es evidente en el dominio del tiempo.

### Transformada discreta de Fourier (DFT)

En matemáticas la DFT [121], denominada también transformada finita de Fourier, es la más utilizada en procesos de tratamiento de señal y puede resolver ecuaciones diferenciales parciales y otras operaciones.

La secuencia de “n” números complejos  $x_0, \dots, x_{n-1}$  es transformada en la secuencia de “n” números complejos  $f_0, \dots, f_{n-1}$  de acuerdo a la fórmula:

$$f_j = \sum_{k=0}^{n-1} x_k e^{-\frac{2\pi i}{n} jk} \quad j = 0, \dots, n-1$$

donde “e” es la base del logaritmo natural, “i” es la unidad imaginaria ( $i^2 = -1$ ), y  $\pi$  es Pi. La transformada, en ocasiones, esta representada por el símbolo F, como en  $f = F(x)$  o  $F_x$ .

### Transformada Rápida de Fourier (FFT)

La FFT (abreviatura en inglés Fast Fourier Transform) [122] & [123] es un algoritmo que permite calcular la Transformada de Fourier Discreta de una onda y encontrar sus armónicos más representativos. Esto significa que es posible encontrar si un comportamiento concreto aparece en un periodo de tiempo considerado. La FFT es de gran utilidad en una amplia variedad de aplicaciones, desde la resolución de ecuaciones diferenciales parciales hasta el tratamiento y filtrado digital de señales.

El objetivo de aplicar esta metodología es encontrar el comportamiento de la carga en un día o en una semana, incluyendo los fines de semana. Por esta razón las curvas de carga diaria (vectores) se trataron de la siguiente forma, ver figura Figura 64:



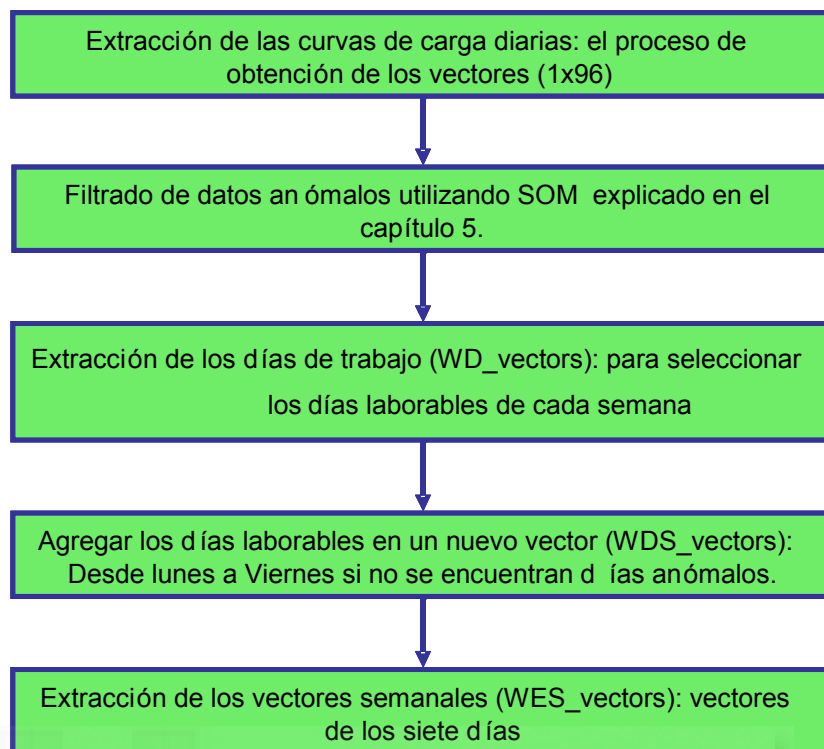


Figura 64. Pasos a seguir para la extracción de vectores semanales

Aplicando la función FFT a cada vector del dominio del tiempo (WD, WDS y WES), y las correspondientes ecuaciones para obtener los coeficientes de las series de Fourier (forma sinusoidal), se obtuvieron los siguientes vectores (F indica vector transformado):

- FWD\_Vector: valor medio ( $n=0$ ) y los términos de los armónicos de seno y coseno ( $n=1, 2, \dots$ ). De esta forma el armónico fundamental ( $n=1$ ) muestra términos de seno y coseno con periodo  $T=1$  día (Frecuencia =  $1/\text{día}$ )
- FWDS-vector: valor medio ( $n=0$ ) y los términos de armónico de seno y coseno ( $n=1, 2, \dots$ ). El primer término ( $n=0$ ) es la demanda promedio en los días laborables y el fundamental ( $n=1$ ) muestra los términos de seno y coseno con el periodo  $T=5$  días.
- FWES\_vector: valor medio ( $n=0$ , valor promedio de la demanda en una semana), el término fundamental ( $n=1$ ) muestra los términos de seno y coseno con el periodo  $T=7$  días (frecuencia= $1/\text{semana}$ )

De esta forma, cada vector de entrada presentado al mapa SOM tendrá una mezcla de tratamientos de frecuencia los cuales contienen información extraída en el dominio de la frecuencia a partir de algunos términos representativos como son los vectores FWD, FWDS y FWES. Después de realizar varias simulaciones y entrenamientos de los mapas con diferentes selecciones de armónicos de importancia, se eligió finalmente la siguiente combinación como vector de entrada (FD-vector) en el dominio de la frecuencia:

$$\text{FD\_vector} = \{\text{FWD-vector}(i); i=0 \dots 12 \ \& \ \text{FWDS\_vector}(j); j=5, 10 \ \& \ \text{FWES-vector}(k); k=0, 1, 7, 14\}$$

Por consiguiente, las curvas de carga diaria fueron transformadas en:

El valor medio y los 12 primeros armónicos del seno y coseno diarios (1+24 términos), 2 términos del patrón de carga diaria filtrado en los días laborables del vector FWDS ( $n=5$  y  $n=10$ ; 4 términos) y algunos términos del vector FWES (valor medio de la demanda, y  $n=1, 7$  y  $14$  términos del seno y coseno, es decir 7 términos) para forzar a la red SOM a tener en cuenta el comportamiento de la carga de los fines de semana. Obviamente los 11 últimos términos del FD\_array son los mismos para cada día de una misma semana.

Por ejemplo en la Figura 65, se muestra el perfil de la demanda a lo largo de una semana y su transformada semanal (FWES) para dos consumidores eléctricos distintos (etiquetas 2 y 14) y en la Figura 65.b recoge los términos armónicos del FWES. El consumidor 14 muestra un comportamiento de la demanda durante los fines de semana similar al comportamiento del resto de los días laborables, por lo tanto el array de armónicos FWES muestra un término  $n=1$  muy bajo. En la figura se puede apreciar el valor del armónico  $n=14$ , debido a la fluctuación de la demanda diaria (2 picos por día). Por otro lado, el consumidor con etiqueta 2 muestra una reducción de su curva de carga en los días del fin de semana, y los términos  $n=1$  y  $n=2$  son bastante altos. El término  $n=0$  no se ha representado en la figura por simplificar la representación de la misma.

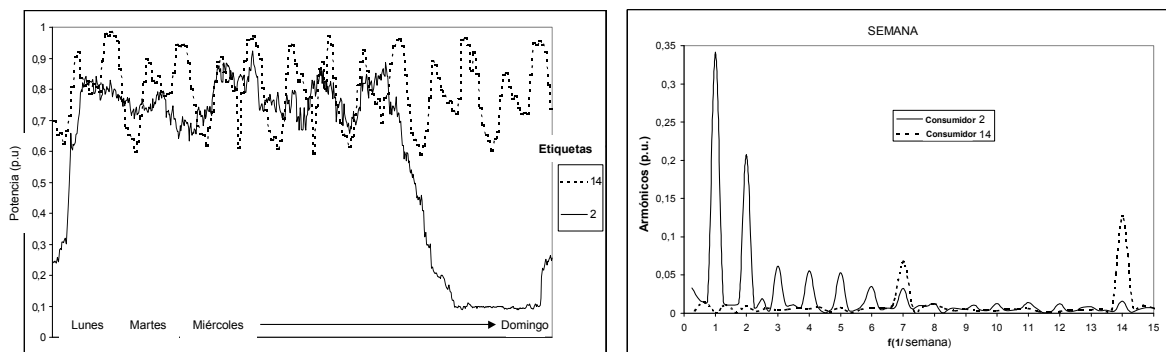


Figura 65a&b. Perfiles de carga semanal y vector FWES para los consumidores etiquetados como 2 y 14

## 6.2. ENTRENAMIENTOS CON SOM

Se han elegido varias opciones para alimentar los mapas SOM con la finalidad comprobar la utilidad y eficacia de la aplicación de la Transformada de Fourier, en concreto:

- Entrenamiento del mapa con datos en el dominio del tiempo. En este caso el espectro de datos de entrada a la red está formado por un conjunto de  $1 \times 96$  vectores de datos los cuales corresponden a perfiles de carga de días laborables, ver investigación presentada en [12].
- Entrenamiento del mapa con datos en el dominio de la frecuencia. En este caso los arrays de entrada son un conjunto de  $1 \times 36$  vectores de datos, tal cual se ha expuesto en el punto 6.1.2 (FD\_vector).

Para lograr una correcta proyección partiendo del conjunto de datos de entrada (374 días laborables filtrados) en un espacio de salida de la SOM coherente, se ha optado por utilizar un espacio de  $N \times N$ . La elección del valor de  $N$  es algo subjetivo; algunos autores sugieren el uso de un número de neuronas del mapa SOM algo más bajo que el número de vectores de entrada [12]. En este caso y para conseguir una mejor definición visual de los clusters del mapa entrenado, se ha elegido un mapa de tamaño  $20 \times 20$  neuronas (400 neuronas), que es un tamaño algo más grande que el número de muestras ( $374 > 19 \times 19$ ).

Para cada uno de los tratamientos de los datos de entrada fueron probados diferentes parámetros de entrenamiento. En las sucesivas simulaciones de prueba efectuadas con los mapas se utilizaron diferentes algoritmos de inicialización de los datos en la red, lineal o aleatorio, diferentes algoritmos de entrenamiento, secuencial o por lotes [112], así como varias combinaciones de los anteriores y diferentes números de pasos (secuencias) de entrenamientos como pueden ser 5000 y 3000 pasos para el entrenamiento primario y secundario, respectivamente.

Para una mejor comprensión de los mapas SOM es importante aclarar que para diferentes sesiones de entrenamientos se producen normalmente diferentes visualizaciones de los mapas entrenados incluso con el mismo conjunto de datos de entrada. También hay que considerar que el tiempo de entrenamiento de cada mapa SOM es bastante diferente debido principalmente a la complejidad asociada a cada conjunto de datos de entrada. En los entrenamientos llevados a cabo en esta investigación los tiempos variaban entre 20 minutos para los datos en el dominio del tiempo y 13 minutos para los datos en el dominio de la frecuencia, utilizando un ordenador Pentium IV con 512 Mb de memoria.

Finalmente, la selección del número de clusters en los mapas entrenados es otra de las tareas importantes de cara a los resultados. Este número, es un valor subjetivo, pero de forma bastante razonable debería tener un valor comprendido entre el número de macro-clusters según el caso de estudio de tipos de consumidores considerados (por ejemplo: residencial, comercial, industrial) y número total de consumidores incluidos en dicho espectro de datos de entrada. En este sentido, el número de clusters debería permitir una agrupación de consumidores promedia de más de un consumidor por cluster.

Es posible encontrar una división automática del mapa entrenado en un número óptimo de clusters, después de aplicar la función k-means [73]. Tan solo hay que indicar a esta función que el número máximo de clusters a encontrar no ha de ser superior a 12 (23 consumidores son los que forman el espectro de estudio, por tanto 12 se ha considerado un valor razonable). Es posible garantizar un óptimo clusterizado o agrupación de consumidores con la minimización del índice de Davies-Bouldin [108],

pero en algunos casos la inspección visual también nos puede ayudar a decidir y elegir cual es la mejor clusterización [124]. Como ya se ha comentado anteriormente valores pequeños del índice de Davies-Bouldin (DB) corresponden a un buen resultado de clusterización, es decir los clusters son compactos y sus centros están bastante separados. De esta forma se ha elegido el mapa que obtiene una configuración de clusters con el índice DB más bajo. En los siguientes puntos se detallan los resultados obtenidos.

### 6.2.1. ENTRENAMIENTOS CON DATOS DE ENTRADA EN EL DOMINIO DEL TIEMPO.

En este caso el espectro de datos de entrada está formado por los 374 perfiles de carga diaria medidos cada 15 minutos. La Figura 66 muestra los resultados obtenidos tras el entrenamiento. El mapa entrenado visualiza la agrupación de las etiquetas y de los clusters encontrados después de aplicar la función k-means [73]. En este caso el índice de Davies-Bouldin ha alcanzado un valor mínimo de 0,82, para un número de 8 clusters. En la **Tabla 13** se muestra un resumen de los principales clusters obtenidos.

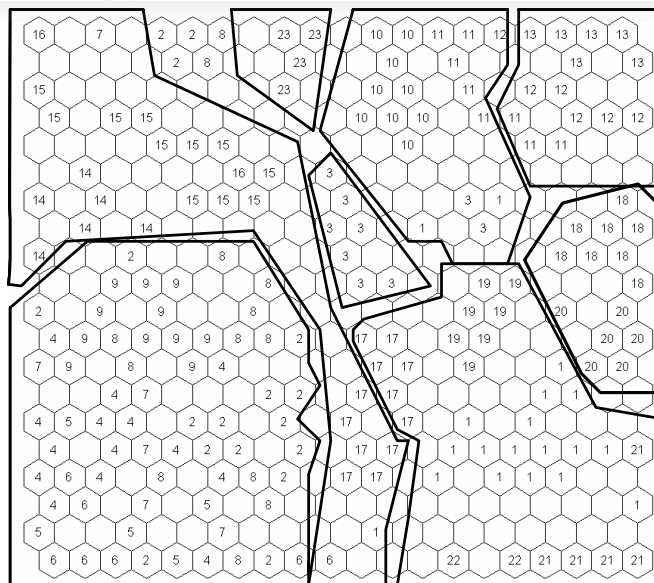


Figura 66. Mapa SOM entrenado con valores en el dominio del Tiempo

Tabla 13. Clusters obtenidos según consumidores

CLUSTER	ETIQUETAS
Universidades de tamaño grande	10 y 11
Universidades de tamaño medio	12 y 13
C.T. residencial	23
Hoteles	14, 15 y 16
Industrias	2, 4, 5, 6, 7, 8 y 9
Hospitales y centros médicos	18 y 20

De los resultados obtenidos se desprenden varias conclusiones. En primer lugar el proceso de agrupación es bastante bueno, y solamente 2 etiquetas aparecen en el mapa como clusters individuales (3 y 23). Las Universidades están divididas en 2 clusters y los principales consumidores industriales están agrupados en una gran área (exceptuando el consumidor con etiqueta 1 y 3). Pero el mapa también tiene algunos inconvenientes: los consumidores 11 y 17 son clasificados en dos clusters diferentes, y por otro lado algunas celdas (etiquetas 2 y 8 en la parte superior izquierda del mapa) no han quedado asignadas a ningún cluster específico.

### 6.2.2. ENTRENAMIENTOS CON DATOS DE ENTRADA EN EL DOMINIO DE LA FRECUENCIA.

Los vectores “FD” con valores de armónicos diarios y de fines de semana han sido utilizados para entrenar un mapa SOM con un tamaño de 20x20 neuronas. La agrupación de las etiquetas de los distintos consumidores y los clusters encontrados después de aplicar la función k-means [73] quedan representados en la Figura 67.a. En este caso el índice de Davies-Bouldin ha alcanzado un valor mínimo de 1.19 para un número de siete clusters.

Una de las principales características de este entrenamiento es la obtención de un mapa entrenado mucho más claro. Las etiquetas están más próximas y no aparecen errores en los clusters durante el proceso de localización y emplazamiento de las etiquetas. Hay que tener en cuenta que una celda del mapa (neurona) puede contener

varios perfiles de de un consumidor de los datos de entrada utilizados para el entrenamiento, pero solamente las etiquetas que más se repiten para una misma celda, son las mostradas en el mapa para una mejor comprensión del mismo, ver el zoom en la Figura 67.

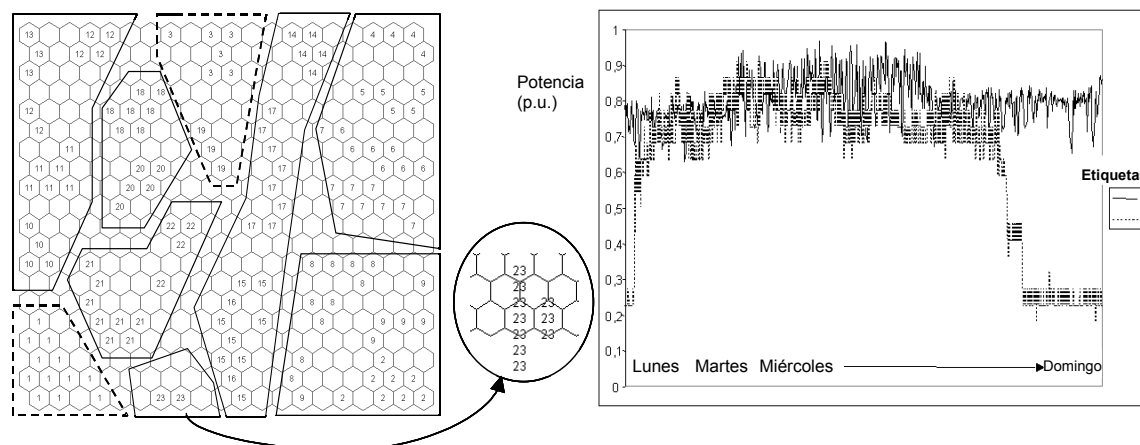


Figura 67 a&b. Mapa SOM entrenado con datos en el dominio de la frecuencia y sus correspondientes perfiles de carga semanal (consumidores etiquetados con 4 y 8)

Las industrias están separadas en varios clusters, por un lado encontramos agrupadas las etiquetas 4, 5, 6 y 7, es decir, industrias con una demanda elevada durante los fines de semana y por otro lado las etiquetas 2, 8 y 9, industrias sin una demanda continuada durante los fines de semana, ver la Figura 67.b.

En el cluster de la parte superior izquierda del mapa nos encontramos agrupadas las etiquetas correspondientes a las universidades (de la 10 a la 13). En la parte central izquierda del mapa aparecen en dos clusters pero muy próximos los consumidores minoristas (etiquetas 21 y 22) y los centros médicos (18 y 20). El CT residencial (etiqueta 23) aparece en un único cluster próximo también a los anteriores.

Por último, quedan dos clusters agrupando consumidores que aun teniendo diferentes actividades (etiquetas de la 14 a la 17 en el cluster de la parte central del mapa, y etiquetas 1, 3 y 19, en el cluster con línea discontinua) tienen similares comportamientos de consumo.

En una primera impresión, tras observar el mapa entrenado y los clusters obtenidos, es posible afirmar que se ha conseguido una mejor y más clara agrupación de los consumidores, pero es necesario medir la calidad del mapa entrenado de forma analítica mediante algunos índices. Esta calidad es analizada mediante el término de Error de Cuantización (“*Quantization error -Qe*”, visto en el capítulo 5) [112], el cual mide la resolución del mapa y mediante el Error Topográfico (“*Topographic error -Te*”) el cual mide la distorsión del mapa. El Error de Cuantización es calculado a partir de los datos de entrada mediante la siguiente fórmula:

$$Q_e = \frac{1}{N} \sum_{i=1}^N \|x_i - m_c(x_i)\|$$

Donde “**N**” es el número de vectores que forman el conjunto de entrada “**x<sub>i</sub>**” es cada vector de entrada, “**m**” es el vector de peso y “**c**” indica la neurona o celda mejor casada o asignada (The Best-Matching Unit) BMU para “**x**”. Tras el entrenamiento, se realiza un proceso de búsqueda de la BMU para cada vector de entrada, devolviendo el valor promedio de los respectivos valores de los errores de cuantización obtenidos para cada vector. Este valor promedio es el “**Qe**”.

El Error Topográfico es calculado mediante el siguiente algoritmo:

$$T_e = \frac{1}{N} \sum_{k=1}^N u(x_k)$$

Donde **u(x<sub>k</sub>)** es 1 si la primera y segunda BMU consecutivas de **x<sub>k</sub>** no se encuentran próximas la una de la otra, de otra forma **u(x<sub>k</sub>)** es 0.

En este sentido, la ventaja de la red entrenada con datos convertidos en el dominio de la frecuencia es superior y queda comparada en la siguiente tabla, ver Tabla 14:



Tabla 14. Análisis de la calidad de los mapas SOM

<b>Casos de entrenamiento</b>	<b>Te</b>	<b>Qe</b>
Dominio del Tiempo	0.008	0.3065
Dominio de la Frecuencia	0.003	0.041

### 6.3. CLASIFICACIÓN E IDENTIFICACIÓN DE NUEVOS CONSUMIDORES

El segundo objetivo de la investigación que recoge este capítulo es estudiar la capacidad de los mapas SOM, entrenados con los diferentes tratamientos, para clasificar consumidores. Dos nuevos consumidores, un restaurante y un centro comercial (no utilizados como datos de entrada en los entrenamientos anteriores) han sido testeados. Se han evaluado tanto los resultados de testear en los mapas entrenados con datos en el dominio del tiempo, así como los mapas entrenados con datos en el dominio de la frecuencia. El objetivo es conseguir para cada nuevo conjunto de datos de entrada la neurona asociada más similar.

Se han desarrollado dos tests para validar los resultados obtenidos, un test visual (ver la Figura 68) y otro de tipo analítico basado en el Error de Cuantización, nombrado anteriormente. Ambos métodos están basados en la función de BMU (Best Matching Unit), suministrada por la “SOM Toolbox” [112]. Esta función proporciona la neurona (y la etiqueta si está disponible en la neurona), más cercana o próxima para cada nuevo vector presentado al mapa ya entrenado. También proporciona el correspondiente Error de Cuantización (Qe) asociado.

El primer mapa, entrenado en el dominio del tiempo, fue capaz de clasificar sin problemas el conjunto de curvas de carga diaria correspondientes al primer nuevo consumidor, un gran almacén (Mall). En este caso, el nuevo usuario es situado cerca del cluster de Universidades (etiquetas 12 y 13 de la Figura 68.a.). El segundo consumidor (Restaurante) presenta una gran incertidumbre y el mapa SOM no es capaz de presentar

un resultado claro de localización. Hasta un total de tres clusters fueron seleccionados por la red para situar (localizar) este nuevo consumidor, entre los que se encontraban las etiquetas 3, 15 y 23, ver Figura 68.b.

El segundo mapa, entrenado en el dominio de la frecuencia, muestra mejores resultados. Los nuevos consumidores presentados al mapa SOM (El gran almacén y el Restaurante) han sido situados en un único cluster. El restaurante ha sido asignado al “Cluster Residencial” y el gran Almacén ha sido situado en el “Cluster Universidades”. Ver las etiquetas 23, 12 y 13 en la Figura 68.b.

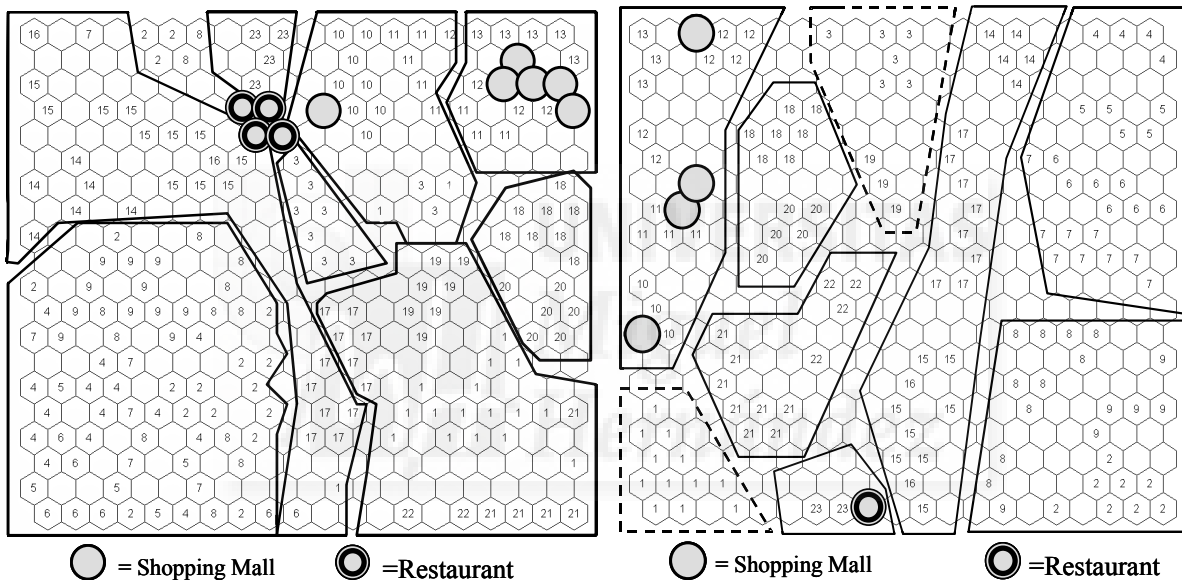


Figura 68 a&b. Clasificación de nuevos consumidores

Los resultados mostrados en la Tabla 15 verifican analíticamente las observaciones establecidas en los párrafos anteriores. El mapa SOM entrenado en el dominio del tiempo muestra altos valores del índice  $Q_e$  (Error de Cuantización), lo que significa una peor capacidad de identificación de nuevos consumidores, tanto para el consumidor tipo Gran Almacén ( $Q_e=1.24$ ) como para el consumidor tipo Restaurante ( $Q_e=0.93$ ). Por el contrario, el mapa SOM entrenado con datos de entrada en el dominio de la frecuencia presenta unos valores de Error de Cuantización para ambos consumidores muy bajos, concretamente  $Q_e=0.28$  para el Gran Almacén y  $Q_e=0.21$  para el restaurante.

Se puede concluir que la clasificación de nuevos consumidores es más precisa cuando el mapa SOM es entrenado con datos de entrada convertidos en el dominio de la frecuencia.

Tabla 15. Resultados del Testeo de Consumidores

Mapa SOM	Consumidores Testeados	Etiqueta más próxima identificada	Clusters Identificados	Qe (promedio)
Dominio del Tiempo	Centro commercial	1 neurona con etiqueta 10 3 neuronas con etiqueta 13 15 neuronas con etiqueta 12 11 neuronas entre las etiquetas 12 & 13	Grandes Universidades (3%) Campus Universitarios (97%)	1.24
	Restaurante	7 neuronas cercanas a la etiqueta 23 2 neuronas cercanas a la etiqueta 23 9 neuronas entre las etiquetas 15&23 17 neuronas cercanas a la etiqueta 3	C.T. Residencial (39%) Hoteles-Centros educativos (13%) Almacenes (48%)	0.93
Dominio de la Frecuencia	Centro commercial	23 neuronas con etiqueta 10 2 neuronas con etiqueta 11 5 neuronas con etiqueta 12	Universidad (100%)	0.28
	Restaurante	35 cells near label 23	C.T. Residencial (100%)	0.21

## 6.4. ESTUDIO DE RESULTADOS DE CLASIFICACIÓN CON MÁS DE 100 CONSUMIDORES

Aunque el estudio anterior se ha realizado para un espectro de 23 consumidores, se han realizado nuevas simulaciones con el objetivo de comprobar la capacidad de clusterización de los mapas SOM ante un número mayor de consumidores, situación más cercana a la realidad. Se pretende demostrar la escasa o nula variación de los resultados de la segmentación, aunque se aumente el número de muestras y tipos de usuarios. Para estas nuevas simulaciones se seleccionó un espectro de datos de entrada formado por un total de 119 consumidores. Estos usuarios han sido etiquetados desde la etiqueta 24 hasta la 119, ver la Tabla 16.

Tabla 16. Nuevo Espectro de Consumidores

Descripción del Consumidor	Etiquetas
Industrias	1,2,3,4,5,6,7,8,9, y desde la 24 hasta la 73
Universidades	10,11,12,13 y desde la 74 hasta la 89
Hoteles	14,15, y desde la 90 hasta la 93
Centros Educativos	16, y desde la 94 hasta la 97
Hospitales y Centros médicos	17,18,19,20 y desde la 98 hasta la 113
Grandes Almacenes	21,22, y desde la 114 hasta la 117
CT Residencial	23,118 y la 119

La capacidad de clusterización de los mapas SOM ha sido evaluada en los casos de datos de entrada convertidos en el Dominio del Tiempo y con datos de entrada convertidos en el Dominio de la Frecuencia. El tamaño del mapa considerado ha sido el mismo que el utilizado a lo largo de la investigación anterior, es decir, 20x20 neuronas. El número de pasos de entrenamiento primario y secundario fue de 5000 y 3000 respectivamente. En este caso aun siendo mayor el número de datos de entrada se ha mantenido el tamaño del mapa para comparar los resultados de la clasificación.

#### Resultados en el Dominio del Tiempo

La Figura 69 muestra los resultados después del proceso de entrenamiento. En este mapa aparecen las etiquetas asociadas a cada consumidor. Resaltar que un número alto de etiquetas coincidentes en una misma neurona del mapa causa una superposición de varios números en una misma celda (neurona).

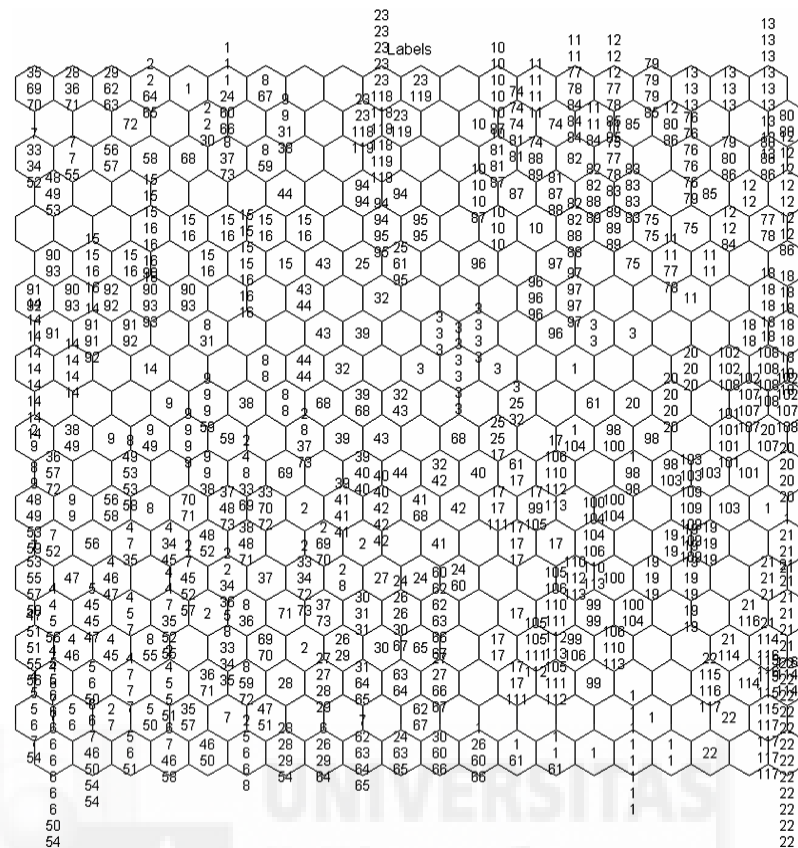


Figura 69. Mapa entrenado con datos de entrada convertidos en el dominio del tiempo (119 consumidores)

Tras el entrenamiento es posible dividir el mapa en clusters aplicando la función k-means [73] limitando el número máximo de clusters a encontrar a 15. El número óptimo de clusters es seleccionado utilizando el índice de Davies-Bouldin, así como una inspección visual del mapa de etiquetas. Los resultados se muestran en la Figura 70 y los clusters encontrados están delimitados en el mapa, en total aparecen 12 nuevos clusters. Para simplificar la comprensión del mapa de clusters, se ha decidido mostrar únicamente las etiquetas más significativas que aparecen en cada celda o neurona.

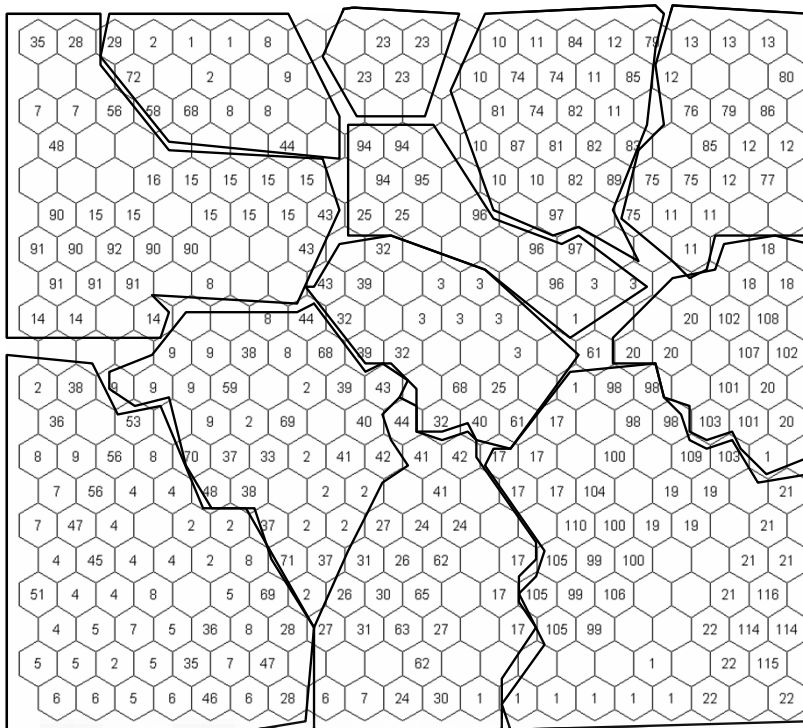


Figura 70. Mapa SOM entrenado y clusterizado con valores en el dominio del tiempo (12 Clusters y 119 Consumidores).

A diferencia del mapa SOM entrenado con menos consumidores en este aparecen 4 nuevos clusters, sin embargo el mapa presenta la misma apariencia que el mapa entrenado en la sección 6.2.2. Por ejemplo la etiqueta 17 aparece en neuronas pertenecientes a dos clusters diferentes, también encontramos un gran cluster que contiene consumidores industriales en la esquina inferior izquierda del mapa. Los consumidores tipo Universidad han quedado divididos en dos clusters diferentes, en la parte superior derecha del mapa. La Figura 71 representa gráficamente el comportamiento de la demanda de los consumidores etiquetados como 10 y 11 con respecto a los consumidores etiquetados como 12 y 13. Obviamente el comportamiento de la carga es muy parecido a lo largo de todo el día. Por último destacar que los consumidores tipo Hospitales y Almacenes han sido agrupados en un mismo cluster, tal como indican las etiquetas 17, 19, 20, 21 y 22 en la parte inferior derecha del mapa.

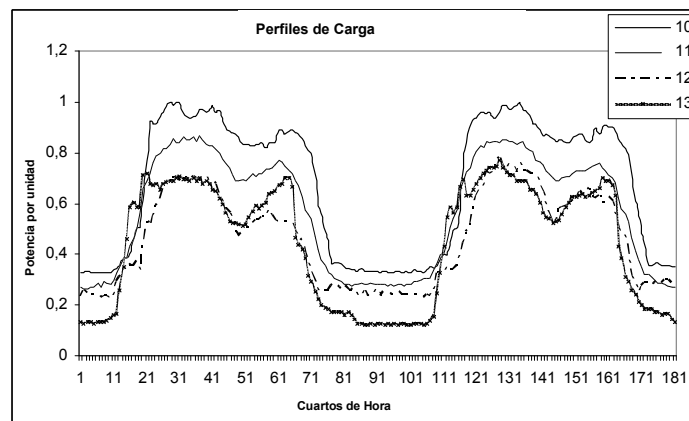


Figura 71. Comparación gráfica de los Perfiles de carga de los consumidores etiquetados como 10, 11, 12 y 13

### Resultados en el Dominio de la Frecuencia

Tras el entrenamiento también se ha limitado el número máximo de clusters a encontrar a 15. En este caso el índice de Davies-Bouldin mínimo se obtiene con el mapa dividido en 13 clusters. Estos clusters son mostrados en el mapa de la Figura 72.

Las principales ventajas obtenidas con este mapa entrenado en el Dominio de la Frecuencia frente al mapa entrenado en el Dominio del Tiempo son una mayor separación y claridad entre clusters y una mejor agrupación de los consumidores.

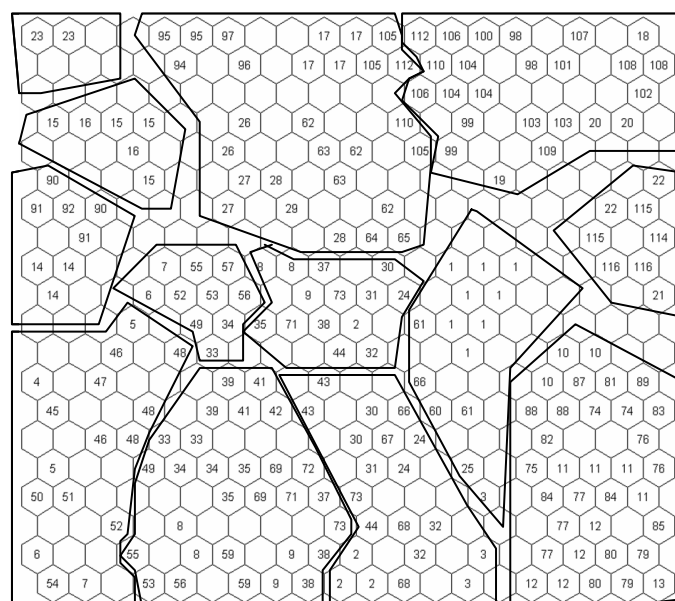


Figura 72. Mapa SOM entrenado y clusterizado con valores en el Dominio de la Frecuencia.

Por ejemplo, las universidades están agrupadas en un único cluster (muy razonable teniendo en cuenta el comportamiento de las curvas de carga de este tipo de consumidores). Por otro lado, las industrias están separadas en varios clusters en función del comportamiento de sus curvas de carga durante los fines de semana. Los consumidores industriales etiquetados como 47 y 4 tienen una demanda continua tal como muestra la Figura 73.

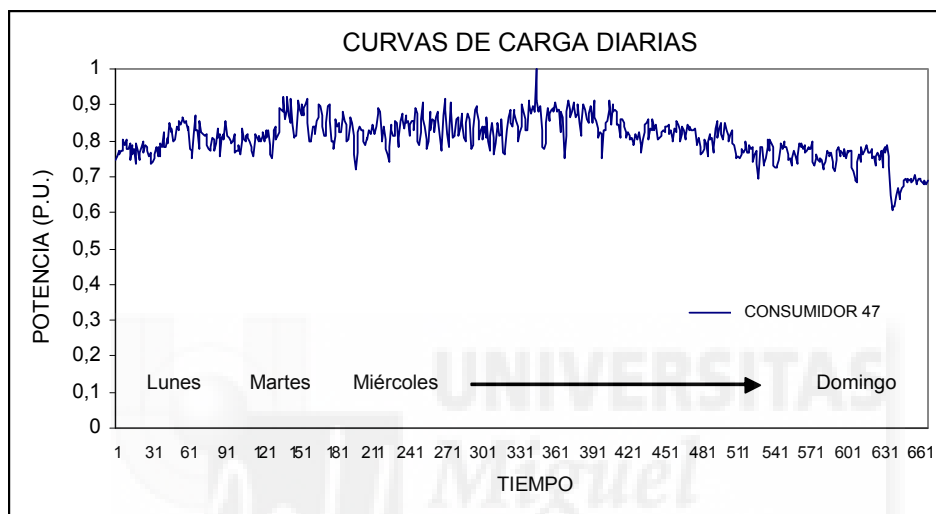


Figura 73. Perfil de carga semanal del consumidor industrial con etiqueta 47

A modo resumen destacar que la utilización de un espectro de datos mayor no cambia de forma significativa los resultados de agrupación obtenidos por los mapas entrenados. Evidentemente la diferencia principal es que aparecen más clusters debido a los nuevos perfiles de demanda de los consumidores que se han añadido.



## 6.5. CONCLUSIONES

A lo largo de la investigación se ha vuelto a demostrar la capacidad de la red neuronal SOM como técnica para afrontar la segmentación y clasificación de patrones de demanda de diferentes consumidores eléctricos, partiendo de las mediciones de potencia existentes en las bases de datos.

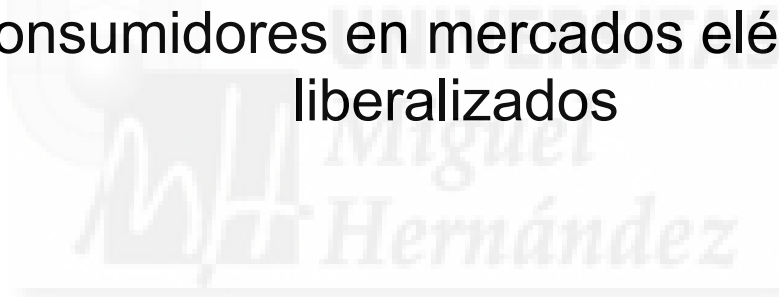
La propuesta que recoge este capítulo de transformar en el dominio de la frecuencia los datos de entrada con el objetivo de extraer más información de los perfiles de demanda originales muestra una mejora en la capacidad de clusterización de los mapas SOM. Esto queda patente en las reducciones conseguidas en los índices de error de cuantización ( $Q_e$ ) y error topográfico ( $T_e$ ), así como en la mejor visualización de los mapas. También es importante resaltar la importancia de algunos armónicos (vectores FWES) en el reconocimiento de patrones de consumo de fines de semana, no considerados en trabajos previos, así como la compresión de los datos de entrada que se ha logrado para los entrenamientos partiendo de los datos originales del espectro de consumidores del estudio.

El método presentado en este capítulo puede ayudar de manera efectiva a comercializadores y distribuidores a la hora de segmentar y clasificar consumidores eléctricos. Este es el primer paso para evaluar los costes de muchas políticas necesarias desde el punto de vista del lado de la demanda: el potencial de alternativas de eficiencia energética, respuesta de los consumidores a tarifas de precios en tiempo real (TOU), el éxito de aplicaciones de almacenamiento de energía, o posibilidades de utilizar generación distribuida en pequeños y medianos consumidores. Las futuras líneas de investigación, algunas de ellas ya en estudio en estos momentos, estarán centradas en el desarrollo de dos objetivos: la mejora de los índices de segmentación utilizados en el mapa SOM y el desarrollo de nuevas herramientas basadas en redes neuronales artificiales para identificar el potencial interés de algunos consumidores a la hora de participar en mercados eléctricos a corto plazo.

# CAPITULO 7

---

Herramienta para la elección de recursos de energía distribuida y consumidores en mercados eléctricos liberalizados





## 7.1. INTRODUCCIÓN

Como se ha venido comentado en capítulos anteriores, los procesos de liberalización de los mercados eléctricos en algunos países industrializados no ha tenido todo el éxito que en un primer momento se esperaba, debido principalmente a los problemas surgidos desde el año 2000; por ejemplo las caídas del sistema eléctrico que se produjeron en Europa, Estados Unidos y Canadá en el año 2003. Debido a estas experiencias, los operadores de los sistemas eléctricos creen, cada vez más, que los Recursos de Generación Distribuida (Distributed Energy Resources - DER) [125, 126] deberían ser aplicados utilizando un proceso integrado que tuviera en cuenta políticas de la demanda. Por ejemplo, obtención de ganancias desde el punto de vista de la demanda (Demand-Side) en el largo plazo o reacción frente a los precios en el corto plazo.

La correcta contribución de estos programas y la necesidad de ofrecer diferentes elecciones energéticas a los consumidores pequeños y medianos, no solamente a los grandes consumidores, necesita de un conocimiento detallado del potencial de la demanda a través de la definición previa de segmentos de la misma, el conocimiento de la capacidad de respuesta de los consumidores, la agrupación y/o agregación de los mismos, y finalmente la caracterización de estos clusters o grupos de consumidores.

La enorme cantidad de información disponible presenta un importante problema a las empresas públicas pero también una oportunidad nada despreciable para la investigación en las empresas distribuidoras y comercializadoras. Es necesario por tanto herramientas avanzadas para sintetizar y modelizar la información de las grandes bases de datos, es decir, nuevas investigaciones relacionadas con el “*Data Mining*” [10] y sus aplicaciones [8]. La evaluación y búsqueda de nuevas aplicaciones en mercados energéticos liberalizados para estas tecnologías de clasificación es uno de los principales objetivos que recoge este capítulo.

## 7.2. EL MERCADO ELÉCTRICO ESPAÑOL

Podemos resumir que el mercado de eléctrico español engloba el conjunto de mecanismos que permiten conciliar la libre competencia en la generación de electricidad con la exigencia de disponer de un suministro que cumpla con los criterios de seguridad y calidad requeridos. Las transacciones de energía que los agentes negocian en el mercado de producción responden a sus previsiones de demanda, de capacidad de generación de los grupos y de disponibilidad de la red de transporte. Este mercado está basado en varios procesos interrelacionados:

- **Mercado Diario:** tiene por objeto llevar a cabo las transacciones de energía para el día siguiente. El Operador del Sistema eléctrico español (REE), comunica a los agentes a las 8:30 horas del día anterior su previsión de demanda, las indisponibilidades de generación y la situación de la red de transporte. Los agentes que desean participar en el mercado diario, en el caso Español presentan al Operador del Mercado entre las 8:30 y las 10 horas sus ofertas de compra o venta de energía, procediendo éste a la casación de dichas ofertas, determinándose de esta forma el precio marginal y el volumen de energía que se acepta para cada unidad de compra y venta en cada periodo horario.
- **Mercado Intradiario:** gestionado por el Operador del Mercado, es un mercado de ajustes de los desvíos en generación o en demanda que se pueden producir con posterioridad a haberse fijado el Programa Diario Viable Definitivo. En España este mercado está organizado en seis sesiones y pueden presentar ofertas de compra o venta de energía aquellos agentes que hayan participado en la sesión del mercado diario. Los precios casados en este mercado son superiores a los precios casados en el mercado diario.

Por último existe el Mercado de Operación en Tiempo Real, lo que se conoce como mercado de balance (conocidos en otros países también como *Real-Time Market*) donde se producen los procesos y mecanismos mediante los cuales se resuelven los desequilibrios que puedan surgir entre generación y demanda. El operador del mercado

identifica y resuelve las incompatibilidades que aparecen entre las transacciones de energía acordadas en el mercado de producción y la gestión de su transporte. Para su resolución parte de los resultados del mercado diario e intradiario, las ventas en régimen especial y los contratos bilaterales físicos y reasigna algunas producciones para asegurar la viabilidad del resto. Dentro de este contexto en el mercado de operación en tiempo real (mercado de balance) existe la regulación secundaria (regulación en tiempo real). Su objetivo es mantener la capacidad de restablecer los desequilibrios entre generación y demanda en un plazo de 30 minutos a un par de horas. El producto que se negocia por los agentes del mercado, es decir, entre productores (unidades generadoras) y demandantes (agregadores, distribuidores y grandes clientes) es la capacidad de subir o bajar generación, y es retribuido por dos conceptos: disponibilidad (banda) y utilización (energía).

En el mercado español la gestión de balances o desvíos tiene por objetivo resolver los desvíos entre la generación, “por averías en los grupos”, y el consumo, si la demanda casada no coincide con la prevista en el plazo que transcurre entre el cierre de cada sesión del mercado intradiario y el inicio del horizonte de efectividad de la siguiente sesión. Sin embargo, en otros mercados eléctricos a nivel internacional nos podemos encontrar que dentro del mercado de balance la regulación del sistema en tiempo real contempla programas de respuesta de la demanda (Load Response Programs) donde la parte consumidora puede influir en la demanda de recursos eléctricos, ya que estas pueden participar en el mercado a través de programas que ofrecen las comercializadoras y distribuidoras. Estos programas van a compensar al consumidor económicamente, dependiendo del tipo de programa al que se haya acogido.

### 7.3. PROGRAMAS DE RESPUESTA DE LA CARGA. OPORTUNIDADES

Diferentes operadores independientes de sistemas eléctricos (Independent System Operators ISO) [17], [17, 127-131] en Europa, Oceanía y Norteamérica han desarrollado Programas de Respuesta de la Carga (LRP- Load Response Programs) [128] con el objetivo de lograr cambiar la demanda eléctrica de grandes consumidores eléctricos. Sin embargo, en estos últimos años han disminuido los requisitos de participación de consumo [15]. En este sentido, han sido muy frecuentes niveles de reducción cercanos a 100kW para poder permitir a los consumidores su participación. Esto quiere decir que algunos consumidores de tipo medio como industrias o centros comerciales podrían lograr ofertas y ahorros en estos nuevos mercados liberalizados.

Dentro de este marco los operadores de los mercados eléctricos deberían animar a las empresas comercializadoras a alcanzar unos niveles mínimos de agregación de la demanda con la finalidad de lograr un nivel mínimo de reducción para facilitar la participación de los pequeños y medianos consumidores. Estos mismos comercializadores serían los responsables de seleccionar los tipos o patrones de carga, y utilizar mecanismos de agregación y/o agrupación de la misma.

El problema, llegados a este punto, es como detectar cuales son los consumidores más idóneos (capaces) de alcanzar estos niveles mínimos de demanda y poder afrontar reducciones de consumo eléctrico en operaciones de mercado a corto plazo y con un tiempo de respuesta lo suficientemente bajo.

Un operador de mercado (ISO) interesado en promover programas de respuesta de la demanda ha sido el Operador Independiente del mercado de Nueva Inglaterra [128], el cual ofrece dos alternativas distintas:

- Programas de respuesta de la carga en el mercado diario, es decir en el mercado de compra-venta de energía que se produce el día anterior al día considerado, (Day-Ahead Load Response Programs)

- Programas de respuesta al precio de la demanda en tiempo real (Real-Time Demand Price-Response Programs)

### **7.3.1. PROGRAMAS DE RESPUESTA DE LA DEMANDA EN EL MERCADO DIARIO (DAY-AHEAD LOAD RESPONSE)**

Estos programas permiten a los participantes (medianos y grandes consumidores, distribuidores o comercializadores) ofrecer reducciones de consumo eléctrico durante las operaciones del mercado de la energía diario (Day-Ahead Energy Market). En este sentido, se les comunicará a los consumidores o participantes ofertas en firme de la misma forma que tienen lugar dentro del mercado diario. Si el participante es un distribuidor, este tiene las siguientes responsabilidades:

- Proporcionar el nivel de carga necesario de acuerdo a los parámetros de la oferta en firme.
- Notificar al consumidor la necesidad de reprogramar o de planificar la demanda resultante para las ofertas aceptadas.
- Ayudar al consumidor en la determinación o elección del mejor conjunto de políticas de respuesta de la demanda (carga) en función del periodo de reducción de la carga y de los niveles de precios.

Para mejorar la participación de empresas distribuidoras, en primer lugar habría que conocer en el espectro de consumidores (clusterización o agrupación de consumidores), la curva de carga demandada que mejor sigue los precios de la energía eléctrica, es decir, una comparación o correlación entre la curva de precios diaria y la curva de carga diaria. El objetivo teórico es alcanzar una gran cantidad de carga disponible para permitir recortes y reducciones. Un segundo objetivo sería comprobar que esta carga (demanda de los consumidores) tuviera capacidad de reducir el consumo (carga) en aquellos periodos donde fuera más interesante aprovechar estas reducciones. Esta tarea puede llevarse a cabo mediante modelos de carga y de generación como los descritos en [132].



### **7.3.2. PROGRAMAS DE RESPUESTA DE LA DEMANDA EN EL MERCADO A TIEMPO REAL (REAL-TIME LOAD RESPONSE PROGRAMS)**

Los consumidores y participantes en los programas de reducción de la carga en tiempo real estarían de acuerdo en alcanzar un cierto nivel de reducción de la carga bajo la supervisión del ISO en un periodo de tiempo a corto plazo, de 30 a 120 minutos, mediante controles en la carga (consumo) o mediante generación propia para autoabastecimiento durante estos periodos.

El periodo de aviso requiere una respuesta rápida por parte de los consumidores acogidos a estos programas y el desarrollo de nuevas herramientas para afrontar las órdenes de reducción de consumo (potencia) para no sufrir multas o penalizaciones económicas.

El operador del mercado comunica la necesidad de la participación a los consumidores acogidos a los programas, a través de un proveedor de servicios de comunicación por Internet [128]. Esta reserva de carga (oferta) debería estar disponible entre 2 y 4 horas, de forma general cuantos más cortos son los periodos de notificación más altos son los precios pagados (por ejemplo desde 100\$/MWh a 500\$/MWh). Obviamente estos precios se presentan como una oportunidad muy interesante para gestionar y administrar de una forma más eficiente los costes de la energía, mientras se contribuye a la mejora del funcionamiento del sistema eléctrico.

De nuevo en este caso, los participantes (comercializadores y/o distribuidores), deben ser capaces de seleccionar a los consumidores y adecuar las políticas de respuesta de la demanda para lograr altos beneficios. Esto supone desarrollar predicciones de precios en tiempo real y crear un abanico de políticas de precios en Tiempo Real, con la finalidad de encontrar la mejor combinación (políticas vs. consumidores vs. precios)

Actualmente el mercado Español no contempla la posibilidad de que la demanda (Demand-Side) participe en los mercados diarios, intradiarios o en el de desvíos (mercado a tiempo real). Por este motivo y para investigar y desarrollar las posibilidades que pueden ofrecer algunas herramientas basadas en Redes Neuronales Artificiales para

ayudar a la gestión de la demanda de los posibles participantes que se puedan dar en un futuro, en esta Tesis se han seleccionado dos clases de precios de mercado para simular la respuesta del consumidor en esos mercados: precios de la energía en el mercado diario y precios de la energía en el mercado de desvíos (mercado de balance). Ambos precios han sido recogidos de la página web del operador del mercado eléctrico español ([www.omel.es](http://www.omel.es)) [24] desde Octubre del 2001 hasta Marzo del 2002. La razón de elegir este periodo de tiempo es porque en él se alcanzaron los mismos niveles de precios que en el caso de los mercados eléctricos de Estados Unidos los cuales si permiten la participación de la demanda. Por ejemplo, con el mercado de Nueva Inglaterra donde se alcanzaron precios entre 150\$/MWh y 500\$/MWh [128].

### **7.3.3. ANÁLISIS DE LOS PROGRAMAS DEL MERCADO DIARIO.**

Para analizar estos mercados y explorar las posibilidades de los participantes, se ha escogido un conjunto de mediciones de potencia del periodo de invierno comprendido entre Enero y Febrero del 2003 y 2004. Las curvas de carga pertenecen a consumidores de tipo industrial, institucional y de tipo comercial o terciario, así como un pequeño Centro de Transformación de una zona residencial, todos ellos pertenecientes a la zona de Alicante. El pico anual de la potencia alcanzada de las muestras varía entre 100kW y 10MW.

La Tabla 17 muestra la descripción y la etiqueta asociada a las curvas de carga diaria de cada uno de los consumidores analizados y utilizados también en capítulos anteriores. Las Figura 74 y Figura 75 muestra las curvas de carga de 4 consumidores distintos.

Tabla 17. Espectro de Consumidores

SECTOR DEL CONSUMIDOR	Actividad	Etiqueta	Número de vectores (curvas de carga)
Industria	Industria Mediana	1,2	21
	Almacenes	3	21
	Grandes Industrias	4 to 9	28
Institucional	Universidad de tamaño medio	10,12, 13	40
	Universidad de gran tamaño	11,19	28
Comercial o Terciario	Pequeños hoteles	14,15	21
	Hoteles de tamaño mediano	17	21
	Hospitales y centros médicos	20 to 23	21
	Retailers	24,25	21
Residencial	Pequeños consumidores	26	40

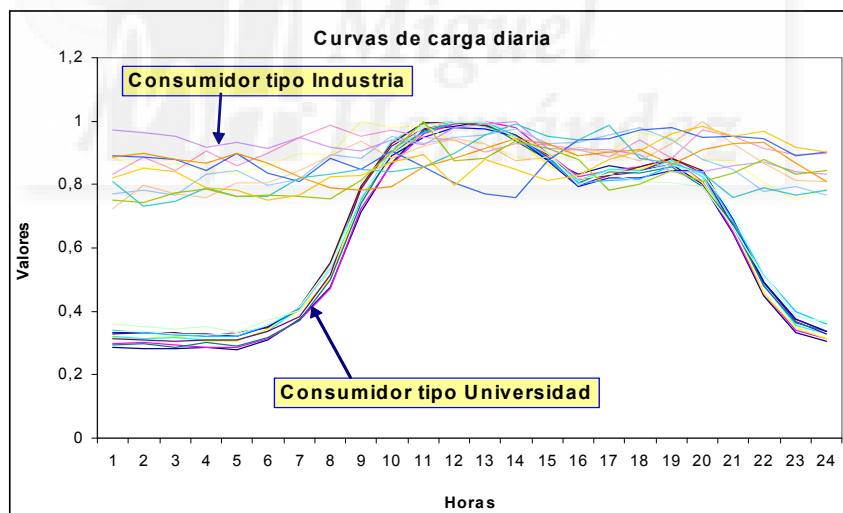


Figura 74. Curvas de carga de los consumidores con Etiquetas 13 y 5

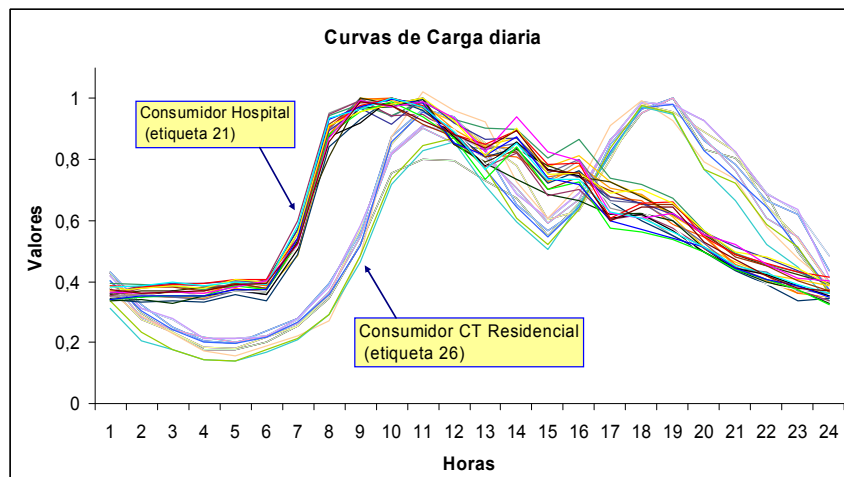


Figura 75. Curvas de carga de dos consumidores con etiquetas 26 y 21

Al igual que se ha realizado en los capítulos anteriores se ha entrenado un mapa SOM con las curvas de carga diaria al objeto de lograr una clusterización o agrupación de consumidores eléctricos.

La Figura 76 muestra nuevamente un mapa obtenido tras el entrenamiento, así como dos ejemplos claros de clusters. Uno de ellos definido por agrupar y contener a las Universidades y al consumidor tipo residencial, es decir, el formado por las etiquetas 10, 11, 12, 13 y 26. Y el otro cluster por contener a un consumidor de tipo Hospital, etiqueta 22.

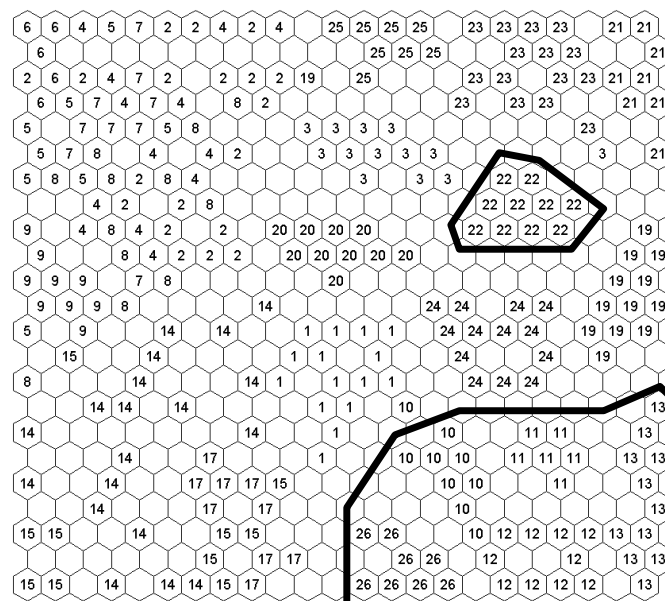


Figura 76. Mapa SOM entrenado con las curvas de carga de los consumidores

Como se ha comentado anteriormente, estos consumidores no tienen capacidad para gestionar su posible participación en el mercado, tal vez por la dificultad de entender la reglas y los procedimientos de participación en los mercados, o quizás porque el consumidor es incapaz por sí mismo de alcanzar los niveles de reducción de la carga requeridos por el sistema en un momento dado (de 100kW hasta 200kW). En ambos casos la participación de un comercializador o distribuidor es bastante probable y necesaria.

Desde el punto de vista del comercializador participante (agregador o comercializador) sería necesario llevar a cabo una coordinación y seguimiento entre la demanda de los consumidores adheridos a los programas y los niveles de los precios en el mercado diario. Teóricamente, aquellos consumidores cuya curva de carga (demanda) sigue a la curva de precios diaria en el mercado diario de la energía serían los primeros candidatos en aceptar un recorte de algunas fracciones de su consumo de energía eléctrica. Es decir, cuantos más altos sean los costes de la energía dentro de los clusters de los consumidores, más posibilidades deberían aparecer para reducir o modificar la carga por parte de algunos consumidores. Para localizar estos clusters de consumidores, se presentaron al mapa entrenado varias curvas (representativas) de precios de la energía en el mercado diario (convertidas en valores por unidad, con un precio base máximo de 12 céntimos de euro). En la Tabla 18 se muestran los resultados de las neuronas ganadoras. El mapa asocia a cada una de las curvas de precio a la neurona o celda del mapa más coincidente o similar con el patrón y forma de la misma (conocida como BMU, Best Matching Unit). De esta forma es posible determinar las etiquetas que contienen las neuronas asignadas a cada una de las curvas de precios. Observando los resultados de la Tabla 18 se desprenden los grupos de consumidores a los que más les interesa administrar y gestionar sus gastos de consumo de energía acogiéndose a posibles políticas DER (Distributed Energy Resources).

Tabla 18. Best Matching Units para cada curva de precio de la energía del mercado diario

Fechas de las curvas de precio de la energía en el mercado diario (Day-Ahead) (mm/dd)	BMU (número de la neurona asociada)	Consumidor asociado a la BMUs
01/13	299	26
01/14	276	26
01/15	299	26
01/24	414	12
01/30	529	13
02/07	466	12-11
02/14	466	12-11
10/09	202	17-14
11/19	253	17-26

La Figura 77 muestra la curva de precio de la energía en el Mercado diario del día 15 de Enero del 2002 frente a un curva de carga diaria del consumidor tipo residencial etiquetado con el número 26.

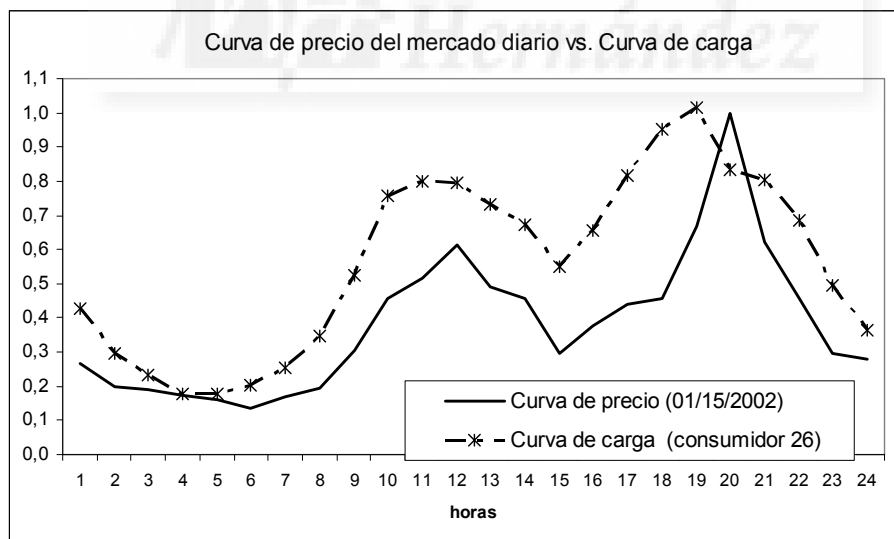


Figura 77. Comparación en valores por unidad de las curvas de precio y demanda

En este sentido el comercializador o agregador debería centrar los esfuerzos de incentivar y promover la participación de la demanda a través de políticas DER de aquellos consumidores seleccionados por el mapa SOM.

### 7.3.4. ANÁLISIS DE LOS PROGRAMAS DEL MERCADO EN TIEMPO-REAL

Aquellos consumidores participantes en los programas de Tiempo Real deberían reducir la carga o aumentarla, dependiendo de si disponen de sistemas de generación para el autoabastecimiento en un periodo de tiempo muy corto, periodos que pueden variar entre 30 minutos y 2 horas desde el momento que el operador del mercado les requiere para su intervención. Si el recorte o reducción de la demanda requerida por el operador (ISO), mediante control de cargas, almacenamiento de energía o mediante autogeneración no fuera alcanzada por los participantes el operador podría aplicar penalizaciones o multas.

Esto requiere la necesidad de conocer de forma precisa la capacidad de respuesta del consumidor y los posibles periodos de recortes o aumentos de energía para así definir los beneficios o pagos que recibirán los consumidores.

La primera tarea, desde el punto de vista de la respuesta de la demanda y las fuentes de generación para autoabastecimiento, sería estimar los posibles modelos de carga. El segundo objetivo, sería definir los patrones y características (si existen) de los precios en el mercado de la energía a tiempo real, esta tarea ha sido asignada a un nuevo mapa SOM.

Para extraer estas características a partir de series de precios anuales (en España correspondería al Mercado de gestión de desvíos, o balance [24]) desde el 2001 al 2002, se realizó un nuevo entrenamiento con un mapa SOM. La Figura 78 muestra el mapa y la mejor clusterización lograda tras el entrenamiento. En este mapa es posible observar 4 zonas o clusters. En cada una de las celdas aparece la etiqueta asociada a cada uno de los vectores o curvas de precios, codificadas en el formato de fecha mm/dd/aa, mes-día y año. La zona más grande es la 1, y es donde se agrupan los días con precios de mercado de desvíos cuyos valores máximos en el mejor de los casos no van a superar los 10 cents.€/kWh. (por ejemplo, alrededor de 125\$/MWh). Esta zona ha sido descartada para seguir con la investigación porque con estos niveles tan bajos de precios (de 100 a 150\$/MWh) es muy poco probable que el operador independiente del sistema (ISO)

inicie una notificación a los consumidores participantes, o abra un periodo de aplicación de uno de estos programas de respuesta en Tiempo Real.

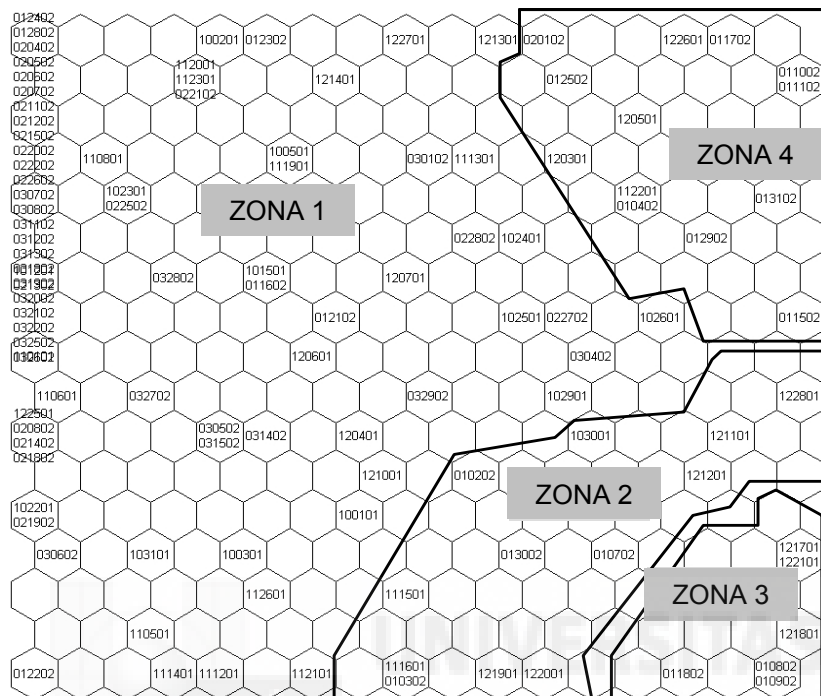


Figura 78. Mapa SOM entrenado con los precios del mercado de gestión de desvíos

Por otro lado, las Zonas 2, 3 y 4 si han sido consideradas para seguir con la investigación y extraer de los días que contienen características y patrones de comportamientos similares. La Figura 79 muestra los precios de la Zona 2 en cents.€/Kwh. (12 días dentro del periodo de estudio considerado), los cuales alcanzan niveles de 250€/MWh (es decir, alrededor de 300\$/MWh).

Tras analizar los resultados anteriores, se puede concluir que los consumidores pueden acogerse a diferentes tipos reducciones:

- Para los días de la zona 2, dos periodos de reducciones por día.
- Para los días de la zona 3, un único periodo de reducción, pero de valor alto.

Es lógico pensar que los participantes en estos programas deberían aplicar diferentes políticas DER para obtener y maximizar beneficios durante los periodos de reducción que notifique el operador del sistema.



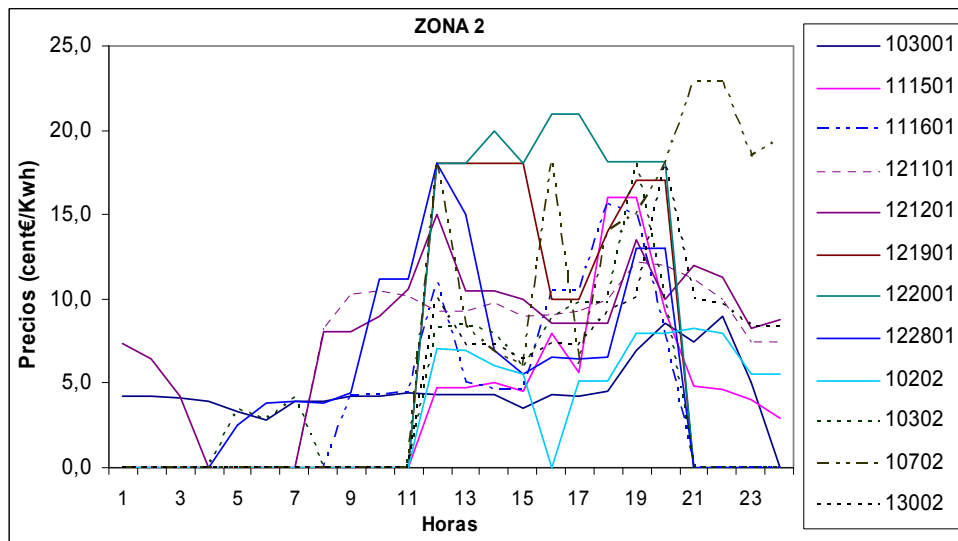


Figura 79. Ejemplo de curvas de precios de la zona 2

#### 7.3.4.1. OBTENCION PATRONES CARACTERISTICOS

Considerando que la clasificación de la red es correcta, podemos interpretar el conjunto formado por las series que clasifica en el mismo grupo como distintas realizaciones de un mismo proceso estocástico, y por tanto podemos caracterizarla por su función de medias la cual vendrá dada por los promedios en cada punto. Si lo que se busca es un patrón (como es el caso) podemos utilizar algún procedimiento no paramétrico para su obtención. Frecuentemente se suelen utilizar modelos de regresión no paramétrica en los que pretendemos estimar los promedios de la respuesta en función de la entrada (o punto en este caso), es decir, si consideramos:

$$E(Y|X) = m(X)$$

Proponiéndose para la variable respuesta un modelo de la forma:

$$Y = m(X) + \varepsilon, \quad E(\varepsilon) = 0$$

Y utilizamos uno de los estimadores más comunes en este tipo de procedimientos como es el de Badallara-Watson, cuya expresión viene dada por:

$$\hat{m}_h(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) y_i}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}$$

siendo  $K(x)$  la función núcleo utilizada (kernel), en nuestro caso se ha utilizado una Normal (0,1), y  $h$  la ventana utilizada.

Los intervalos de confianza para las estimaciones vienen dados por:

$$\left[ \hat{m}_h(x) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_h^2(x) \|K\|_2^2}{nh \hat{f}_h(x)}}, \hat{m}_h(x) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_h^2(x) \|K\|_2^2}{nh \hat{f}_h(x)}} \right],$$

donde:

$$\hat{\sigma}_h^2(x) = \frac{\frac{1}{n} \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) \{y_i - \hat{m}_h(x)\}^2}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}$$

Y  $\|K\|_2^2$  es la norma L2 del núcleo utilizado, y  $f_h(x)$  es la estimación de la densidad asociada en el punto  $x$  (estimada también por técnicas kernel).

Aplicando esta técnica es posible obtener los siguientes patrones para cada una de las zonas. Para la Zona 2 utilizamos regresión de Kernel con núcleo de Gauss y ventana de valor  $h=0.78$ , con validación cruzada e intervalo de confianza del 95%. Ver Figura 81 y Figura 80.

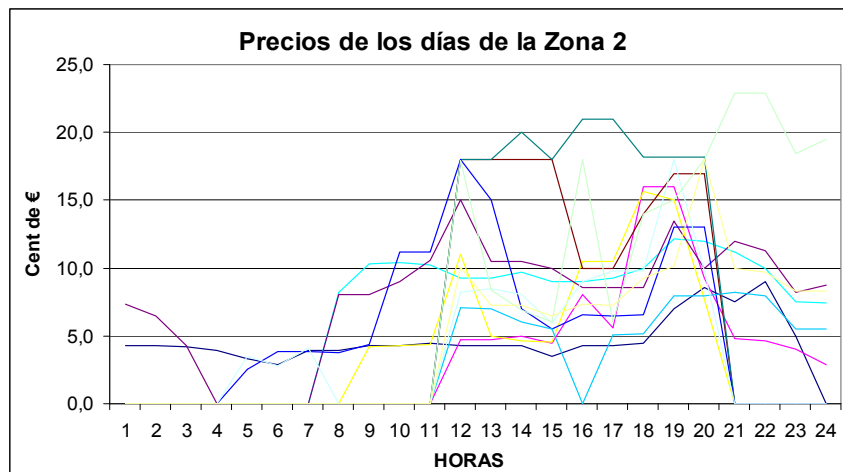


Figura 80. Precios de los días correspondientes al cluster denominado zona 2

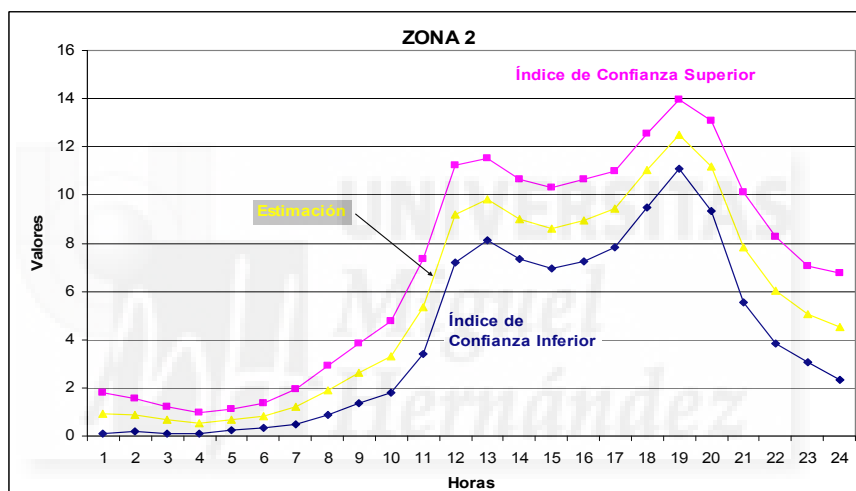


Figura 81. Curva de Estimación e índices de confianza para la zona 2 de precios

Para la Zona 3 volviendo a utilizar regresión de Kernel con núcleo de gauss y ventana de valor 1.07, con validación cruzada e intervalo de confianza del 95%. Ver Figura 83 y Figura 82.

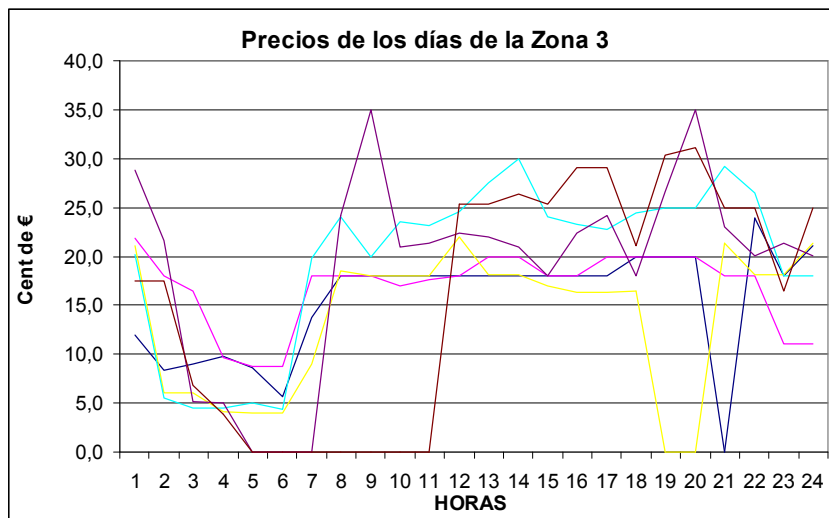


Figura 82. Precios de los días correspondientes al cluster denominado zona 3

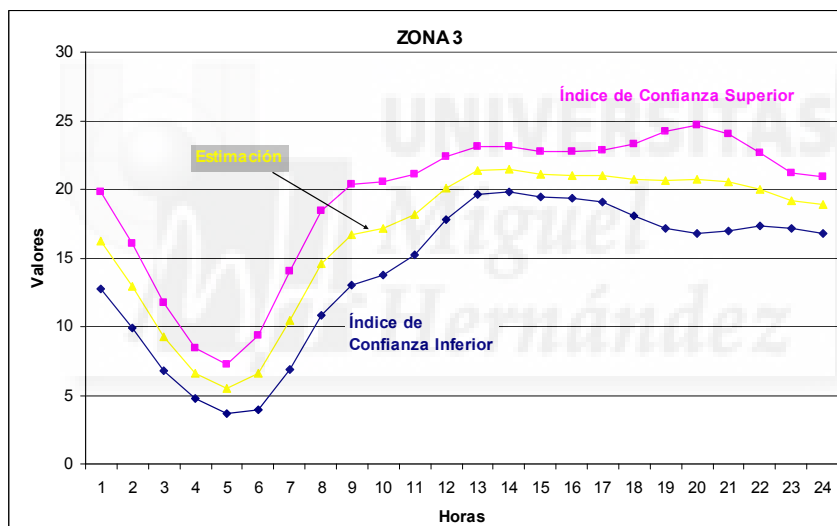


Figura 83. Curva de Estimación e índices de confianza para la zona 3 de precios

Para la Zona 4 volvemos a utilizar regresión de Kernel con núcleo de gauss y ventana de valor 0.85, con validación cruzada e intervalo de confianza del 95%. Ver Figura 84 y Figura 85.

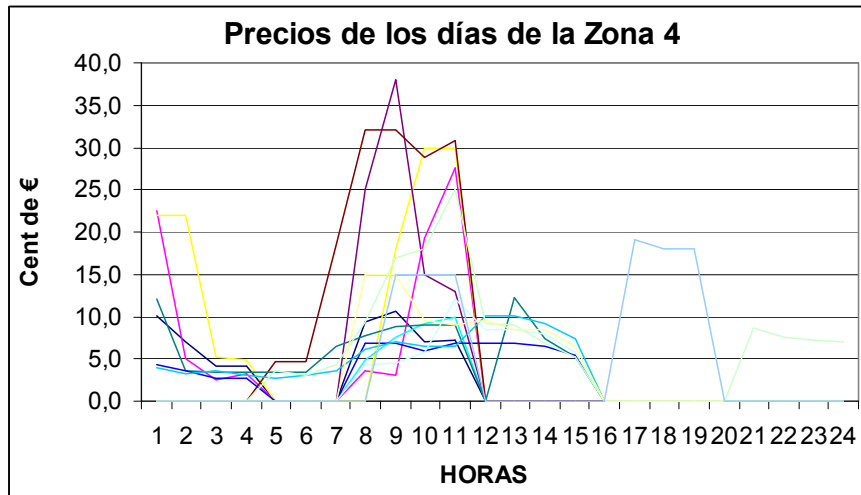


Figura 84. Precios de los días correspondientes al cluster denominado zona 4

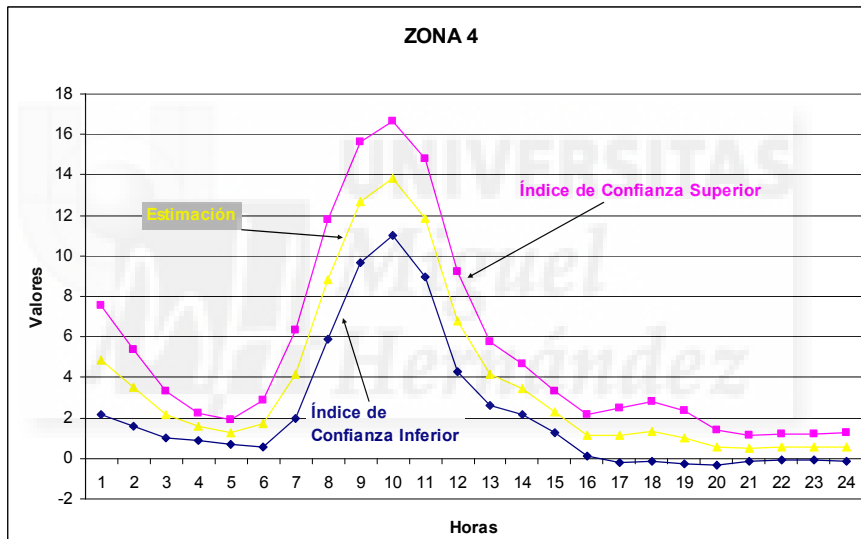


Figura 85. Curva de Estimación e índices de confianza para la zona 4 de precios

## 7.4. IDENTIFICACIÓN DE CONSUMIDORES CON CAPACIDAD DE APLICAR POLÍTICAS DER PARA PARTICIPAR EN PROGRAMAS DE RESPUESTA DE LA DEMANDA EN TIEMPO REAL.

Para realizar este análisis se ha seleccionado la zona 2 por tener un elevado número de días con precios altos y con una frecuencia de repetición importante, el siguiente paso es evaluar el potencial de la participación de los consumidores dependiendo de los niveles específicos de los precios de la energía en tiempo real. Para ello se ha utilizado la curva promedio de precios de la zona 2 con la finalidad de diseñar varias estrategias de control y gestión de la carga en función de la forma de esta curva.

Las políticas DER que se han utilizado corresponden a resultados de otra Tesis en curso dentro del grupo de investigación. Se han utilizado, por tanto, algunos resultados de aplicación de políticas de demanda, tales como:

Estrategias de intercambio de carga, LTS

Generación Distribuida, DG

Estrategias combinación de las dos anteriores, SC

Concretamente se han aplicado:

- LTS 1: Control del ciclo de funcionamiento de aparatos de calefacción eléctrica.
- LTS 2: Control del ciclo de funcionamiento calentadores eléctricos de agua.
- DG: Generación eléctrica in-situ por medio de placas solares fotovoltaicas
- SC 1: Combinación de los efectos de las estrategias LTS 1 y LTS 2 más la estrategia DG
- SC 2: Combinación de los efectos de las estrategias LTS 1 y LTS 2
- SC 3: Combinación de los efectos de las estrategias LTS 1 y DG
- SC 4: Combinación de los efectos de las estrategias LTS 2 y DG

Las estrategias de control propuestas fueron diseñadas de forma independiente para cada uno de los usuarios basándose en la forma de la curva de carga de cada uno de ellos, pero intentando que las reducciones de consumo obtenidas siguieran el patrón y tendencia marcada por el promedio de la curva de precios de los días obtenidos para la zona 2 del mapa SOM. El objetivo es conseguir el máximo beneficio económico. La Figura 86 muestra una curva de carga diaria para un consumidor de tipo Universidad, así como las curvas de demanda del consumo dedicado a calefacción eléctrica y a calentadores eléctricos de agua.

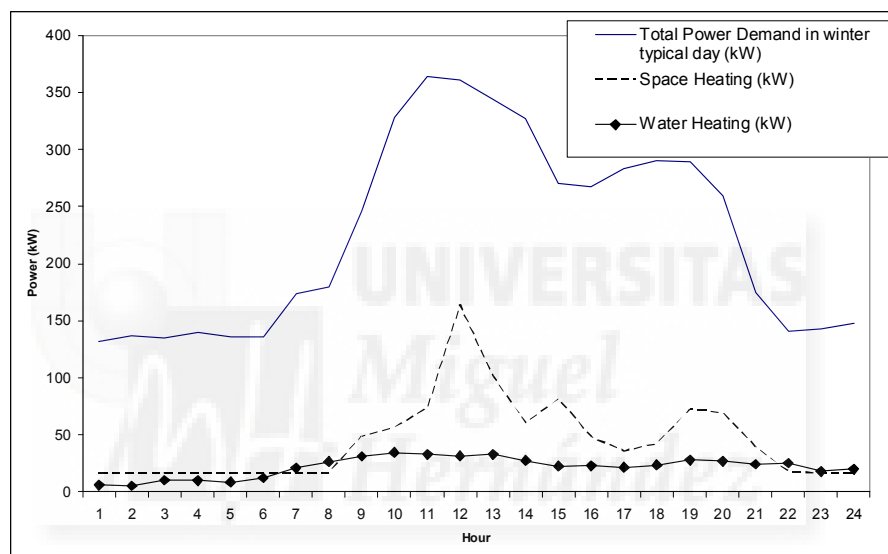


Figura 86. Curva de carga y usos finales de un consumidor tipo Universidad

Para reproducir el consumo real de los equipos (HVAC, Calentadores de agua, Iluminación, Paneles fotovoltaicos,...) se utilizó un software que permite implementar modelos de carga basados en dispositivos físicos reales [5], [133]. El autor de esta Tesis vuelve a insistir en que el uso de estos modelos de carga corresponden a resultados de otra Tesis en curso dentro del grupo de investigación.

Los usos finales de la energía, por parte de los consumidores, se estimaron, usando este software y teniendo en cuenta sus situaciones geográficas, así como sus características físicas de sus recursos. Una vez determinados los comportamientos de los usos finales, el siguiente paso fue evaluar con este software las diferentes acciones de

control de carga (reducción del ciclo de funcionamiento de los equipos eléctricos) y generación fotovoltaica.

Se testeó sobre el consumidor etiquetado como 24 (Centro comercial): 2 políticas de LTS, el control de la calefacción eléctrica y el control de los calentadores de agua eléctricos, y una estrategia de DG. Los resultados son mostrados en la Figura 87.

En el gráfico los valores positivos significan que el consumidor está reduciendo su consumo (Políticas de respuesta de la demanda), y las áreas negativas significan un aumento de consumo, es decir una recuperación de energía consumida y reflejada en la curva de carga diaria.

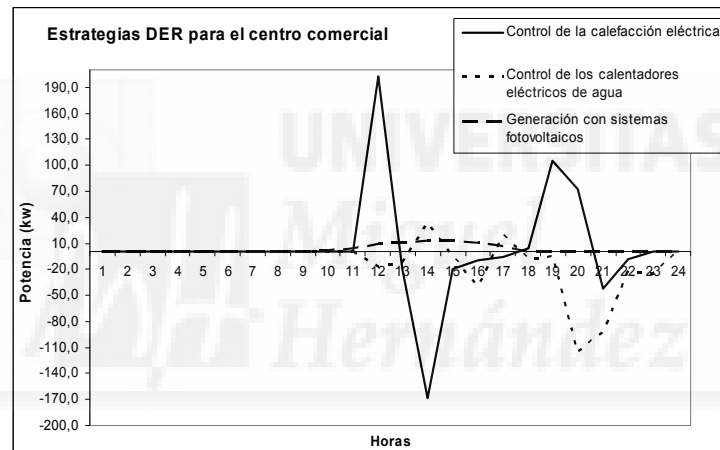


Figura 87. Ejemplo de políticas DER para un consumidor de tipo Centro Comercial

Del mismo modo la Figura 88 muestra los efectos del control de los calentadores eléctricos de agua caliente en la curva de carga diaria de un consumidor tipo Universidad, concretamente el etiquetado con el número 13.



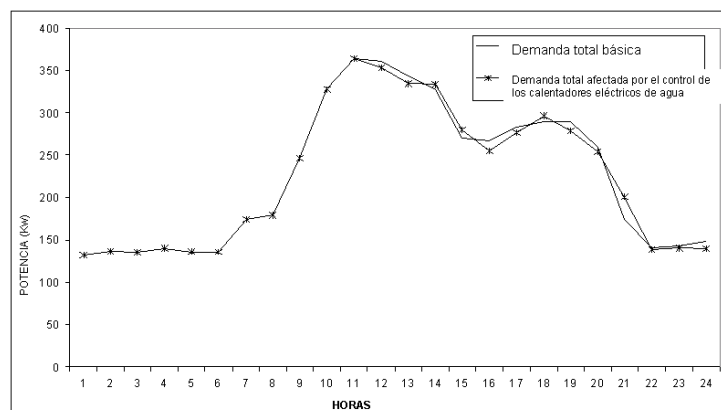


Figura 88. Curva de carga diaria con y sin política de control de los calentadores eléctricos de agua caliente

Las Tablas 19, 20 y 21 presentan los valores numéricos (recorte de picos, reducción de energía y retornos de energía) necesarios para evaluar los efectos económicos y técnicos de algunos consumidores, en este ejemplo en particular, de los consumidores etiquetados como 13, 24 y 26.

Tabla 19. Coeficientes de evaluación de políticas DER del consumidor Universidad

	Recorte de Picos (kW)	Reducción de Energía (kWh)	Retornos de energía (kWh)
LTS 1	3,52	25,69	69,20
LTS 2	0,00	13,53	49,80
DG	19,56	260,85	-
SC 1	31,94	300,08	119,01
SC 2	-3,78	39,23	119,01
SC 3	33,58	286,54	69,20
SC 4	19,56	274,38	49,80

Tabla 20. Coeficientes de evaluación de políticas DER del consumidor centro comercial

	Recorte de Picos (kW)	Reducción de Energía (kWh)	Retornos de energía (kWh)
LTS 1	4,42	112,57	272,73
LTS 2	-8,74	-286,75	340,79
DG	0,63	72,49	-
SC 1	-3,70	-101,69	613,52
SC 2	-4,32	-174,18	613,52
SC 3	5,04	185,06	272,73
SC 4	-8,11	-214,26	340,79

Tabla 21. Coeficientes de evaluación de políticas DER del consumidor CT Residencial

	Recorte de Picos (kW)	Reducción de Energía (kWh)	Retornos de energía (kWh)
LTS 1	28,77	63,27	127,84
LTS 2	10,43	-1,98	50,83
DG	1,35	75,02	-
SC 1	27,33	136,31	178,68
SC 2	26,48	61,28	178,68
SC 3	29,62	138,30	127,84
SC 4	11,79	73,04	50,83

Tras definir para cada consumidor las curvas de reducción el objetivo era intentar encontrar cuales son las mejores políticas de participación en los programas de respuesta de la demanda en Tiempo Real y cuales no. Este proceso viene detallado en el siguiente punto de este capítulo.

## 7.5. SELECCIÓN DE LAS POLÍTICAS DE RESPUESTA EN TIEMPO REAL

Una vez definidas las políticas y diferentes estrategias DER es el momento de elegir los productos adecuados, para que los consumidores puedan afrontar los recortes requeridos por los Operadores del sistema (ISO) en los mercados a Tiempo Real. Las reducciones de la carga (generación de energía) desde el punto de vista del mercado son contempladas como una oferta de la demanda (Demand-Side Offer). Por otro lado, el aumento de la carga debido al fenómeno de retorno que se produce después de los periodos de actuación requeridos por el Operador del Mercado (ISO), han de gestionarse y tratarse como una operación de compra adicional de energía a precios del mercado a tiempo real.

Volviendo al esquema mostrado en la figura 9 del capítulo 2, en los puntos 5 y 6 de dicho esquema quedan recogidas las últimas fases de la estructura de un sistema de participación de la demanda. Por tanto, la aportación de este capítulo se basa en estas dos últimas fases, la agregación de la demanda y la definición y evaluación de paquetes (programas de ahorro, ofertas, etc...) para la demanda, ver Figura 89.

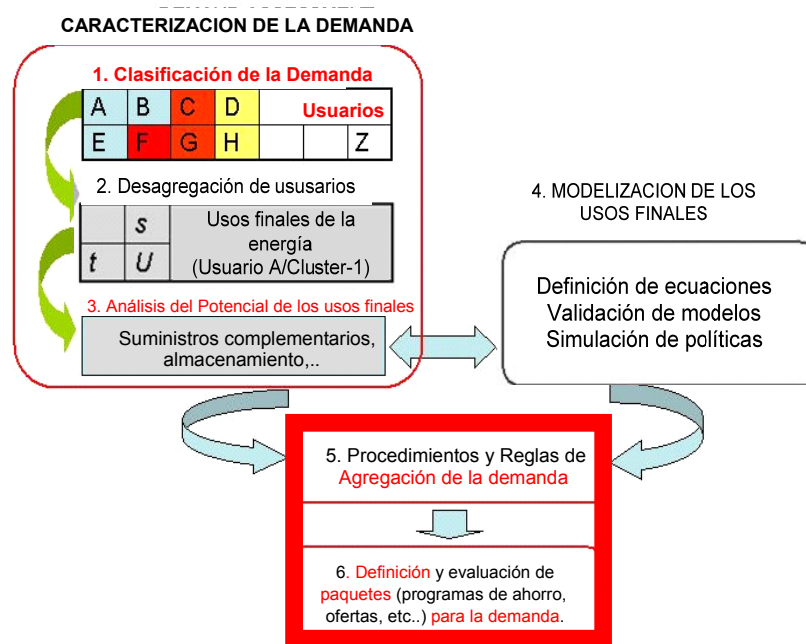


Figura 89. Aportaciones de este capítulo al esquema de la estructura del sistema de participación de la demanda.

Se ha utilizado un coeficiente para la elección de la mejor política para cada consumidor, considerando una determinada curva de precios y una determinada estrategia o política DER. Este coeficiente fue el producto de las reducciones de energía (puntos positivos de la curva, ver la Figura 87), o aumentos del consumo de energía (puntos negativos de la curva) por los precios de la energía en el mercado de operación a Tiempo Real. Por tanto, este producto tendrá para cada punto de la curva un valor positivo o negativo. La suma total nos proporciona el beneficio de la política para un día si el resultado global o suma es positivo. Sumando los coeficientes más altos se obtendrá por tanto el mayor beneficio.

Para validar esto, se realizó un entrenamiento de un mapa SOM con las 7 estrategias de reducción para cada consumidor, y se obtuvieron diferentes ejemplos de clusterización. En la Figura 90, se muestran estos mapas para los consumidores 13 y 24.

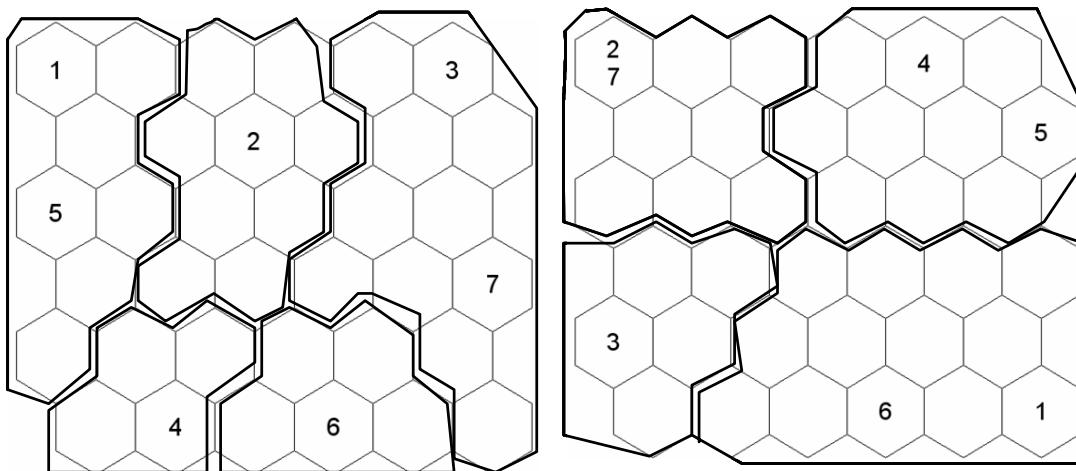


Figura 90. Mapas SOM entrenados con las estrategias DER (Consumidores: Universidad y Centro Comercial)

Después cada uno de los mapas entrenados son testeados con las curvas de precios de la ZONA 2 (esta zona fue elegida como la zona de días con precios más altos, que hacía pensar que el operador del mercado -ISO- requeriría a los consumidores para que redujeran su consumo, ver los puntos anteriores), con el objetivo de encontrar la estrategias DER más beneficiosas para cada consumidor.

Para introducir los datos en la red SOM se consideraron los siguientes pasos. En un primer lugar, las curvas de precio en tiempo real y las reducciones de energía obtenidas (cents./kWh.; kW.), y en segundo lugar, los valores por unidad de ambas curvas, tomando como base los valores máximos de ambas magnitudes (precio máximo de la zona 2 y valor máximo de pico de la reducción). Los primeros resultados obtenidos no fueron muy buenos, principalmente porque la solución que ofrecía el mapa SOM tras el entrenamiento no era la mejor desde el punto de vista económico.

Observando la forma de trabajar del algoritmo de Kohonen [54], quedó claro que la red (mapa entrenado) casa el vector de precio con el vector de reducción de energía con menor distancia entre ellos, no eligiendo o casando reducciones de formar similares pero con valores altos, lo cual desde un punto de vista económico es bastante más beneficioso.

La mejor solución fue escalar ambas curvas. Las curvas de precios fueron escaladas entre 0 y 1 tomando como base el valor más alto de los precios de todas las curvas de precios de los días de la zona 2. Las curvas de las estrategias de reducción fueron también escaladas utilizando como base el máximo pico de la reducción y multiplicando las curvas normalizadas de las estrategias por el valor medio de las curvas de precios normalizadas. Con este tratamiento las curvas de reducción con valores altos fueron localizadas cerca del promedio de los precios eliminando la desventaja producida por las diferentes escalas, permitiendo a la red casar de forma apropiada las curvas. No solamente teniendo en cuenta la forma sino el beneficio económico. Los resultados de estos test han sido presentados en la Tabla 22.

Es importante resaltar que las mejores alternativas han sido las combinaciones de las estrategias LTS y DG, además no aparecen exactamente las mismas para cada día. Las estrategias 4, 6 y 7 son las propuestas para ser aplicadas a la zona 2 de precios). La peor alternativa (WMU-Worst Matching Unit) que proporciona el Mapa SOM entrenado, para los hipotéticos participantes considerados, también aparece en la misma tabla. En estos casos y debido a los retornos de la energía el mapa SOM no considera esta opción.

Tabla 22. Las Mejores Y Peores Estrategias Seleccionadas Por El Mapa SOM (Consumidor: Universidad)

Day (mm/dd/yy)	BMU (Mejor neurona asociada)	Estrategia DER asociada con la BMU	WMU (Peor neurona asociada)	Estrategia DER asociada con la WMU
103001	24	6	14	2
111501	34	7	1	1
111601	12	4	13	2
121101	24	6	13	2
121201	18	6	14	2
121901	34	7	1	1
122001	34	7	1	1
122801	12	4	14	2
10202	30	6	3	5
10302	34	7	1	1
10702	18	6	14	2
13002	18	6	1	1

En la Tabla 23 se muestra la mejor estrategia de reducción para cada día de precios de la zona 2. El texto resaltado en negrita son los valores más altos que se han alcanzado resultado del producto de los precios en tiempo real por las curvas de reducción de la energía. En la misma tabla las celdas seleccionadas por la red SOM aparecen con un sombreado. Es posible apreciar fácilmente que ambas presentan similares resultados. A lo largo de la investigación fueron efectuados diferentes Test para cada uno de los clusters del espectro de consumidores, alcanzando índices de éxito en las elecciones (casamientos de curvas de precios con estrategias) realizadas por el mapa SOM (las conocidas como bmus- best matching Units) del orden del 95%. Recordar que mediante el, calcula la distancia euclídea entre los vectores entrenados y el nuevo vector presentado a la red.

De esta forma los mapas SOM permiten a los participantes, ya sean comercializadores o grandes consumidores, una rápida y fácil elección de las políticas DER a aplicar y cuales son las mejores y peores estrategias a desde un punto de vista económico, es decir, conocer las estrategias de mayor beneficio y las de mayor penalización.

Tabla 23. Comparación de los Resultados de la SOM

UNIVERSIDAD												
Días de la Zona 2												
Estrategias	10/30/01	11/15/01	11/16/01	11/12/01	12/12/01	12/19/01	12/20/01	12/28/01	02/01/02	03/01/02	07/01/02	1/30/02
LTS 1	183,0	85,2	404,6	296,7	469,6	-34,1	-151,2	781,1	99,7	142,5	524,4	215,2
LTS 2	8,4	163,0	340,5	95,0	131,7	362,6	533,5	457,8	6,0	374,0	170,5	152,4
DG	1.092,0	1.280,1	1.852,9	2.423,7	2.812,1	3.830,6	4.508,7	2.786,4	1.281,0	1.926,0	2.631,3	1.833,8
SC 1	<b>1.283,4</b>	<b>1.528,3</b>	<b>2.598,1</b>	<b>2.815,4</b>	<b>3.413,4</b>	4.159,1	4.891,0	<b>4.025,3</b>	<b>1.386,7</b>	<b>2.442,6</b>	<b>3.326,1</b>	<b>2.201,5</b>
SC 2	191,4	248,2	745,1	391,7	601,3	328,5	382,3	1.239,0	105,7	516,6	694,9	367,6
SC 3	1.275,0	1.365,3	2.257,6	2.720,4	3.281,6	3.796,5	4.357,4	3.567,5	1.380,7	2.068,5	3.155,6	2.049,0
SC 4	1.100,4	1.443,1	2.193,4	2.518,7	2.943,8	<b>4.193,2</b>	<b>5.042,2</b>	3.244,2	1.287,0	2.300,1	2.801,8	1.986,3

CENTRO COMERCIAL												
Días de la Zona 2												
Estrategias	10/30/01	11/15/01	11/16/01	11/12/01	12/12/01	12/19/01	12/20/01	12/28/01	02/01/02	03/01/02	07/01/02	1/30/02
LTS 1	920,5	2.013,7	3.323,5	1.383,4	2.334,2	2.846,3	2.576,7	4.314,8	1.179,3	2.550,2	3.786,9	2.357,4
LTS 2	-2.148,2	-2.119,3	-1.416,8	-3.116,6	-3.085,4	-2.361,1	-2.688,7	-2.021,5	-1.993,2	-1.456,3	-6.117,7	-3.727,9
DG	301,9	355,9	503,7	675,7	756,9	1.029,2	1.245,1	699,6	339,0	524,2	678,5	491,6
SC 1	-925,8	250,3	2.410,3	-1.057,4	5,7	1.514,4	1.133,1	2.992,9	-474,9	1.618,1	-1.652,2	-878,9
SC 2	-1.227,7	-105,6	1.906,6	-1.733,1	-751,2	485,2	-112,0	2.293,3	-813,9	1.093,9	-2.330,8	-1.370,5
SC 3	<b>1.222,4</b>	<b>2.369,6</b>	<b>3.827,1</b>	<b>2.059,1</b>	<b>3.091,1</b>	<b>3.875,6</b>	<b>3.821,8</b>	<b>5.014,3</b>	<b>1.518,2</b>	<b>3.074,4</b>	<b>4.465,5</b>	<b>2.849,0</b>
SC 4	-1.846,3	-1.763,4	-913,1	-2.440,9	-2.328,5	-1.331,9	-1.443,6	-1.321,9	-1.654,2	-932,1	-5.439,2	-3.236,4

CT RESIDENCIAL												
Días de la zona 2												
Estrategias	10/30/01	11/15/01	11/16/01	11/12/01	12/12/01	12/19/01	12/20/01	12/28/01	02/01/02	03/01/02	07/01/02	1/30/02
LTS 1	231,4	743,0	1.532,8	614,0	985,0	1.878,1	2.124,8	2.244,1	512,4	1.451,3	669,1	677,0
LTS 2	-67,0	144,0	417,4	-28,0	23,5	287,7	426,5	398,4	-18,5	334,0	-198,1	-20,2
DG	325,1	367,5	548,2	720,2	765,5	860,9	1.071,0	727,3	296,3	480,1	630,6	447,5
SC 1	489,6	<b>1.254,5</b>	<b>2.498,4</b>	1.306,2	<b>1.773,9</b>	<b>3.026,6</b>	<b>3.622,3</b>	<b>3.369,9</b>	790,2	<b>2.265,4</b>	1.101,7	1.104,3
SC 2	164,5	887,0	1.950,2	586,0	1.008,4	2.165,8	2.551,3	2.642,6	493,8	1.785,3	471,1	656,8
SC 3	<b>556,5</b>	1.110,5	2.081,0	<b>1.334,2</b>	1.750,4	2.738,9	3.195,8	2.971,5	<b>808,7</b>	1.931,4	<b>1.299,7</b>	<b>1.124,5</b>
SC 4	258,2	511,6	965,7	692,2	788,9	1.148,5	1.497,5	1.125,8	277,8	814,1	432,6	427,3

## 7.6. CONCLUSIONES

Como conclusión final, aquellos consumidores que encuentren dificultad para participar directamente en el mercado eléctrico debido a su complejidad, podrían beneficiarse de las operaciones de compra-venta de energía en los mercados, con la ayuda de un comercializador o un agregador, el cual actuaría en este caso como participante e intermediario en los programas de respuesta de la carga en el operador del mercado (ISO).

Mediante esta propuesta de integración de herramientas tales como los mapas SOM y los métodos de modelización de cargas basados en modelos físicos es posible saber cuando y como producir las reducciones (según patrones), de esta forma agentes (terceros) podrían identificar en los mercados de la energía, diario y a tiempo real, los grupos de consumidores y las estrategias DER con mayores posibilidades de obtener grandes beneficios dentro de las operaciones de compra-venta de la energía en dichos mercados.

Los participantes de estos productos (agentes) deberían notificar y ayudar a sus consumidores en la gestión de la demanda (consumo) y de la autogeneración (según cada caso) con el fin de maximizar los beneficios o minimizar costes. De esta forma, los participantes deberían cumplir con los recortes de consumos requeridos por el operador del mercado (ISO) y evitar en lo posible penalizaciones o multas, es decir, estos agentes (terceros) participantes han de elegir la cantidad de demanda agregada necesaria para alcanzar la carga (energía) requerida por el ISO, pero sin reducir de más, debido principalmente a los periodos de retorno o recuperación. En este sentido una integración o combinación de opciones ofrecidas por las partes demandantes y opciones ofrecidas por las partes ofertantes en los mercados, permitiría obtener grandes beneficios y como consecuencia una mejor operación del mercado. Al mismo tiempo los consumidores podrían tener la oportunidad de gestionar sus costes, gracias a la ayuda de los agregadores y comercializadores como puente a una participación efectiva de la demanda en mercados energéticos. Esta metodología ayudaría a los agentes ofertantes



(productores y distribuidores) a mejorar sus competencias y servicios mediante un amplio abanico de productos.



# CAPITULO 8

---

Conclusiones finales y trabajos futuros

*Mh Miguel  
Hernández*



## 8.1. APORTACIONES

Las conclusiones finales, fruto de esta investigación, han sido recogidas a lo largo de los capítulos que forman esta Tesis. No obstante, a modo resumen destacaremos las aportaciones más importantes:

1. Identificación de las posibilidades de participación de la demanda (de medianos y pequeños consumidores) en entornos energéticos liberalizados, no solo directamente debido a las dificultades que plantea el mercado, sino a través de empresas comercializadoras capaces de ofrecer programas de participación.
2. La definición de una metodología que permite obtener segmentos de demanda eléctrica (grupos de consumidores) con patrones y comportamientos de consumo similares, siendo los Mapas Autoorganizados de Kohonen una de las herramientas más eficaces para este fin. Dentro de este campo y como ampliación de la utilidad de los Mapas SOM también se ha contrastado la capacidad de los mismos para filtrar diferentes patrones, usos diferentes con el tiempo o patrones anómalos de consumo de uno o de varios consumidores. Los mapas de etiquetas obtenidos tras entrenar varios consumidores sin filtrado previo de los datos de entrada mostraron que existe posibilidad de localización de curvas de carga de un consumidor dentro del cluster o región de otro consumidor con un perfil de consumo totalmente distinto al patrón o perfil de carga normal del consumidor en cuestión. Tras realizar un filtrado de días festivos y/o anómalos, los nuevos mapas obtenidos mostraron una agrupación y localización más precisa en el mapa de etiquetas de los perfiles agrupados según el comportamiento del consumidor sin llegar a mezclar etiquetas. El objetivo de esta herramienta es ayudar, por un lado a los consumidores en la toma de decisiones y por otro a las empresas comercializadoras o agregadoras en el desarrollo de políticas y programas de participación de la demanda en los mercados liberalizados.

3. Se ha medido la sensibilidad de los Mapas SOM frente al tratamiento de los datos utilizados como entrada a la red. En un primer comienzo de la investigación se realizaron ensayos con datos de entrada tratados como curvas de carga normalizadas. Este tratamiento proporcionó buenos resultados cuando los datos de entrada pertenecían a días laborales, sin incluir fines de semana. Posteriormente la investigación se centró en realizar entrenamientos de los mapas con datos de entrada convertidos al dominio de la frecuencia aplicando a los mismos la Transformada de Fourier Rápida (Fast Fourier Transformation). Estos ensayos permitieron obtener importantes conclusiones: la primera de ellas afecta a la posibilidad de incluir en cada vector de entrada más información que la curva de carga de los días laborales. Es decir, es posible incluir información a cerca del comportamiento de consumo durante los fines de semana y durante varias semanas seguidas. La segunda conclusión obtenida de los resultados de los mapas entrenados es que mediante vectores de armónicos el mapa SOM es capaz de agrupar o separar consumidores con perfiles de consumo similares o distintos durante sus fines de semana aun teniendo los mismos patrones de consumo durante los días laborales de lunes a viernes. Por tanto, las principales ventajas han sido: una mejor y más clara clasificación en el dominio de la frecuencia, la inclusión de los fines de semana como información adicional frente al dominio del tiempo, reducción del número de datos a utilizar permitiendo una mayor velocidad de tratamiento de los mismos (ahorro de espacio y tamaño de las bases de datos). En definitiva, pasar los datos al dominio de la frecuencia ha permitido comprimir los datos sin perder información para entrenar la red y ha aumentado la rapidez de entrenamiento.
  
4. Se ha desarrollado una metodología para facilitar la toma de decisiones de los consumidores o agregadores con vistas a su participación en mercados de energía eléctrica, y que se puede extender a otros mercados. Debido a la dificultad de funcionamiento del mercado los pequeños y medianos

consumidores podrían beneficiarse de las operaciones de compra-venta de energía a través de programas ofrecidos por comercializadores y agregadores. Gracias a la aportación de los mapas SOM como herramienta de segmentación e identificación de consumidores eléctricos y a la utilización de modelos de carga basados en elementos físicos, los comercializadores pueden identificar en los mercados de compra-venta de energía (diario, intradiario, tiempo real, etc..) los segmentos de consumidores y las estrategias DER con mayores posibilidades de obtener beneficios dentro de las operaciones de compra-venta de energía. Ambos, agentes comercializadores y consumidores, obtendrían importantes ventajas. La Figura 91 muestra la integración de consumidores y agentes comercializadores en el mercado eléctrico mediante el desarrollo y uso de esta metodología.



Figura 91. Integración de consumidores y agentes comercializadores en el mercado eléctrico mediante el desarrollo y uso de esta metodología.

## 8.2. LÍNEAS FUTURAS

Dentro de los posibles trabajos futuros estaría la comparación en mayor profundidad de los resultados obtenidos con los mapas SOM con otras técnicas estadísticas, redes neuronales y técnicas de lógica Fuzzy con vistas a generar técnicas híbridas de segmentación e identificación.

Otra de las futuras líneas de investigación sería la utilización de los mapas SOM en la predicción de curvas de demanda o de precios. En este sentido ya existen investigaciones realizadas en el campo de la predicción de curvas de demanda utilizando redes neuronales.

El problema principal en la predicción de series temporales mediante redes neuronales es calcular el valor de unas determinadas variables no conocidas a priori en función de los valores pasados que tuvieron estas y en función de otras variables que puedan afectar a la que se quiere predecir.

La aplicación de las redes neuronales a problemas de predicción se puede encontrar en una gran variedad de campos, como la predicción en los mercados de valores, predicción de ventas en empresas, predicción de bajas de una compañía de seguros etc.. También se pueden encontrar en el campo de la electricidad aplicaciones de redes neuronales a la predicción. Han sido varios los modelos de Redes Neuronales empleados en la predicción de la demanda, la mayoría de ellos basados en las Redes MLP (Multi Layer Perceptron), pero también se han utilizado otros modelos de redes como las de aprendizaje no supervisado, redes recurrentes, e incluso técnicas de lógica difusa conjuntamente con redes neuronales.

Vermaak y Botha [134] realizaron un estudio comparativo, para la predicción de la demanda eléctrica, mediante el empleo de una red MLP y una red de tipo recurrente. El estudio demostró que la red recurrente, aún con su menor número de entradas, daba resultados tan buenos como la red MLP. En este sentido se utilizó un tamaño de patrón

de entrenamiento de 3 semanas, que corresponden a las 3 semanas previas a la fecha de predicción. De acuerdo con esto, conforme se va variando la fecha a predecir se varía el contenido del patrón de entrenamiento, esta técnica es conocida como “*moving window strategy*”.

Por otro lado, los estudios realizados por Taylor y Buizza [135] le da una gran importancia a las condiciones climatológicas a la hora de realizar la predicción de la demanda eléctrica. Consideran las condiciones climatológicas como un aspecto clave a tener en cuenta en la entrada del sistema cuando se emplean redes neuronales artificiales para la predicción de carga.

Siguiendo esta línea de investigación [136] crearon modelos para la predicción de la energía y de los picos de carga mediante redes neuronales. Utilizaba una red MLP entrenada con el algoritmo BP y toma en cuenta para el entrenamiento de la red los datos de un histórico de carga, variables climatológicas y variables que reflejan la variación económica. En este caso los factores que reflejaban las variaciones climatológicas y económicas son seleccionados mediante unos coeficientes de correlación.

La utilización de los mapas SOM en este campo de la predicción se ha basado hasta el momento en dos líneas, principalmente. Por un lado se han utilizado estos mapas para realizar una preclasificación de curvas de carga, con la finalidad de obtener conjuntos de curvas agrupadas para ser utilizados como entrada en otro tipo de técnica, esta sí, utilizada como herramienta de predicción. En este sentido nos encontramos nuevamente con el uso de redes neuronales del tipo MLP [115] u otras técnicas como el SVM (Support vector machine) [137], [138]. Por otro lado, la utilización de modelos basados en mapas SOM como herramientas de predicción, han empleado como datos de entrada los datos de carga pasados (históricos de demanda). En este sentido no hace falta emplear ni datos de temperatura, ni códigos indicativos del tipo de día y hora que se quiere predecir. Este modelo, utilizado en investigaciones anteriores [139], se basa en que los valores de la curva de carga que se va a predecir ya se han dado de forma



aproximada en un pasado, es decir, probablemente si buscamos en el histórico de datos de carga pasados encontraremos una curva que se aproxime a la curva futura que se quiere predecir.

Sin embargo, uno de los trabajos futuros sería analizar la forma de introducir al modelo basado en una red SOM la dependencia existente entre la demanda eléctrica y otros factores como la temperatura, variables económicas, etc.. También sería conveniente investigar como afectarían estos nuevos factores a los resultados de los mapas SOM comparando los resultados con otras técnicas. Por último, destacaría la creación de modelos híbridos de predicción utilizando los mapas SOM y otras herramientas. Siguiendo estas líneas ya se han realizado algunos primeros estudios destacando los presentados en el año 2005 y 2006 por Fan S. y Chen L.N. [140] y [141].



# BIBLIOGRAFIA

---





## 9.1. BIBLIOGRAFÍA

- [1] E.Hirst, "Reability benefits of price-responsive demand," *Oak Ridge*, 2002.
- [2] E.Hirst, "The financial and physical insurance benefits of price-Responsive demand," *The electric journal*, vol. 15(4), pp. 66-73, 2002.
- [3] E.Hirst and B.Kirby, *Retail load Participation in competitive wholesale electricity markets*. Washington.D.C, 2001.
- [4] J. A. Jardini, C. M. V. Tahan, M. R. Gouvea, S. U. Ahn, and F. M. Figueiredo, "Daily load profiles for residential, commercial and industrial low voltage consumers," *Ieee Transactions On Power Delivery*, vol. 15, pp. 375-380, 2000.
- [5] A. Molina, A. Gabaldon, J. A. Fuentes, and C. Alvarez, "Implementation and Assessment of Physically Based Load Models: Application to Direct Load Control Residential Programmes," presented at IEE Proceedings Transmission and Distribution, 2003.
- [6] E. Hirst, R. Goeltz, and B. Bronfman, "Comparisons Among Participants And Nonparticipants In The Bonneville-Power-Administration Residential Weatherization Pilot Program," *Journal Of Environmental Systems*, vol. 14, pp. 103-114, 1984.
- [7] E. Hirst and J. Jackson, "Energy Demand And Conservation In United-States Residential And Commercial Buildings - Impact Of The United-States National Energy Plan," *Energy Policy*, vol. 7, pp. 208-214, 1979.
- [8] B. D. Pitt and D. S. Kirschen, "Application of data mining techniques to load profiling," in *Proceedings Of The 21st International Conference On Power Industry Computer Applications*, 1999, pp. 131-136.
- [9] E. M. Ruiz, "Data Mining: Técnicas y Herramientas," October 2001.
- [10] C. Olaru and L. Wehenkel, "Data Mining tutorial," *IEEE Computer Applications in Power*, vol. vol.12, pp. 19-25, 1999.
- [11] M. Daneshdoost, M. Lotfalian, G. Bumroonggit, and J. P. Ngoy, "Neural network with fuzzy set-based classification for short-term load forecasting," *Ieee Transactions On Power Systems*, vol. 13, pp. 1386-1391, 1998.
- [12] R. Chicco, F. Napoli, P. Piglione, Postolache, M.Scutariu, and C.Toader, "Load Pattern-Based Classification of Electricity Customers," *IEEE Trans. On Power Systems*, vol. 19, pp. 1232-1238, 2004.
- [13] S. Valero, M. Ortiz, F. García, and A. Gabaldón, "Application of Self-Organizing Maps for Classification and Filtering of Electrical Customer Load Patterns," presented at 7th IASTED International Conference on Power and Energy Systems, Florida, USA, 2004.
- [14] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. M. Toader, "Emergent Electricity Customer Classification," *IEE Proc. -Generation Transm. Distribution*, vol. Vol.152, pp. pp. 164-172, 2005.
- [15] E. Hirst, "Electric Utilities - Demand-Side Management Programs," *Energy Policy*, vol. 17, pp. 69-72, 1989.
- [16] E. Hirst, "Effects Of Utility Demand-Side Management Programs On Uncertainty," *Resource And Energy Economics*, vol. 16, pp. 25-45, 1994.
- [17] K. W. Cheung, P. Shamsollahi, D. Sun, J. Milligan, and M. Potishnak, "Energy and ancillary service dispatch for the interim ISO New England Electricity Market," *Ieee Transactions On Power Systems*, vol. 15, pp. 968-974, 2000.
- [18] L. J. Hill and E. Hirst, "Demand-Side Management - Research Opportunities For Electric Utilities," *Electric Power Systems Research*, vol. 8, pp. 137-142, 1985.
- [19] F. Schweppe, M. Caraminis, R. Bohn, and R. Tabors, "Spot pricing of electricity," 1998.
- [20] *"Electricity market reform. An IEA handbook"*: International Energy Agency, 1999.
- [21] L. J. S. Tembleque, *Nuevos marcos reguladores para la energía. Referencias internacionales: La nueva regulación eléctrica*, CNE ed, 2000.
- [22] S. Hunt and G. Shuttleworth, "Unlocking the grid," *IEEE Spectrum*, pp. pp. 20-25, July 1996.
- [23] *"Electricity sector deregulation in the APEC region"*: Asia Pacific Energy Research Centre, March 2001.

- [24] "Spanish Electricity Market Operator (OMEL)," [http://www.omel.es/frame/es/resultados\\_index.htm](http://www.omel.es/frame/es/resultados_index.htm).
- [25] "The Lessons of Privatisation," in *Business Review Weekly*, 1-03-1999, pp. p. 48.
- [26] G. English, "Putting Consumers First," *IEEE Spectrum*, pp. 19, 06-2001.
- [27] R. Rudolph and S. Ridley, *Power Struggle: The Hundred-Year War over Electricity*. New York: Harper & Row, 1986.
- [28] EIA, "Privatization and the Globalization of Energy Markets," US Dept. of Energy, <http://www.eia.doe.gov/emeu/pgem/p.v.>, 10-1996.
- [29] EUA, "Opening up to Choice: The single Electricity Market," 1999.
- [30] V. Mallet, "France Goes on Sale," in *Financial Times*, 18-06-2002, pp. 22.
- [31] Comisión Europea, "First Benchmarking Report on the Implmmentation of the Internal Electricity and Gas Market," 2002, pp. 3 & 19.
- [32] Agencia Internacional de la Energía, "Reforma al mercado de electricidad: California y después," Marzo del 2005, pp. 5.
- [33] <http://www.iec.ch/index.html>, "IEC Standards," International Electrotechnical Commission.
- [34] MITC, "Ley 54/1997 del sector eléctrico," 1997.
- [35] MITC, "Real Decreto 1955/2000, de 1 de diciembre, por el que se regulan las actividades de transporte, distribución, comercialización, suministro y procedimientos de autorización de instalaciones de energía eléctrica. Dedicar el Capítulo II, del Título VI, a establecer el contenido y extensión de la calidad del suministro eléctrico.," 2000.
- [36] C. Europea, "[http://ec.europa.eu/comm/off/green/index\\_es.htm](http://ec.europa.eu/comm/off/green/index_es.htm)."
- [37] C. Europea, "[http://ec.europa.eu/comm/off/white/index\\_es.htm](http://ec.europa.eu/comm/off/white/index_es.htm)."
- [38] Energy Information Administration Department of Energy, "The Changing Structure of the Electric Power Industry 2000, An Update," DOE/EIA-0562 (00), October 2000.
- [39] B. Lescoeur; J.B. Galland; E. Husson., ""Marginal cost pricing: an efficient tool to ensure electricity demand side management. Demad- side management and electricity end-use efficiency"," pp. pp. 191-205, 1998.
- [40] C.W. Gellings, "Bibliography on load management," *IEEE Trans. On Power Apparatus Systems*, vol. PAS-100, no. 5, pp. 2597-2601, 1981.
- [41] M.L. Chan; F. Albuyeh, "Integrating Load Management Into Energy Management Systems Normal Operations," *IEEE Trans. On Power Systems*, vol. 1, no. 4, pp. 152-157, 1986.
- [42] D. R. Limaye, "Implementation of Demand-side Management Programs," *IEEE Proceedings*, vol. 73, no. 10, pp. 1503-1511, 1985.
- [43] H. Nilsson, "The many faces of demand-side management," *Power Engineering Journal*, vol. 85, pp. 207-210, 1994.
- [44] N. Redondo, "Metodología de Diseño y Evaluación de Productos de Respuesta de la Demanda en Mercados Eléctricos Competitivos." Valencia, 2006.
- [45] R. Luff, "Economic Principles of Demand Response in Electricity," Edison Electric Institute, October 2002.
- [46] I.E. Agency, "The Power to Choose: Demand Response in Liberalised Electricity Markets," OECD Energy- Organisation for Economic Co-operation and Development, Octubre 2003.
- [47] S. Borestein; M. Jaske; and A. Rosenfeld, "Dynamic Pricing, Advanced Metering and Demand Response in Electricity Markets." University of California Energy Institute: Center for the Study of Energy Markets, working paper series, October 2002.
- [48] C. E. Commission", "Integrated Energy Policy Report," vol. 2003: California Energy Commission.
- [49] J. Nazarko and Z. A. Styczynski, "Application of Statistical and Neuronal Approaches to the Daily Load Profiles Modelling in Power Distribution Systems," presented at Proc. IEEE Transm. and Distrib. Conference, New Orleans, 1999.
- [50] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. Toader, "A Review of Concepts and Techniques for Emergent Customer Categorization," 2002.
- [51] C. S. Chen, J. C. Hwang, and C. W. Huang, "Application of load survey systems to proper tariff design," *IEEE Transactions On Power Systems*, vol. 12, pp. 1746-1751, 1997.

- [52] C. S. Chen, J. C. Hwang, Y. M. Tzeng, C. W. Huang, and M. Y. Cho, "Determination of customer load characteristics by load survey system at taipower," *Ieee Transactions On Power Delivery*, vol. 11, pp. 1430-1436, 1996.
- [53] M.Sforma, "Data Mining in a Power Company Customer Database," presented at Electric Power Systems Research, 2000.
- [54] T.Kohonen, *Self-organisation and associative memory*, 3rd Ed. ed. Berlin, 1989.
- [55] T. Kohonen, "Self-organized formation of topologically correct feature maps," in *Biological cybernetics*, 1982, pp. 43:59-69.
- [56] B.D.Pitt and D.S.Kirschen, "Application of Data Minig Tecniques to Load Profiling," presented at IEEE PICA'99, Santa Clara, 1999.
- [57] D. Gerbec, S. Gasperic, I. Smon, and F. Gubina, "Consumers' load profile determination based on different classification methods," in *2003 Ieee Power Engineering Society General Meeting, Vols 1-4, Conference Proceedings*, 2003, pp. 990-995.
- [58] A. P. Birch and C. S. Ozveren, "An Adaptive Classification For Tariff Selection," in *Seventh International Conference On Metering Apparatus And Tariffs For Electricity Supply*, vol. 367, *Iee Conference Publications*, 1992, pp. 53-57.
- [59] C. S. Ozveren, C. Vechakanjana, and A. P. Birch, "Fuzzy classification of electrical load demand profiles - A case study," in *Fifth International Conference On Power System Management And Control, Iee Conference Publications*, 2002, pp. 353-358.
- [60] Y. H. Pao and D. J. Sobajic, "Combined Use Of Unsupervised And Supervised Learning For Dynamic Security Assessment," in *1991 Power Industry Computer Application Conference: Conference Paper*, 1991, pp. 278-284.
- [61] B. D. Pitt and D. S. Kirschen, "Application of Data Mining Techniques to Load Profiling," presented at Proc. IEEE PICA'99, Santa Clara, 1999.
- [62] A. Seppälä, "Statistical distribution of customer load profiles," *IEEE Catalogue*, vol. No. 95TH8130, 1995.
- [63] Carlson R.E.; Foley T.A., "The Parameter R2 in Multiquadratic Interpolation," *Computers Math. Applic.*, vol. 21, pp. 29-42, 1991.
- [64] W. Mendenhall; J.E. Freund, *Probabilidad y Estadísticas para Ingeniería y Ciencias.*: Prentice Hall, 1997.
- [65] Bland JM; Altman DG, *Statistic Notes: Transforming data*, BMJ ed, 1996.
- [66] DeShon R.P.; Alexander R.A., *Alternative Procedures for Testing Regression Slope Homogeneity when group error variances are unequal*, Psychological Methods ed, 1996.
- [67] Altman E. I., "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, 1968.
- [68] Huberty C. J., *Applied Discriminant Analysis*: Wiley. Interscience, 1994.
- [69] Yoon Y.; Swales G.; y Margavio T.H., "A Comparison of Discriminant Analysis versus Artificial Neural Networks," *Journal Operational Research Society*, vol. 44.
- [70] J. B. Kruskal and M. Wish, "Multidimensional Scaling," in *Sage Publications, Beverly Hills and London*.
- [71] D. A. Cole, S. E. Maxwell, R. Arvey, and E. Salas, "Multivariate Group Comparisons Of Variable Systems - Manova And Structural Equation Modeling," *Psychological Bulletin*, vol. 114, pp. 174-184, 1993.
- [72] R. J. Harris, "Multivariate analysis of variance," in *Applied analysis of variance in behavioral science. Statistics: Textbooks and monographs*, vol. vol.137, In L. K. Edwards (Ed.), Ed. New York, 1993, pp. 255-296.
- [73] J. Hartigan and M. Wong, "A k-means clustering algorithm," in *Applied Statistics*, 1979, pp. 100-108.
- [74] C. Chinrungrueng and C. Sequin, "Optimal adaptive k-means algorithm dynamic adjustment of learning rate," *IEEE transactions on Neural Networks*, pp. 6(1):157-169, 1995.
- [75] H. Andrade, T. Kurc, A. Sussman, and J. Saltz, "Decision tree construction for data mining on clusters of shared-memory multiprocessors," Technical Report CS-TR-4203 and UMIACS-TR.
- [76] A. Nazarko and Z. Styczynski, "Application of statistical and neural approaches to the daily load profiles modeling in power distribution systems," presented at Proc. IEEE Trans. Dist. Conf, N.orleans, 1999.

- [77] Kohonen T.; Kaski S.; Lagus K.K.; Salojärvi J.; Paatero V.; Saarela A., "Organization of a massive document collection," *IEEE Transactions on Neural Networks*, vol. 11, pp. 574-585, 2000.
- [78] R. F. Chang and C. N. Lu, "Load profile assignment of low voltage customers for power retail market applications," *IEE Proc. -Generation Transm. Distribution*, vol. Vol. 150, May. 2003.
- [79] B. Widrow, "Adaptive samplet-data systems- a statical theory of adaption," *WEscon*, Ed., 1959, pp. 4:74-85.
- [80] B. Widrow, "The original adaptative neural network broom balancer," in *International symposuim on cicuits and systems*, 1988, pp. 351-357.
- [81] Widrow B.; Winter R., "Neural nets for adaptive filtering and adaptive pattern recognition," *IEEE Computer*, pp. 25-39, marzo 1988.
- [82] D. Rumelhart and J. McClelland, *Parallel Distributed Processing.*, vol. 1: Foundations: MIT Press, 1986.
- [83] D. Rumelhart, G. Hinton, and R. J. Williams, "Learning representations by backpropagation errors," *Nature*, vol. 323, pp. 533-536, 1986.
- [84] García Villares J. L., Blasco J., Martín del Brío B., Domínguez J.A., J. Barquillas, I. Ramírez, and N. Medrano, "Short-term electric load-forecasting using ANN. Part II: Multilayer Perceptron for hourly electric-demand forecasting," presented at 14th IASTED Int. Conf. on Modeling, Identification and Control, Austria, 1995.
- [85] P. Werbos, "Backpropagation throught time: What it does and how to do it," presented at In proceedings of IEEE, 1990.
- [86] P. Werbos, "Backpropagation and neurocontrol: a review and propectus," presented at In international joint conference on neural networks, 1989.
- [87] H. Ritter, K. Obermayer, K. Schulten, and J. Rubner, "Self-Organizing Maps and Adaptive filters. En: Domany E., Van Hemmen J. L., Schulten K (eds). Models of Neural Networks.," Springer-Verlag, Berlín, Heidelberg, Nueva York, 1991, pp. 281-306.
- [88] T. Kohonen, "The Self-Organizing Map," *Proc. of the IEEE*, vol. 78, pp. 1464-1480, 1990.
- [89] J. Elman, "Finding structure in time," in *Cognitive science*, 1990, pp. 14:179-211.
- [90] V. M. Mladenov., A. C. Tsakoumis., and S. S. Vladov., "Electric load forecasting with multilayer percetron and elman neural network.," *IEEE Neurel*, 2002.
- [91] M. Cohen and S. Grossberg, "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks.," *IEEE Trans. Syst., Man and Cyb.*, 1983.
- [92] C. Von der Malsburg, "Self-organization of orientated sensitive cells in the striate cortex.," *Kybernetik*, vol. 14, pp. 85-100, 1973.
- [93] G. Carpenter and S. Grossberg, "The ART of Adaptive Pattern recognition by a Self-organizing neural network," *IEEE Computer*, pp. 77-88, marzo 1988.
- [94] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," in *Biological cyebrnetics*, 1980, pp. 193-202.
- [95] S. Reanals, "Radial basis function network for speech pattern classification," in *Electronics letters*, 1989, pp. 437-439.
- [96] P. D. Wasserman, in *Advanced Methods in Neural Computing*, V. N. Reinhold, Ed. New York, 1993, pp. 155-61 and 35-55.
- [97] D. Specht, "A general regression neural network," *IEEE Transactions on Neural Networks*, vol. 2, pp. 568-576, 1991.
- [98] D. F. Specht, "Probabilistic neural networks and the polynomial adaline as complementary techniques for classification," *IEEE Trans. on Neural Networks*, vol. vol.1, pp. 111-121, 1990.
- [99] J. J. DeGrujter, McBartney, A. B., "A Modified Fuzzy k means for Predictive Classification," presented at Classification and Related Methods of Data Analysis, Amsterdam, 1988.
- [100] J.-S. R. Jang, "ANFIS: Adaptive-network-based fuzzy inference system," *IEEE Trans. Systems, Man and Cybernetics*, vol. vol.23, 1993.
- [101] J. F. Leary, S. R. McLaughlin, L. N. Reece, C. Stewart, S. Stewart, and J. Smith, "Data mining by subtractive clustering," *Cytometry*, pp. 47-47, 2002.



- [102] J. N. Smith, T. Prow, L. Reece, A. Fontenot, E. Salazar, N. Wang, W. Rose, P. Szaniszló, and J. F. Leary, "Subtractive clustering analysis: A novel data mining method for finding cell subpopulations," *Cytometry Part A*, vol. 59A, pp. 140-140, 2004.
- [103] K. Demirli and P. Muthukumar, "Higher order fuzzy system identification using subtractive clustering," *Journal Of Intelligent & Fuzzy Systems*, vol. 9, pp. 129-158, 2000.
- [104] M. P. Windham, "Cluster Validity For The Fuzzy C-Means Clustering-Algorithm," *Ieee Transactions On Pattern Analysis And Machine Intelligence*, vol. 4, pp. 357-363, 1982.
- [105] J. C. Bezdek, R. Ehrlich, and W. Full, "Fcm - The Fuzzy C-Means Clustering-Algorithm," *Computers & Geosciences*, vol. 10, pp. 191-203, 1984.
- [106] R. L. Cannon, J. V. Dave, and J. C. Bezdek, "Efficient Implementation Of The Fuzzy C-Means Clustering Algorithms," *Ieee Transactions On Pattern Analysis And Machine Intelligence*, vol. 8, pp. 248-255, 1986.
- [107] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, and C. M. Toader, "Customer characterization options for improving the tariff offer," *Ieee Transactions On Power Systems*, vol. 18, pp. 381-387, 2003.
- [108] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Patt Anal. Machine Intell*, vol. vol. PAMI-1, pp. 224--227, 1979.
- [109] "<http://www.usyd.edu.au/su/agric/acpa/fkme/FkME.html>." The University of Sydney, Australia.
- [110] G. Chicco, R. Napoli, F. Piglion, P. Postolache, M. Scutariu, and C. Toader, "A review of Concepts and Techniques for Emergent Customer Categorisation," 2002.
- [111] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, 1981.
- [112] "Helsinki University of Technology. SOM toolbox for Matlab 5.0. Available: <http://www.cis.hut.fi/projects/somtoolbox/download>."
- [113] S. Valero, M. Ortiz, F. García, and A. Gabaldón, "Characterization and Identification of Electrical Customer Through the Use of SOM and Daily Load Parameters," presented at IEEE PSCE2004, New York, 2004.
- [114] G. Chicco, R. Napoli, and F. Piglion, "Load Pattern Clustering for Short-Term Load Forecasting of Anomalous Days," presented at 2001 IEEE Porto Power Tech Conference, Porto, Portugal.
- [115] R. Lamedica, A. Prudenzi, M. Sforna, M. Caciotta, and V. Orsolini Cencelli, "A Neural Network Based Technique for Short-Term Load Forecasting of Anomalous Load Periods," *IEEE Transactions on Power Systems*, vol. Vol. 11, pp. pp. 1749-1756, 1996.
- [116] "List of Nace codes. Gateway to the EU, [Online]. Available: [http://europa.eu.in/comm/competition/mergers/cases/index/nace\\_all.html](http://europa.eu.in/comm/competition/mergers/cases/index/nace_all.html)."
- [117] F. D. Galiana, "Identification of stochastic electric load models from physical data," *IEEE Tran. on Automatic Control*, vol. vol.19, pp. 887-893, 1974.
- [118] C. M. Huang and H.T. Yang, "Evolving wavelet-based networks for short-term load forecasting," presented at IEE Proc-Gener. Transm, 2001.
- [119] F. H. Mangano and A. Abur, "Fault location using Wavelets," *Ieee Transactions On Power Delivery*, vol. Vol. 13, pp. pp. 1475-1480, 1998.
- [120] E.C. Titchmarsh, *Introduction to the theory of Fourier integrals*: Oxford Univ. Press, 1948.
- [121] R. N. Bracewell, *The Fourier Transform & Its Applications*: Hardcover, Jun 8, 1999.
- [122] E. D. Brigham, *The fast Fourier transform*: Prentice-Hall, 1974.
- [123] R.W. Ramirez, *The FFT fundamentals and concepts*: Prentice-Hall, 1985.
- [124] H. T. Yang, S. C. Chen, and W. N. Tsai, "Classification of direct load control curves for performance evaluation," *Ieee Transactions On Power Systems*, vol. 19, pp. 811-817, 2004.
- [125] B. F. Hobbs, H. B. Rouse, and D. T. Hoog, "Measuring The Economic Value Of Demand-Side And Supply Resources In Integrated Resource Planning-Models," *Ieee Transactions On Power Systems*, vol. 8, pp. 979-987, 1993.
- [126] G. SW, ""Potential World Development throught 2100:the Impacts on Energy Demand Resources and the Enviroment," *WEC journal*, 1992.



- [127] L. B. N. L. Neenan Associates, "How and Why Customers Respond to Electricity Price Variability: A Study of NYISO and NYSEDA 2002 PRL Program Performance," LBNLReport 52209, Available: <http://eetd.lbl.gov/EA/EMP/drlm-pubs.html> January 2003.
- [128] ISO, "ISO New England Load Response Program Manual [On line], Available: [http://www.iso-ne.com/rules\\_proceeds/isone\\_mnls/index.html](http://www.iso-ne.com/rules_proceeds/isone_mnls/index.html)," 2005 JUNE.
- [129] N. Associatesy, "New York ISO Price-Responsive Load Program Evaluation: Executive Summar," *New York Independent sysytm operator*, 2002.
- [130] P. Cramton and R. Wilson, "A review of ISO New England's market rules," *Working paper*, 1998.
- [131] L. Cameron and P. Cramton, "The role of the ISO in U.S. elctricity markets: A review of restructuring in California and PJM," *Electricity Journal*, vol. 12(3), pp. 71-81, 1999.
- [132] C. Alvarez, A. Gabaldón, and A. Molina, "Assessment and Simulation of the Responsive Demand Potential in End-User Facilities," *IEEE Trans. Power Systems*, vol. vol 19, pp. 1223-1231, 2004.
- [133] M. Cohen, H. Franco, N. Morgan, D. Rumelhart, and V. Abrash, "Advances in neural information processing sytems," M. Kaufmann, Ed., 1993.
- [134] J. Vermaak and E. C. Botha, "Recurrent neural networks for short-term load forecasting," *IEEE Transactions On Power Systems*, vol. 13, pp. 126-132, 1998.
- [135] J. W. Taylor and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Transactions On Power Systems*, vol. 17, pp. 626-632, 2002.
- [136] S. C. Phimpachan, K. Kumhom, P. Jittiwangkul, N. Sangswang, A., "Energy and peak load forecast models using neural network for fast developing area," presented at Communications and Information Technology, 2004. ISCIT 2004. IEEE International Symposium on, Japan, 26-29 Oct 2004.
- [137] M. Martin-Merino and J. Roman, "Electricity load forecasting using self organizing maps," in *Artificial Neural Networks - Icnan 2006, Pt 2*, vol. 4132, *Lecture Notes In Computer Science*, 2006, pp. 709-716.
- [138] Z. J. Bao, D. Y. Pi, and Y. X. Sun, "Short-term load forecasting based on self-organizing map and support vector machine," in *Advances In Natural Computation, Pt 1, Proceedings*, vol. 3610, *Lecture Notes In Computer Science*, 2005, pp. 688-691.
- [139] T. G. Baumann, A.J., "Application of the Kohonen network to short-term load forecasting," presented at Neural Networks to Power Systems, 1993. ANNPS '93., Proceedings of the Second International Forum on Applications of, Yokohama, Japan, 19-22 Apr 1993.
- [140] S. Fan, C. X. Mao, and L. N. Chen, "Peak load forecasting using the self-organizing map," in *Advances In Neural Networks - Isnn 2005, Pt 3, Proceedings*, vol. 3498, *Lecture Notes In Computer Science*, 2005, pp. 640-647.
- [141] S. Fan and L. N. Chen, "Short-term load forecasting based on an adaptive hybrid method," *IEEE Transactions On Power Systems*, vol. 21, pp. 392-401, 2006.

# ANEXO

---





## 10.1. ARTICULOS PUBLICADOS EN REVISTAS CON INDICE DE IMPACTO


### IEEE, Transactions on Power Systems

Título del artículo: “*CLASSIFICATION, FILTERING, AND IDENTIFICATION OF ELECTRICAL CUSTOMER LOAD PATTERNS THROUGH THE USE OF SELF-ORGANIZING MAPS*”

Volumen: 21, número: 4. Páginas: 1672-11682 ISSN: **0885-8950**.


Fecha: Noviembre del 2006.



FACTOR DE IMPACTO: **0,951**

 **Journal Citation Reports®**


WELCOME HELP RETURN TO LIST PREVIOUS JOURNAL **2005 JCR Science Edition**

Journal: **IEEE TRANSACTIONS ON POWER SYSTEMS**

Mark	Journal Title	ISSN	Total Cites	Impact Factor	Immediacy Index	Articles	Cited Half-life	Citing Half-life
	<a href="#">IEEE T POWER SYST</a>	0885-8950	4165	<a href="#">0.951</a>	<a href="#">0.209</a>	239	<a href="#">8.3</a>	<a href="#">7.8</a>


[Cited Journal](#)  [Citing Journal](#)  [Source Data](#)

[CITED JOURNAL DATA](#) [CITING JOURNAL DATA](#) [IMPACT FACTOR TREND](#) [RELATED JOURNALS](#)

**Journal Information** 

**Full Journal Title:** IEEE TRANSACTIONS ON POWER SYSTEMS  
**ISO Abbrev. Title:** IEEE Trans. Power Syst.  
**JCR Abbrev. Title:** IEEE T POWER SYST  
**ISSN:** 0885-8950  
**Issues/Year:** 4  
**Language:** ENGLISH  
**Journal Country/Territory:** UNITED STATES  
**Publisher:** IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC  
**Publisher Address:** 445 HOES LANE, PISCATAWAY, NJ 08855  
**Subject Categories:** ENGINEERING, ELECTRICAL & ELECTRONIC [SCOPE NOTE](#)

[VIEW JOURNAL SUMMARY LIST](#) [VIEW CATEGORY DATA](#)

**Journal Impact Factor** 

Cites in 2005 to articles published in:	2004	160	Number of articles published in:	2004	261
	=			=	
		2003		2003	208
	=			=	
		Sum:	446	Sum:	469

Calculation:  $\frac{\text{Cites to recent articles}}{\text{Number of recent articles}} = \frac{446}{469} = \mathbf{0.951}$


## IET-Proceedings-Generation Transmission and Distribution

Título del artículo: "METHODS FOR CUSTOMER AND DEMAND RESPONSE POLICIES SELECTION IN NEW ELECTRICITY MARKETS"

Volumen: 1, NÚMERO: 1, Páginas: 104-110. ISSN: 1751-8687


Fecha: Enero del 2007



INDICE DE IMPACTO: 0,377

 **Journal Citation Reports®**

WELCOME ? HELP RETURN TO LIST 2005 JCR Science Edition

**Journal: IEE PROCEEDINGS-GENERATION TRANSMISSION AND DISTRIBUTION**

Mark	Journal Title	ISSN	Total Cites	Impact Factor	Immediacy Index	Articles	Cited Half-life	Citing Half-life
	<a href="#">IEE P-GENER TRANSM D</a>	1350-2360	766	<a href="#">0.377</a>	<a href="#">0.008</a>	124	<a href="#">8.6</a>	<a href="#">9.1</a>

[Cited Journal](#)  [Citing Journal](#)  [Source Data](#)

[CITED JOURNAL DATA](#)   
 [CITING JOURNAL DATA](#)   
 [IMPACT FACTOR TREND](#)   
 [RELATED JOURNALS](#)

**Journal Information** ⓘ

**Full Journal Title:** IEE PROCEEDINGS-GENERATION TRANSMISSION AND DISTRIBUTION  
**ISO Abbrev. Title:** IEE Proc.-Gener. Transm. Distrib.  
**JCR Abbrev. Title:** IEE P-GENER TRANSM D  
**ISSN:** 1350-2360  
**Issues/Year:** 6  
**Language:** ENGLISH  
**Journal Country/Territory:** ENGLAND  
**Publisher:** IEE-INST ELEC ENG  
**Publisher Address:** MICHAEL FARADAY HOUSE SIX HILLS WAY STEVENAGE, HERTFORD SG1 2AY, ENGLAND  
**Subject Categories:** ENGINEERING, ELECTRICAL & ELECTRONIC [SCOPE NOTE](#)

[VIEW JOURNAL SUMMARY LIST](#)   
 [VIEW CATEGORY DATA](#)

**Journal Impact Factor** ⓘ

Cites in 2005 to articles published in:	2004 = 22	Number of articles published in:	2004 = 105
	2003 = 59		2003 = 110
	Sum: 81		Sum: 215
Calculation: $\frac{\text{Cites to recent articles}}{\text{Number of recent articles}}$	$\frac{81}{215}$	=	<b>0.377</b>

## 10.2. ARTICULO ACEPTADO A PUBLICACIÓN OFICIALMENTE.

### International Journal of Power and Energy Systems

Título del artículo: "APPLICATION OF SELF-ORGANIZING MAPS FOR CLASSIFICATION AND FILTERING OF ELECTRICAL CUSTOMER LOAD PATTERNS"

Fecha de Aceptación: 20/04/2006



**ACTA Press**

A Scientific and Technical Publishing Company

---

• Home
• Login
• My Cart
• Reviewers Only
• FAQ
• Contact

- Find/Buy Articles
- Journals
- Proceedings
- Subscriptions
- Submissions
- Call for Papers

Publication Search:

[advanced search](#)

■ [Publishing Services](#)

### International Journal of Power and Energy Systems

2006 Issue

Editor-In-Chief(s): Dr. A. Domijan, Jr.  
Frequency: 3 issues per year



**Journal Information**

**Rates (USD):** \$365.00 (HardCopy); \$330.00 (Online)

**Individual Articles (USD):** \$25.00 (Online)

**ISSN:** 1078-3466

[Editorial Board and Indexing Information](#)

**Tasks**

**Description**

First published in 1972, this journal serves a worldwide readership of power and energy professionals. As one of the premier referred publications in the field, this journal strives to be the first to explore emerging energy issues, featuring only papers of the highest scientific merit. The subject areas of this journal include power transmission, distribution and generation, electric power quality, education, energy development, competition and regulation, power electronics, communication, electric machinery, power engineering systems, protection, reliability and security, energy management systems and supervisory control, economics, dispatching and scheduling, energy systems modelling and simulation, alternative energy sources, policy and planning. It appears three times per year.

Please choose:

**Add to Shopping Cart**

Abstracts and reference lists may contain minor errors and formatting inconsistencies. Please contact us if you have any concerns or questions.