



UNIVERSIDAD DE CANTABRIA

**ESCUELA TÉCNICA SUPERIOR DE INGENIEROS DE CAMINOS,
CANALES Y PUERTOS**

DEPARTAMENTO DE CIENCIAS Y TÉCNICAS DEL AGUA Y DEL
MEDIOAMBIENTE INSTITUTO DE HIDRÁULICA AMBIENTAL “IH
CANTABRIA”

Másteres oficiales en Hidráulica Ambiental. Curso 2008-2010

**TESINA DE MÁSTER UNIVERSITARIO EN
GESTIÓN INTEGRADA DE ZONAS COSTERAS**

DESARROLLO DE UN MODELO NEURODIFUSO (ANFIS) PARA LA
PREDICCIÓN DE LA SALINIDAD EN ZONAS AFECTADAS POR EL
VERTIDO DE PLANTAS DESALINIZADORAS

PRESENTADA POR:

Guiomar López Fernández

DIRIGIDA POR:

**Roberto Minguez
Pilar Palomar**

Santander, Abril de 2010

A mis padres.

A todos los que me ayudaron.

TABLA DE CONTENIDOS

1	INTRODUCCIÓN Y OBJETIVOS	10
1.1.	INTRODUCCIÓN	10
1.2.	OBJETIVOS	14
1.3.	ORGANIZACIÓN DEL DOCUMENTO.....	15
2	LA DESALINIZACIÓN Y SU PROBLEMÁTICA AMBIENTAL	17
2.2.	LA DESALINIZACIÓN. DEFINICIÓN Y TIPOS	20
2.3.	CARACTERIZACIÓN DEL VERTIDO DE DESALADORAS.....	22
2.4.	IMPACTO AMBIENTAL DE LA SALMUERA SOBRE EL MEDIO MARINO	24
2.5.	PLANES DE VIGILANCIA AMBIENTAL. NECESIDAD Y CONTENIDOS.....	29
3	REVISIÓN BIBLIOGRÁFICA DE LAS HERRAMIENTAS PREDICTIVAS PARA LA MODELIZACIÓN DEL VERTIDO DE LAS DESALINIZADORAS	34
3.1.	MODELOS FÍSICOS.....	35
3.2.	MODELOS DE ECUACIONES INTEGRADAS	37
3.3.	MODELOS HIDRODINÁMICOS	37
3.4.	SISTEMAS BASADOS EN INTELIGENCIA ARTIFICIAL	41
3.4.1.	<i>Redes neuronales</i>	41
3.4.2.	<i>Lógica difusa</i>	49
3.4.3.	<i>Sistemas neuro-difusos</i>	53
3.4.4.	<i>Comparación de los tres métodos</i>	57
4	METODOLOGÍA	59
4.1.	DESCRIPCIÓN DE LOS AGENTES DEL MEDIO FÍSICO Y DE EXPLOTACIÓN QUE PARTICIPAN EN EL MODO DE FALLO PRINCIPAL: SUPERACIÓN DE LOS UMBRALES DE SALINIDAD.....	59
4.2.	DESARROLLO DEL MODELO PREDICTIVO BASADO EN ANFIS	62
4.2.1.	<i>Selección de las variables de entrada</i>	64
4.2.2.	<i>División de los datos</i>	66
4.2.3.	<i>Transformación de las variables de entrada</i>	67
4.2.4.	<i>Calibración del modelo</i>	69
4.2.5.	<i>Validación y fiabilidad de las predicciones</i>	71

5 APLICACIÓN DE LA PROPUESTA A PLANTAS DESALADORAS REALES:	
PLANTAS I Y II DEL CANAL DE ALICANTE.....	72
5.1. INTRODUCCIÓN.....	72
5.2. DESCRIPCIÓN DEL ÁREA DE ESTUDIO.....	74
5.2.2. <i>Áreas y ecosistemas protegidos</i>	75
5.3. CARACTERÍSTICAS DEL EFLUENTE DE SALMUERA Y DEL SISTEMA DE VERTIDO	76
5.4. CARACTERIZACIÓN DEL CLIMA MARÍTIMO.....	78
5.4.1. <i>Oleaje</i>	79
5.4.2. <i>Nivel del mar</i>	86
5.5. RESULTADOS COMPUTACIONALES.....	88
5.5.1. <i>Datos disponibles</i>	88
5.5.2. <i>Selección de variables</i>	90
5.5.3. <i>Tratamiento de la información</i>	94
5.5.4. <i>Calibración</i>	97
5.5.5. <i>Validación</i>	107
5.5.6. <i>Generación de pronósticos</i>	108
6 CONCLUSIONES.....	112
6.1. CONCLUSIONES.....	112
6.2. FUTURAS LÍNEAS DE INVESTIGACIÓN	113
REFERENCIAS.....	116

LISTA DE FIGURAS

Figura 2.1. Localización de las actuaciones previstas en el Programa A.G.U.A. Fuente: ACUAMED.	19
Figura 2.2. Pradera de <i>Posidonia oceanica</i> . La Zohía, 14 m de profundidad.	27
Figura 2.3. Césped de <i>Cymodocea nodosa</i> . Calabardina, 5 m de profundidad.	29
Figura 3.1. Izq: Experimento de laboratorio para simular el vertido de salmuera procedente de una planta RO. Derch: Modelado con CORMIX.	38
Figura 3.2. Estructura de un sistema de inferencia difusa	51
Figura 3.3. Funciones de pertenencia de un sistema de inferencia difusa con dos inputs x_1 y x_2	52
Figura 3.4. Estructura en cinco capas de la arquitectura ANFIS.....	54
Figura 5.1. a) Plano del LIC. La línea roja marca el límite del LIC. b) Mapa de contornos batimétricos y situación del punto de descarga (v).	75
Figura 5.2. Funciones de distribución de mejor ajuste de a) Hs y b)Tp	82
Figura 5.3. Rosa de oleaje de la altura significativa	83
Figura 5.4. Excedencias sobre el umbral del cuartil 99% de Hs y Tp.....	84
Figura 5.5. a) Función de distribución de mejor ajuste para la velocidad del viento. b) Direcciones de procedencia del viento.	85
Figura 5.6. Excedencias sobre un umbral de la velocidad del viento.....	85
Figura 5.7. Series de nivel del mar (azul), marea astronómica (rojo) y residuo meteorológico (verde).....	87
Figura 5.8. Esquema lógico del sistema ASDECO. Documentación del proyecto, 2008.	88
Figura 5.9. Posición de las estaciones de medida SADO I y SADO II.....	89
Figura 5.10. Resultados del método selección "hacia delante".	92
Figura 5.11. Funciones de a) Auto-correlación, ACF (el eje negativo no se muestra puesto que los resultados son simétricos a ambos lados) y b) Auto-correlación parcial de la salinidad.	93
Figura 5.12. Serie temporal de altura máxima de ola empleada. En la tabla se muestran los estadísticos discretos descriptivos de la variable y en la figura inferior la función de distribución de mejor ajuste.....	94

Figura 5.13. Serie temporal parámetro de sales empleada. En la tabla se muestran los estadísticos discretos descriptivos de la variable y en la figura inferior la función de distribución de mejor ajuste.	95
Figura 5.14. Serie temporal de salinidad empleada. En la tabla se muestran los estadísticos discretos descriptivos de la variable y en la figura inferior la función de distribución de mejor ajuste.	95
Figura 5.15. Funciones de pertenencia de la variable Hmax. La campana de gauss marcada en rojo corresponde a la primera función de pertenencia, con centro $p_1=0.1217$	101
Figura 5.16. Estructura del modelo ANFIS generado.	102
Figura 5.17. Curvas de error RMSE.	103
Figura 5.18. Series temporales de salinidad medida y simulada por ANFIS	104
Figura 5.19. Recta de regresión entre valores medidos y simulados por ANFIS.	104
Figura 5.20. a) Residuos del modelo. b) Función de auto-correlación de los residuos. c) Función de correlación cruzada entre la altura de ola máxima y los residuos del modelo. d) Función de correlación cruzada entre el parámetro de sales y los residuos del modelo	107
Figura 5.21. Resultados de la validación de la red entrenada con el conjunto de datos de test.	108
Figura 5.22. Pronósticos de salinidad para horizontes de predicción de 24 h en tres días consecutivos.	110
Figura 5.23. Pronósticos de salinidad para horizontes de predicción de 24 h (21/02/10, 0:00h -21/02/10, 23:00)	111

LISTA DE TABLAS

Tabla 1.1. Aplicación de sistemas de inteligencia artificial a estudios de calidad de agua e hidrología. (¹ La referencia proviene de Maier et al., 2000.).....	12
Tabla 2.2. Principales tecnologías de desalación.....	20
Tabla 2.3. Propiedades típicas del efluente de salmuera de las plantas desaladoras de osmosis inversa. Adaptado de S. Lattemann, T.Höpner, 2008.	23
Tabla 3.1. Diluciones en campo cercano obtenidas a partir de ensayos modelos 2D y 3D y medidas realizadas en la naturaleza. Fuente: Ruiz Mateo, 2007.	36
Tabla 5.1. Composición del vertido en un muestreo realizado el 26/02/2004. U. Alicante. PVA, 2004.	77
Tabla 5.2. Tabla de estadísticos básicos de la variable Hs	83
Tabla 5.3. Valores estadísticos de los tres grupos de datos	98
Tabla 5.4. RMSE obtenido de la aplicación de los diferentes tipos de funciones de pertenencia	99
Tabla 5.5. RMSE obtenido con el método de agrupamiento substractivo para diferentes radios de influencia	100

1 INTRODUCCIÓN Y OBJETIVOS

En este primer capítulo, se pretende introducir el concepto de la predicción de valores futuros y su importancia en el campo de la desalación. Por otra parte, se indican los objetivos perseguidos en esta tesina y, finalmente se resumen los contenidos de la misma.

1.1. Introducción

La predicción de valores futuros es una técnica que juega un papel importante en muchos aspectos de la vida cotidiana, habiéndose convertido en los últimos años en un paso esencial de muchas actividades humanas, especialmente las económicas. Desde el punto de vista científico, la predicción constituye un proceso de búsqueda de la relación entre el valor actual de una variable y los valores pasados de la misma, o de otras variables que puedan influir sobre el valor de la primera. Así, según Amara y Salanik, 1972, la predicción de valores futuros incluye cualquier actividad que mejore el conocimiento de las consecuencias futuras de desarrollos y decisiones presentes.

A pesar del hecho de que la imprecisión es una parte inherente a cualquier proceso predictivo, las predicciones generadas por el mismo nos ayudan a planificar el futuro, tomar decisiones racionales y reducir el riesgo en una atmósfera de incertidumbre.

Los procedimientos predictivos pueden clasificarse en cualitativos y cuantitativos, dependiendo del grado en el que éstos empleen métodos matemáticos y estadísticos (Montgomery et al., 1990). Los métodos cualitativos se basan principalmente en el juicio experto, lo que conlleva una estimación subjetiva, mientras que los métodos cuantitativos se basan en técnicas matemáticas y estadísticas. Los métodos estadísticos consisten en el estudio de datos históricos para determinar los procesos subyacentes y en utilizar este conocimiento para extrapolar el proceso al futuro.

Han sido varios los métodos estadísticos desarrollados para tratar con diferentes problemas predictivos. Estos métodos incluyen: análisis de regresión, medias móviles, suavizado exponencial, métodos Box-Jenkins, etc., métodos tradicionales que son

apropiados si el problema es simple, o si tanto la variable dependiente como la de decisión están bien definidas. Sin embargo, estos métodos presentan diversas deficiencias y limitaciones. En primer lugar, los entornos de predicción se han vuelto más y más complejos y las relaciones entre la variable a predecir y los factores de decisión que la influyen no siempre pueden ser expresados por un modelo matemático. Y en segundo lugar, los entornos de predicción se encuentran en constante evolución, lo que hace que los límites de decisión cambien. Los métodos tradicionales carecen de los mecanismos para tratar estos cambios y por lo tanto no pueden aprender de ellos para adaptarse.

Recientemente, ha habido muchos intentos de paliar estas limitaciones mediante la aplicación de técnicas de inteligencia artificial o *soft computing*, que constituyen una aproximación innovadora para la construcción de sistemas computacionalmente inteligentes (Jang et al., 1997). Herramientas tales como las redes neuronales, la lógica difusa (Zadeh, 1965), los algoritmos genéticos, etc., han sido desarrolladas dentro este campo. La fortaleza de estas técnicas se deriva de su capacidad para inferir relaciones subyacentes complejas. Desde el punto de vista estadístico, estas herramientas son análogas a los modelos de regresión no paramétricos y no lineales (Refenes, 1995; Smith 1993). Se ha demostrado que el dominio no lineal puede ser descrito de forma más precisa con estas tecnologías que con los métodos estadísticos no lineales que han sido empleados como las principales técnicas de análisis durante las últimas décadas.

Estos sistemas inteligentes, que combinan conocimiento, técnicas y metodologías de varias fuentes, tratan de imitar el razonamiento humano y adaptarse para aprender de ejemplos. A la hora de afrontar problemas del mundo real desde el punto de vista computacional, suele ser ventajoso el empleo de varias técnicas diferentes de forma sinérgica, lo que resulta en la construcción de sistemas híbridos complementarios. Dentro de este grupo de sistemas, una de las técnicas más exitosas es la neuro-difusa, en la que las redes neuronales se encargan de reconocer patrones y aprender a partir de ejemplos, mientras que los sistemas de inferencia difusa incorporan el conocimiento humano.

Tanto redes neuronales, como lógica difusa y sistemas híbridos neuro-difusos se han venido empleando en numerosos estudios relacionados con la hidrología y la calidad de

aguas durante algo más de una década (Tabla 1.1). Los modelos desarrollados demuestran la capacidad de los sistemas basados en inteligencia artificial para predecir valores futuros de diferentes variables.

Autor (es) y año	Variable modelada	Sistema empleado
¹ Whitehead et al., 1997	Concentración algal	Red neuronal
¹ Yabunaka et al., 1997	Concentración algal	Red neuronal
¹ Maier y Dandy, 1997b	Conc. de cianobacterias	Red neuronal
¹ Lorrai y Sechi, 1995	Corriente	Red neuronal
¹ Fernando y Jayawardena, 1998	Corriente	Red neuronal
¹ Golob et al., 1998	Corriente	Red neuronal
¹ Allen y le Marshall, 1994	Precipitación	Red neuronal
¹ Hsu et al., 1997	Precipitación	Red neuronal
Luk et al., 2001	Precipitación	Red neuronal
¹ Loke et al., 1997	Escorrentía	Red neuronal
Gautam et al., 2001	Precipitación-escorrentía	Neuro-difuso
Khonder et al., 1998	Inundaciones	Red neuronal
¹ DeSilets et al., 1992	Salinidad	Red neuronal
Maier y Dandy, 1996b	Salinidad	Red neuronal
¹ Bastarache et al., 1997	Salinidad	Red neuronal
¹ Bastarache et al., 1997	pH	Red neuronal
¹ Bastarache et al., 1997	pH	Red neuronal
¹ Thirumalaiah y Deo, 1998b	Nivel de agua	Red neuronal
Chang y Chang, 2006	Nivel de agua	Neuro-difuso
Tutmez et al., 2006	Conductividad	Neuro-difuso
Sahoo et al., 2005	Conc. pesticidas	Neuro-difuso
Marcé et al., 2007	Oxígeno	Neuro-difuso
Lu y Lo, 2003	Calidad de agua	Neuro-difuso
Nayak et al., 2004	Series hidrológicas	Neuro-difuso

Tabla 1.1. Aplicación de sistemas de inteligencia artificial a estudios de calidad de agua e hidrología. (¹ La referencia proviene de Maier et al., 2000.)

El campo de la desalación no es ajeno a la necesidad de predicción de las variables involucradas en el proceso, particularmente la salinidad que se alcanza en el medio donde se vierten las aguas de rechazo generadas como resultado del proceso desalinizador. En España, la actual apuesta por la desalinización como alternativa para generar un agua en cantidad y calidad necesarias para satisfacer las actuales demandas de este recurso, supone un aumento significativo de la salmuera vertida al mar, donde el carácter estenohalino de algunas especies de alto valor ecológico, justifica la necesidad de disponer de herramientas que permitan predecir los valores de salinidad que serán alcanzados en el medio receptor.

En la actualidad, las Autorizaciones de Vertido y las Declaraciones de Impacto Ambiental de las plantas desalinizadoras, establecen límites de salinidad para especies de fanerógamas como *Posidonia oceanica* o *Cymodocea nodosa*. Para asegurar el cumplimiento de estos límites, y con carácter previo a la aprobación del proyecto, se modela el fenómeno de dispersión de la salmuera para tomar decisiones acerca del tipo de dispositivo de descarga más adecuado, según la dilución necesaria para cumplir con los límites establecidos. Esta modelación se realiza principalmente empleando modelos comerciales, más o menos sencillos, que resuelven las ecuaciones de continuidad, Conservación de la Cantidad de Movimiento y de Transporte. Sin embargo, estos modelos presentan ciertas limitaciones así como un alto grado de incertidumbre, y su utilidad se limita a la fase de diseño del proyecto, pero no a la de explotación. En esta fase de explotación deben cumplirse los objetivos establecidos en los Planes de Vigilancia designados tanto en las Declaraciones de Impacto Ambiental, como en las Autorizaciones de Vertido, para lo cual es de vital importancia el desarrollo de un modelo matemático que sirva para realizar predicciones fiables acerca de los valores de salinidad alcanzados en determinados puntos sensibles del medio.

Dentro de este contexto se desarrolla el proyecto ASDECO (Sistema Automático de Control del Vertido de Desaladoras), investigación subvencionada dentro del Programa Nacional de Ciencias y Tecnologías Medioambientales del Plan Nacional I+D+i 2004-2007. El objetivo principal del proyecto es diseñar, desarrollar y validar un prototipo que, analizando en tiempo real datos océano-meteorológicos del medio receptor y datos del efluente, permita aprovechar la capacidad asimiladora del medio marino –físico,

químico y biológico-, maximizando la eficiencia del sistema de vertido, sin perjuicio para los ecosistemas marinos.

ASDECO pretende proporcionar un sistema eficiente de la gestión del vertido respetuoso con el medioambiente informando a los gestores de la planta de las condiciones del medio receptor y de cómo este incide sobre la dilución del vertido para que, en caso de condiciones desfavorables, se pueda aplicar un protocolo de gestión del vertido diferente al habitual (aumento de la dilución, parada temporal del vertido, etc.). Además, permitirá a los gestores y agentes implicados cumplir con el programa de vigilancia ambiental requerido por las Declaraciones de Impacto Ambiental.

En el presente trabajo, se realizará un ejercicio de recapitulación de las herramientas de predicción de salinidad, así como un ensayo de la herramienta neuro-difusa ANFIS, con la finalidad de analizar su idoneidad para ser integrada en los desarrollos del proyecto ASDECO.

1.2. Objetivos

El objetivo fundamental de este trabajo ha sido el desarrollo de un modelo neuro-difuso, basado en la arquitectura ANFIS (Adaptative Neuro-Fuzzy System), para la predicción a corto plazo, de la salinidad alcanzada en un punto sensible de una zona afectada por el vertido de salmuera. La meta del trabajo es que el modelo sea capaz de alcanzar unas predicciones lo suficientemente fiables como para constituir una herramienta de apoyo a la gestión del vertido y poder ser integrado en los desarrollos del ya mencionado, proyecto ASDECO.

Más específicamente, los objetivos principales son:

- Desarrollar un modelo capaz de generar pronósticos de salinidad, basado en la arquitectura neuro-difusa ANFIS.
- Calibrar y validar el modelo desarrollado con datos procedentes de una zona afectada por el vertido de plantas desalinizadoras reales.

- Generar pronósticos empleando el modelo desarrollado, de forma que se pueda estudiar la idoneidad del mismo para su integración en los desarrollos del proyecto ASDECO.

1.3. Organización del documento

El trabajo que aquí se presenta, consta de seis capítulos. Este primero, constituye una introducción al resto del documento y presenta los principales objetivos perseguidos en el desarrollo de esta tesina.

En el Capítulo 2, se describe la técnica de la desalinización de agua de mar, la evolución de la misma, y la problemática ambiental que el proceso lleva asociada. Esta problemática, relacionada principalmente con el vertido hipersalino generado como resultado del proceso desalinizador, genera la necesidad del desarrollo de medidas de control y mitigación de los efectos perjudiciales generados por el proceso, objetivos que se persiguen con el establecimiento de Planes de Vigilancia Ambiental, cuyos contenidos se repasarán en la última parte del capítulo.

El Capítulo 3 corresponde a un repaso del estado del arte de la modelización del comportamiento de la salmuera en el medio marino, para la predicción de la salinidad alcanzada en zonas afectadas por este vertido. Se presenta un resumen de los modelos empleados en la actualidad para este propósito, así como de las técnicas de inteligencia artificial en las cuales se basará la herramienta que será presentada en capítulos posteriores.

El Capítulo 4 describe las metodologías más empleadas en el desarrollo de modelos basados en redes neuronales. Se repasan los puntos más importantes de cada una de las etapas, así como las opciones disponibles, extraídas de la literatura relacionada con la materia.

En el Capítulo 5 se presenta el caso de estudio. En primer lugar, se ponen antecedentes de las desalinizadoras en las cuales se basa el estudio (localización, áreas y ecosistemas afectados por la presencia de la obra, clima marítimo de la zona) y, por último, se exponen los resultados computacionales obtenidos tras la aplicación del modelo

propuesto al caso de estudio presentado. En este último punto, se utilizan algunas de las metodologías expuestas en el Capítulo 4 para el desarrollo del modelo y se presentan los resultados obtenidos.

Finalmente, en el Capítulo 6 se enumeran algunas de las conclusiones extraídas del trabajo desarrollado y se proporcionan recomendaciones para investigaciones futuras.

2 LA DESALINIZACIÓN Y SU PROBLEMÁTICA AMBIENTAL

La desalinización de agua de mar constituye, tanto una solución para paliar los problemas relacionados con la escasez de agua, como un problema medioambiental si nos referimos a los subproductos generados durante el proceso, entre los que destaca el vertido hipersalino (salmuera) que, generalmente es devuelto al mar. En este capítulo se repasarán varios puntos relacionados con esta técnica, desde la importancia que ha cobrado en los últimos años o las técnicas empleadas, hasta la problemática asociada al proceso.

2.1. Introducción

Uno de los problemas más serios de la humanidad en materia de abastecimiento de agua hoy en día, es la escasez de este recurso, limitado tanto en cantidad como en calidad apta para todo tipo de consumos.

La escasez de agua se refiere a los limitados recursos disponibles para cumplir con la demanda actual o proyectada en una localización específica. Dicha demanda está muy influenciada por el rápido crecimiento de la población mundial, que alcanzó los 6700 millones en 2007 y probablemente aumentará hasta 9200 millones en 2050, según estimaciones de las Naciones Unidas. Este rápido aumento, y el cambio producido en los hábitos de consumo de la población, han causado unas demandas sin precedentes en los recursos de agua disponibles, que acaban generando escasez y una intensa explotación de los recursos superficiales y subterráneos.

En cuanto a cantidad, los cinco países con mayor disponibilidad de recursos de agua en el mundo son Brasil, Rusia, Canadá, Estados Unidos, China e Indonesia, que en suma reciben 22.070 km³/año (Shiklomanov, 1999). En comparación con estos países, la totalidad de la Unión Europea recibe 171 km³/año. En el otro extremo, los países con menos recursos según la misma fuente son Chipre, Israel, Jordania, Singapur, Libia y Emiratos Árabes, cuya disponibilidad de agua únicamente alcanza los 4 km³/año.

En España, el agua también se reparte de una manera muy desigual, concentrándose el 70% de los recursos totales en el tercio norte peninsular. Los territorios insulares, con gran atractivo turístico, junto con la mayoría del litoral sur-levante, padecen un déficit crónico de agua que amenaza seriamente con estrangular su potencial desarrollo industrial, turístico y agrícola. En los periodos de sequía prolongada, la agricultura sufre pérdidas catastróficas y ni siquiera se pueden garantizar las demandas para el abastecimiento de la población.

Para encarar este problema existen tres líneas principales de actuación: aumentar la oferta, reducir la demanda y redistribuir los recursos (Ruiz Mateo, 2007). Entre las medidas orientadas a aumentar los recursos disponibles, la desalinización ha experimentado un notable crecimiento en las últimas décadas como alternativa tanto para resolver los problemas de escasez de agua, como para disminuir las presiones sobre las masas de agua superficiales y subterráneas y mejorar la calidad de las aguas destinadas al consumo humano o al regadío. Este crecimiento se traduce en un incremento de la capacidad mundial instalada para desalación de agua de mar, que en 2005 alcanzó aproximadamente los 24.5 millones m³/día, según el informe número 19 de la IDA Worldwide Desalting Plant Inventory. De esta capacidad instalada, el 6% de las plantas se encuentran situadas en la región Pacífica de Asia, 7% en América, 10% en Europa y 77% en Oriente Medio y el Norte de África (Lattemann et al., 2008).

En lo que respecta al Mediterráneo, la producción total es de unos 4.2 millones m³/día (17% de la capacidad mundial). Dentro de estas cifras, España, con un 7% de la capacidad mundial, es el mayor productor de la región: alrededor del 70% de las plantas desalinizadoras españolas están localizadas en la costa mediterránea y las Islas Baleares, y el resto en las Islas Canarias (Lattemann et al., 2008).

En España, la capacidad desaladora ha pasado de 750 plantas, con una producción total de 1.2 Mm³/día en el año 2000, a 900 plantas con una producción superior a los 1,5 Mm³/día en 2005 (Torres, 2005). Por otra parte, con la aprobación del vigente Plan Hidrológico Nacional (Ley 11/2005, de 22 de junio, por la que se modifica la Ley 10/2001, de 5 de julio, del Plan Hidrológico Nacional) y a través del programa A.G.U.A. (Actuaciones para la Gestión y Utilización del Agua) se prevé la puesta en marcha de unas treinta plantas desaladoras, la mayoría en el Arco Mediterráneo (Figura

2.1). Con estas actuaciones, que incluyen tanto la construcción de nuevas instalaciones como la ampliación de algunas de las existentes, se pretende alcanzar una producción de unos 1,1 Mm³/día de agua desalinizada.

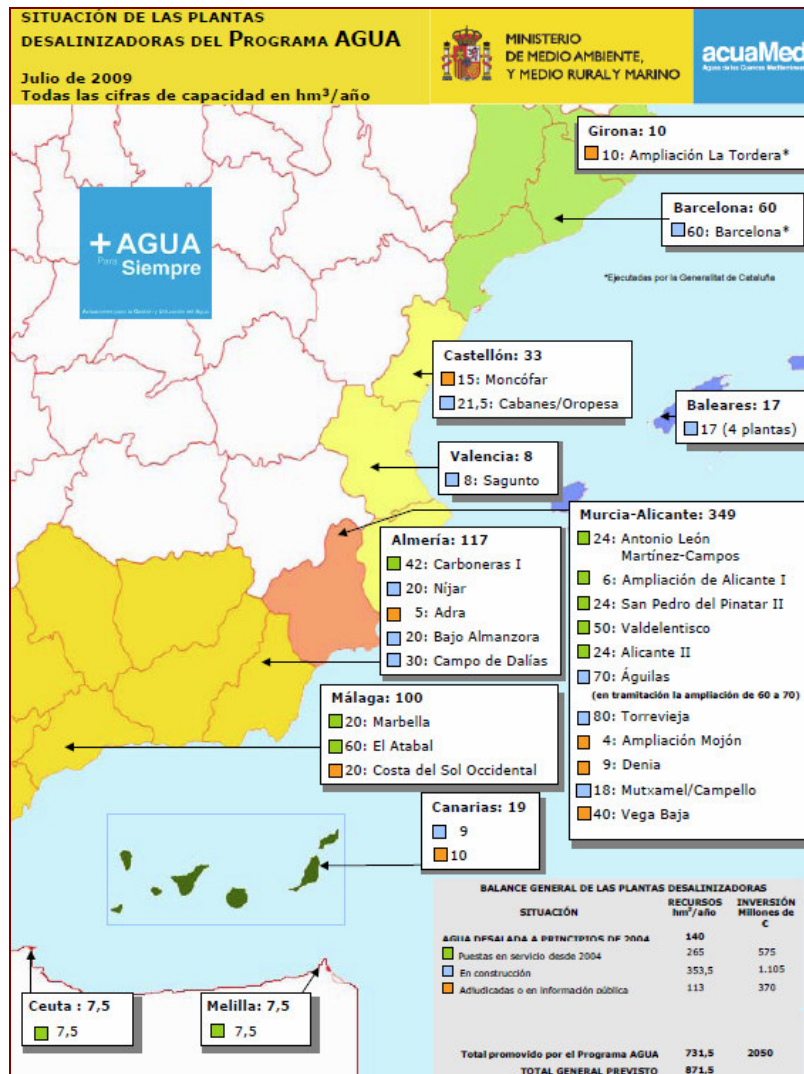


Figura 2.1. Localización de las actuaciones previstas en el Programa A.G.U.A. Fuente: ACUAMED.

Este crecimiento de la capacidad desaladora en España implica un aumento muy significativo del efluente hipersalino salmuera, que es generalmente vertido al mar. La salmuera se caracteriza por su exceso de salinidad con respecto al medio donde se vierte, lo que le confiere una condición de fluido hiperdenso que hace que se hunda hacia el fondo y avance sobre el lecho marino, generando un riesgo significativo de afección y pérdida de hábitats y ecosistemas marinos. Este hecho nos lleva a plantearnos preguntas acerca de si las presiones antes ejercidas sobre las aguas

continentales están siendo trasladadas a las aguas costeras, así como a reconocer la necesidad de evaluar el efecto de la salmuera sobre los ecosistemas marinos.

En el mar Mediterráneo, donde se sitúa la mayor parte de la capacidad desaladora de España, son de especial importancia y fragilidad las fanerógamas marinas, en concreto *Posidonia oceanica*, por lo que es imprescindible el desarrollo de herramientas que permitan un diseño y control eficaz del vertido, que asegure el correcto cumplimiento de las Declaraciones de Impacto Ambiental de las plantas desalinizadoras e incluso ayude a mejorar la definición y redacción de las mismas.

2.2. La desalinización. Definición y tipos

La desalación se define como cualquier tratamiento de aguas (superficiales, subterráneas, salobres o de mar y residuales) que permita reducir los sólidos totales disueltos por medio de procesos térmicos o tecnologías de membrana con el fin de obtener un recurso aprovechable tanto para el abastecimiento humano, como para riego o usos industriales.

El proceso para separar sales del agua en que se encuentran disueltas requiere de una u otra forma de energía, y es precisamente la forma en que se aplica esta energía, uno de los elementos diferenciadores de los procesos disponibles. En la siguiente tabla se recogen las principales tecnologías de desalación en función del principio que aplican y del tipo de energía empleada.

Principio	Energía utilizada	Tecnología
Resina intercambio	Difusión-atracción	Intercambio iónico(II)
Destilación	Vapor	Destilación Multiefecto (MED) Destilación Flash Multietapa (MSF) Compresión de Vapor (CV)
Membrana densa	Presión hidráulica	Osmosis inversa (RO) y Nanofiltración (NF)
Membrana porosa	Campo eléctrico	Electrodiálisis

Tabla 2.2. Principales tecnologías de desalación

En la actualidad, las tecnologías de destilación flash multietapa (MSF) y ósmosis inversa (RO) son las más ampliamente utilizadas a nivel mundial, presentando la última una clara tendencia al alza en los últimos años, sobre todo en España donde el 87% de la capacidad desaladora instalada emplea esta técnica.

Destilación flash multietapa (MSF)

La destilación flash multietapa (MSF en inglés) es el proceso predominante dentro del grupo de procesos térmicos basados en cambios de estado. Esta técnica opera en base al principio de que el agua se hierve a temperaturas progresivamente menores cuando se somete a presiones también menores en la cámara de evaporación. El vapor que se obtiene en el proceso es convertido en agua dulce por condensación.

El principal inconveniente de este tipo de desalación es su elevado consumo energético. En cuanto al agua de rechazo, presenta temperaturas superiores a 100° C, por lo que se mezcla con agua fría para disminuir dicha temperatura hasta aproximadamente unos 10° C por encima del agua receptora del vertido. Este incremento de temperatura minimiza la diferencia de densidad generada por el elevado contenido en sales. Además, y puesto que las plantas MSF suelen estar acopladas a plantas de generación de energía, el efluente es mezclado de forma adicional con el agua de refrigeración de estas. Como resultado el efluente es menos denso que el agua receptora (Lattemann et al., 2003).

Ósmosis Inversa (RO)

En las operaciones de ósmosis inversa (RO en inglés), el agua a tratar en forma de disolución concentrada se convierte en agua dulce al pasar a través de una membrana semipermeable, mediante la aplicación de una cierta presión.

La ósmosis inversa es una tecnología ampliamente utilizada en la desalación de aguas salobres y agua de mar, cuyo principal atractivo es su bajo consumo de energía, gracias al cual se ha convertido en el proceso desalinizador más empleado en la actualidad fuera de los países del Golfo, donde se siguen utilizando principalmente tecnologías de destilación.

La principal ventaja de esta técnica es, que la mejora de los sistemas de recuperación de energía, así como el uso de membranas específicas para trabajar a altas presiones, han conseguido rebajar el coste energético de 4-5 kwh/m³ a menos de 3 kwh/m³ con grados de recuperación del 60%.

Por otra parte, y como principal desventaja de la técnica de ósmosis inversa, se encuentra la elevada salinidad del agua de rechazo generada como resultado del proceso de desalación, cuya composición y comportamiento se describirán en el siguiente apartado.

2.3. Caracterización del vertido de desaladoras.

La principal característica de la salmuera es su elevada concentración de sales. En las plantas desaladoras con tecnología de ósmosis inversa (a cuyo rechazo nos referiremos en adelante) las tasas de conversión están en torno al 45%. Por lo tanto, si un agua marina aporta una salinidad de 37.5 psu, después del proceso de desalación se producirá un desecho con una salinidad entre 43 y 90 psu (Ruiz Mateo, 2007), lo que supone un incremento considerable con respecto a la salinidad natural del medio, que en el caso de las aguas costeras del sureste peninsular se encuentra en torno a 37.5 psu con pequeñas variaciones de ± 0.5 psu (Fernández-Torquemada et al., 2004).

En cuanto a su **composición**, es la misma que la del agua de mar pero en mayor concentración (casi el doble en los procesos de ósmosis inversa). Sin embargo, y debido a los procesos de pretratamiento, depuración y a la limpieza de membranas, puede contener determinados aditivos químicos (Tabla 2.2).

La **caracterización hidrodinámica** del comportamiento de la salmuera se divide generalmente en dos regiones: campo cercano y campo lejano, que se diferencian en los mecanismos físicos dominantes en cada una de ellas (Jirka et al., 2008).

La región de **campo cercano** es la zona inicial de mezcla provocada por las turbulencias que se generan debido a la energía cinética del efluente. En esta zona las condiciones ambientales juegan un papel secundario, y son las características físicas del

efluente y el sistema de vertido empleado los que controlan el comportamiento de la salmuera.

A cierta distancia del punto de descarga, la turbulencia deja de ser el proceso dominante y la salmuera se hunde dando lugar a una pluma hipersalina que avanza siguiendo la línea de máxima pendiente. Esta zona, que se extiende desde cientos de metros a decenas de kilómetros, se denomina **campo lejano**, y en ella las condiciones ambientales (batimetría, rugosidad del fondo, corrientes, oleaje, etc.) y la diferencia de densidad entre fluidos controlan el proceso de mezcla (García et al., 2009; Palomar et al., 2008).

<i>Propiedades físicas</i>	
Salinidad	65.000-85.000 mg/l
Temperatura	Temperatura ambiente
Densidad de la pluma de salmuera	Flotabilidad negativa
Oxígeno disuelto	Si se utilizan pozos como sistema de toma: por debajo del agua de mar del medio receptor debido a bajo contenido en OD del agua de toma. Si se utilizan tomas abiertas: aprox. el mismo contenido en OD que el agua receptora.
<i>Aditivos para el control de biofouling y sub-productos</i>	
Cloro	Si se emplea, normalmente se neutraliza antes de que el agua pase por las membranas para prevenir que estas se dañen.
Compuestos orgánicos halogenados	Bajo contenido, normalmente por debajo de los niveles dañinos.
<i>Eliminación de sólidos disueltos</i>	
Coagulantes (ej. cloruro férrico)	Puede generar la coloración del efluente.
<i>Aditivos para el control de pH</i>	
Acido (H ₂ SO ₄)	No presente (reacciona con el agua de mar para generar compuestos peligrosos, ej. agua y sulfatos; la acidez es absorbida por la alcalinidad natural del agua de mar, con lo que el pH del vertido es normalmente similar o ligeramente superior al agua de mar del medio.

Tabla 2.3. Propiedades típicas del efluente de salmuera de las plantas desaladoras de osmosis inversa. Adaptado de S. Lattemann, T.Höpner, 2008.

Mientras el campo cercano se caracteriza por altas diluciones en escalas de tiempo y espacio pequeñas (segundos y metros), que dependen del método de vertido adoptado, las diluciones que se alcanzan en el campo lejano son bajas, ocurriendo en escalas

espaciales y temporales grandes (km y horas), de forma que la pluma hipersalina puede desplazarse grandes distancias sin apenas dilución, afectando incluso a especies alejadas de la zona de vertido (García et. al, 2009).

2.4. Impacto ambiental de la salmuera sobre el medio marino

Atendiendo a su composición química, la salmuera no puede ser considerada un contaminante, ya que sus componentes son los mismos que los del agua de mar pero en mayor concentración. Sin embargo, esta alta concentración de sales inorgánicas así como los aditivos químicos empleados en el pretratamiento del agua de toma, pueden constituir una fuente de impacto, afectando a los ecosistemas locales. La magnitud de este impacto dependerá de las características de la planta desaladora y de su vertido, pero también de la naturaleza física (batimetría, hidrodinamismo, etc.) y de las condiciones biológicas del ambiente marino receptor (Ahmed et al., 2000; Einav et al., 2002).

Algunos de los efectos de este tipo de vertidos sobre el medio marino son:

- **Incremento de la temperatura.** Los cambios en los perfiles de temperatura de las áreas de descarga pueden acelerar los procesos biológicos, resultar en stress o provocar un abrupto descenso en la actividad cuando los valores críticos son excedidos. Sin embargo, este impacto está asociado a las plantas MSF, ya que las aguas de rechazo procedentes de procesos de ósmosis inversa se vierten con una temperatura igual a la del medio.
- **Incremento de la salinidad.** La descarga de grandes volúmenes de salmuera puede llevar a un incremento de la salinidad en la zona de vertido, y puesto que tanto la salinidad como la temperatura son parámetros ambientales fundamentales para la vida marina, las altas concentraciones de sal pueden ser peligrosas e incluso letales. En general, la toxicidad depende de la sensibilidad de las especies frente al incremento de salinidad, las variaciones de los valores naturales de su hábitat y el estadio del ciclo de vida en el que se encuentren. Por ejemplo, estudios de laboratorio realizados (Fernandez-Torquemada et al., 2005) sobre la especie endémica mediterránea *Posidonia oceanica* han mostrado que salinidades superiores

a 39.1 psu generan una reducción de la vitalidad de la planta (crecimiento de las hojas, necrosis en los tejidos y senescencia prematura en las hojas) y una elevada mortalidad (a 45 psu el 50% de las plantas morían en 15 días). Sin embargo, dos especies de fanerógamas comunes en las aguas del oeste de Australia, *Posidonia australis* y *Posidonia amphibolis*, parecen adaptarse mejor a las altas salinidades. Así, los estudios disponibles indican que algunas especies son más tolerantes a las condiciones hipersalinas que otras, y que *Posidonia oceanica* se encuentra entre las más sensibles.

Además de las fanerógamas, algunas especies animales frecuentes en las praderas, tales como equinodermos o misidáceos, también parecen ser sensibles a las variaciones de salinidad. Experimentos realizados en laboratorio (Sanchez-Lizaso et al., 2008) han mostrado que en ambos casos, la mortalidad se incrementaba con salinidades superiores a los 40.5-41 psu, resultados que sin embargo, dependían siempre de la temperatura.

- **Influencia en los procesos de mezcla.** La descarga de elevadas cantidades de salmuera con una densidad diferente a la del medio puede afectar a los procesos de mezcla y la estratificación de las aguas, especialmente en áreas caracterizadas por corrientes y oleaje débiles. En estas condiciones de calma y escasa renovación de las aguas, la presencia de especies bentónicas puede llevar a episodios de anoxia en el fondo (Hodges, 2006)
- **Incremento de la salinidad en las aguas intersticiales.** El caudal de salmuera que se expande por el lecho marino puede también difundir hacia las aguas intersticiales del sedimento debido a su elevada densidad, e incrementar la salinidad de este agua.

Tal como se ha visto en párrafos anteriores, el mayor perjuicio causado por los vertidos hipersalinos se produce sobre las especies bentónicas, que en el caso del Mediterráneo suelen ser fanerógamas tales como *Posidonia oceanica* o *Cymodocea nodosa*.

En las siguientes líneas se procederá por tanto a dar una descripción de estas dos especies.

Posidonia oceanica

La *Posidonia oceanica* (L.) Delile es una especie endémica del Mediterráneo, donde esta especie longeva y de crecimiento lento forma extensas praderas entre los 5 y 30 metros de profundidad, dependiendo de la transparencia del agua y la granulometría de los fondos.

Posidonia oceanica es una especie “ingeniera” (Jones et al. 1997), constructora de un hábitat, la pradera, que realiza diversas funciones y servicios ambientales, entre los cuales:

- El entramado vegetal de la pradera sirve como sustrato colonizable por una rica comunidad epífita de organismos sésiles, y proporcionan **alimento, lugar de puesta y cría o refugio** a una gran diversidad de organismos vegetales y animales, entre ellos algunas especies de peces de interés gastronómico o comercial (Bell et al., 1982; Francour, 1997).
- Su estructura firmemente arraigada en el sedimento **previene la erosión por corrientes y oleaje** de los fondos arenosos (Gacia et al., 2001).
- Su elevada producción, además de sostener la persistencia y crecimiento de la pradera, tiene otros destinos, como la alimentación de multitud de especies de herbívoros. Además, la fracción de materia orgánica que no es consumida por los herbívoros, pasa a formar parte de la hojarasca con la caída de las hojas. Esta hojarasca, junto a la materia orgánica depositada puede ser exportada por corrientes y oleajes a sistemas vecinos (p.e. los deficitarios sistemas profundos, o las playas donde se acumula formando las denominadas “banquetas” que **protegen la playa frente a la acción erosiva** de los temporales), entrar en la red trófica a través de descomponedores o detritívoros, o ser enterrada junto a los sedimentos retenidos pasando a formar parte de la denominada “mata” (Mateo et al., 1997; Pergent et al., 1994; Cebrián et al., 2001). Estos materiales orgánicos acumulados bajo las praderas representan “sumideros” de carbono y otros nutrientes, por lo que podrían ayudar a **atenuar los efectos del cambio climático y de la eutrofización** de las aguas costeras.

Todos los bienes y servicios ambientales, ecológicos y económicos enumerados, se encuentran ampliamente reconocidos, considerándose de gran relevancia. Tanto es así, que constituyen un patrimonio legalmente protegido nacional (Ley 42/2007 del Patrimonio Natural y la Biodiversidad), e internacionalmente (Directiva Hábitat 92/43/CEE).

Sin embargo, y pese a la protección bajo la que se encuentra, esta especie se encuentra amenazada por diferentes actividades humanas, que generan una serie de impactos marcadamente multifactoriales, es decir, que modifican simultáneamente diferentes factores ambientales. Uno de estos impactos es el producido por el vertido de aguas procedente de plantas desalinizadoras, que con una cantidad de sales superior a la del agua de mar, puede provocar alteraciones fisiológicas y del crecimiento, necrosis y mortalidad de haces (Fernández-Torquemada et al., 2005; Gacia et al. 2007).

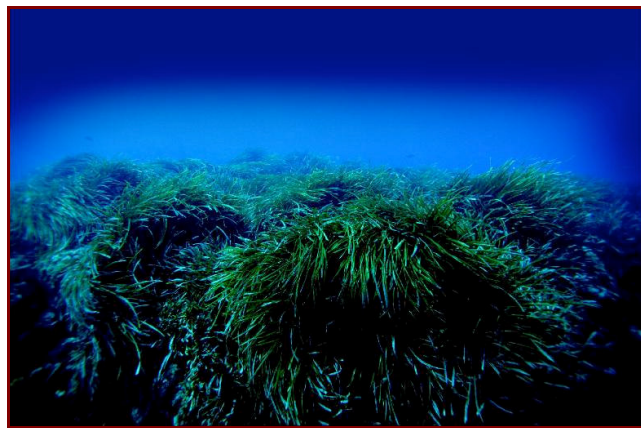


Figura 2.2. Pradera de *Posidonia oceanica*. La Zohía, 14 m de profundidad.

Circunstancias tales como sus altos requerimientos de luz, que implican la necesidad de una alta transparencia de las aguas, su reducida capacidad de colonización o su baja tolerancia a las fluctuaciones de salinidad, explican la elevada sensibilidad de *Posidonia oceanica* a los cambios originados por la actividad antrópica, causa principal de la regresión de estas comunidades en el litoral mediterráneo.

El vacío de conocimiento científico acerca de la importancia de la salinidad en la autoecología de *Posidonia oceanica*, junto con las perspectivas de desarrollo de la desalinización y la necesidad de evaluar su impacto ambiental, ha propiciado la puesta en marcha de varios proyectos científicos dirigidos a establecer los límites de tolerancia

de esta especie a la salinidad y la influencia de este factor en la composición y funcionamiento de este ecosistema.

El primero de estos proyectos fue el elaborado por diferentes expertos pertenecientes al CEDEX, CSIC, IEO, al Departamento de ecología de la universidad de Barcelona y al Departamento de Ciencias Ambientales de la Universidad de Alicante, en el cual se establecieron unos límites de salinidad según los cuales se garantiza la no afección a la Pradera de *Posidonia oceanica* (Sánchez-Lizaso et al., 2008). Estos criterios son: no superar en ningún punto de la pradera 40 psu de salinidad en más del 5% de las observaciones, o 38.5 psu en más del 25% de las observaciones a lo largo de un año. Sin embargo, cabe destacar la ambigüedad de las citados criterios, en los que no se especifica la frecuencia con la que se han de realizar las mediciones necesarias para calcular los percentiles indicados anteriormente.

Cymodocea nodosa

Las comunidades de *Cymodocea nodosa* se instalan sobre arenas finas o algo fangosas no expuestas a un hidrodinamismo muy acentuado. El crecimiento conjunto de un elevado número de plántulas de *Cymodocea nodosa* es lo que da origen a las praderas, que tienen un efecto fijador sobre los fondos móviles donde se implantan. Suele ocupar los fondos arenosos del litoral comprendidos entre 6 y 20 m de profundidad, donde suele formar una banda continua previa a la pradera de *Posidonia oceanica*.

Pese a que las praderas de *Cymodocea nodosa* ocupan una menor extensión que las *Posidonia oceanica*, su papel ecológico es igualmente relevante. Asimismo, también son sensibles a la alteración de su hábitat, por lo que un incremento de salinidad en el medio puede implicar efectos negativos tales como los vistos para *Posidonia oceanica*. Sin embargo, el rango de salinidades en el cual habita esta especie es mayor que en el caso anterior. De hecho existen praderas en el Mediterráneo, que presenta salinidades sobre 37.5 psu, en Canarias, donde aparecen salinidades menores, de unos 35 psu y hasta en el Mar Menor de Murcia, donde se encuentran salinidades superiores a 43 psu.

Los criterios de calidad para la protección de las praderas de *Cymodocea nodosa*, fueron publicadas en la Declaración de Impacto Ambiental de la Nueva desalinizadora

de Águilas-Guadalestín (BOE nº 117, 17 de mayo de 2006), donde se establecen los siguientes requisitos de calidad: en ningún punto de la pradera la salinidad se podrá superar 41 psu en más del 5% de las observaciones y 39.5 psu en más del 25% de las observaciones.



Figura 2.3. Césped de *Cymodocea nodosa*. Calabardina, 5 m de profundidad.

2.5. Planes de Vigilancia Ambiental. Necesidad y contenidos

La actual apuesta por la desalinización ha supuesto un gran aumento del caudal de salmuera vertido al mar, y de las consecuencias que ésta tiene sobre el medio. Por ello, es necesario establecer unas medidas con el fin de reducir, eliminar o compensar los efectos ambientales negativos.

A este respecto, el Real Decreto Legislativo 1/2008, de 11 de enero, aprueba el texto refundido de la Ley de Evaluación de Impacto Ambiental de proyectos (EIA), publicada en el BOE de 26 de enero de 2008. El artículo 7 de dicho Real Decreto regula el contenido de los Estudios de Impacto Ambiental (EsIA), incluyendo en su apartado e el Programa de Vigilancia Ambiental (PVA), cuyo objeto es el seguimiento de la efectividad de las medidas citadas anteriormente.

En el caso de las instalaciones desalinizadoras de agua marina, los objetivos generales de los mencionados planes de vigilancia se enumeran en la siguiente página:

- Analizar el grado de ajuste entre el impacto que teóricamente generará la actuación, y el impacto real producido durante la ejecución de la obra y la fase de funcionamiento.
- Detectar las eventualidades aparecidas durante el desarrollo de la actuación.
- El seguimiento y control de la eficacia y cumplimiento de las medidas establecidas en el EsIA y la autorización de vertido para el efluente y el medio receptor.
- Realizar las modificaciones necesarias o expansiones convenientes de las medidas anteriores a la luz de nueva información y tomar acciones para gestionar impactos no previstos
- Adquirir conocimientos para la mejora de proyectos similares y la redacción de sus EsIA.

El desarrollo de los Planes de Vigilancia Ambiental para la consecución de los anteriores objetivos se divide en tres fases, la preoperacional, la fase de obra y la fase de explotación.

El propósito de la vigilancia durante la **fase preoperacional** es la definición de la situación cero del área que se verá afectada con la construcción de la planta desalinizadora. Durante esta fase se deberá realizar un control de la calidad del medio receptor, un mapa de curvas de salinidad y temperatura, el análisis del clima marítimo y la caracterización del estado de conservación de las praderas de fanerógamas marinas y otros organismos marinos. Una vez se comienza el proyecto, y durante toda la **fase de obra**, se llevarán a cabo controles de la calidad del medio receptor y del material de dragado, así como la caracterización del estado de conservación de las praderas de fanerógamas marinas y otros organismos marinos (arrecifes biogénicos, equinodermos, moluscos).

Por último, y durante su desarrollo en la **fase de explotación** de la planta, el plan de vigilancia deberá contener controles de la calidad del medio receptor y del efluente, una caracterización espacio-temporal de la pluma de salinidad producida por el vertido, el control de posibles afecciones sobre las praderas de fanerógamas marinas y otros

organismos marinos (arrecifes biogénicos, equinodermos, moluscos), la vigilancia estructural de las conducciones marinas, y un protocolo de corrección del exceso de salinidad. A continuación se detallarán algunos de los aspectos más importantes de cada uno de estos puntos:

- **Vigilancia estructural de la conducción de vertido.** Se examinará al menos con una periodicidad anual toda la longitud del tramo sumergido de la conducción del vertido y sus elementos.
- **Caracterización de la pluma.** Se realiza durante las primeras semanas de funcionamiento de la planta, y se basa principalmente en la toma de medidas de salinidad en diferentes puntos que permitan caracterizar la geometría y salinidad del efluente.
- **Vigilancia ambiental de la calidad del efluente y del medio receptor.** El muestreo deberá realizarse de una manera sistemática, manteniendo constantes los puntos, la periodicidad y períodos de muestreo, cuya descripción y localización deberán estar claramente detallados en el PVA.
 - *Control de la calidad del efluente.* El objetivo de este control es la determinación de la salinidad y temperatura del vertido, así como del caudal que se está vertiendo.
 - *Control y seguimiento de las aguas receptoras.* Para el muestreo de las aguas receptoras, se establecen por término general entre 12 y 16 estaciones de muestreo: tres puntos sobre la línea de costa, otro en la zona de vertido del efluente y el resto a diferentes distancias y direcciones de la zona de vertido, de acuerdo con las predicciones del modelo de dispersión del vertido. Para cada estación se mide un perfil continuo de temperatura, salinidad y densidad, a lo largo de toda la columna de agua. La frecuencia de muestreo será trimestral y se realizará a dos niveles (en la superficie y a 20 cm del fondo).

Además de los controles trimestrales anteriores, cuando el vertido se produce en zonas próximas a especies de fanerógamas marinas, se establece la obligación de instalar equipos que registren de modo automático la salinidad y

temperatura del agua al menos en cuatro estaciones situadas en la zona próxima a dichas especies. Los datos se recogen semanalmente y mediante análisis estadístico se determinan, en el caso de *Posidonia oceanica*, los percentiles 5% y 25%, para comprobar que los umbrales 40 y 38.5 psu no se superan en el 5% y 25% de las observaciones, respectivamente.

Adicionalmente, se instalan correntímetros ópticos para la monitorización de las condiciones hidrodinámicas asociadas a la dispersión del vertido, con sensor de presión para poder determinar el nivel medio del mar y el oleaje.

- ***Control de los ecosistemas a proteger.*** Este seguimiento se refiere frecuentemente a las praderas de fanerógamas marinas situadas en las proximidades del vertido. Se realizará un seguimiento anual coincidiendo con la época estival, durante el cual se seleccionarán alrededor de 4 o 5 estaciones de muestreo en zona próxima, media y alejada del vertido. En cada una de estas estaciones se medirán los parámetros indicadores de la calidad de las especies presentes. Estos trabajos se repetirán en dos estaciones control alejadas del punto de vertido a igual profundidad, lo que proporcionará información para analizar si los cambios observados se deben a la influencia del vertido o la variabilidad natural de las praderas.
- ***Control de sedimentos y organismos.*** Se establece un control anual y en la misma época del año en los fondos marinos y en las comunidades bentónicas.
- **Protocolo de corrección del exceso de salinidad.** En caso de detectar lecturas de salinidad superiores a los umbrales marcados para las especies presentes en el área influenciada por el vertido, se deberá aplicar alguna de las diferentes medidas correctoras existentes: aumento del caudal de vertido manteniendo el diámetro de las boquillas y de la predilución, mantenimiento del caudal y el diámetro y aumento de la predilución, disminución del diámetro de boquilla manteniendo el caudal y predilución.

En caso de que las medidas correctoras no sean suficientes para la disminución de la salinidad por debajo de los límites establecidos en la autorización de vertido, la

planta deberá cesar inmediatamente el mismo. Además, el titular de la planta deberá remitir un informe detallado cada vez se que se detecten lecturas de salinidad superiores a las permitidas, antes de transcurrida una semana desde su detección y anualmente junto con el PVA. En el informe se indicarán los valores obtenidos en los controles de las zonas a proteger y el control de las variables de flujo, el motivo y las medidas adoptadas, así como cualquier otra información que sea relevante para conocer el funcionamiento de la planta y, en su caso, de la afección al medio marino.

La revisión realizada de los PVA demuestra que este tipo de vigilancia sólo permite una gestión reactiva del impacto ambiental, ya que tanto los incumplimientos como los cambios adversos se detectan una vez han tenido lugar. Sin embargo, dada la alta sensibilidad de algunas especies como *Posidonia oceanica* y su fauna asociada a los incrementos de salinidad, incluso cuando estos son moderados, la gestión reactiva no resuelve adecuadamente el problema.

3 REVISIÓN BIBLIOGRÁFICA DE LAS HERRAMIENTAS PREDICTIVAS PARA LA MODELIZACIÓN DEL VERTIDO DE LAS DESALINIZADORAS

Tal como se ha visto en el capítulo anterior, los vertidos procedentes de las instalaciones desalinizadoras tienen numerosos efectos adversos sobre el medio marino, lo que hace imprescindible la necesidad de disponer de un modelo matemático que sirva para realizar predicciones del comportamiento de la salmuera al ser vertida en el medio receptor.

A este respecto existen numerosas herramientas y software comerciales que se han venido empleando en la simulación del vertido de salmuera al mar. Algunas de ellas son: CORMIX, VISUAL PLUMES y VISJET, principalmente para la predicción de campo cercano, y COHERENS, DELFT 3D, etc., para la predicción del campo lejano. La mayoría de las aplicaciones se centran en la dispersión en campo cercano y presentan simplificaciones importantes, siendo aún más acusada la carencia de modelos de campo lejano y de modelos que integren el comportamiento en ambas regiones.

La necesidad de disponer de un único modelo matemático que sirva para realizar predicciones del comportamiento de la salmuera tanto al ser vertida en el medio marino, como en campo lejano, es algo que todos los implicados en el ámbito de la desalación necesitan cada vez con mayor urgencia. Disponer de un sistema predictivo fiable permitirá adoptar las medidas correctoras que realmente se necesitan para reducir el impacto sobre el medio ambiente, haciendo que las acciones llevadas a cabo sean en la dirección correcta y con las inversiones que realmente se precisen. Lo contrario puede llegar a hacer inviables plantas desalinizadoras, por la necesidad de llevar a cabo medidas excesivas de protección marina debido a una predicción incorrecta del comportamiento de la salmuera.

En la mayoría de los casos, el comportamiento de la salmuera se ha venido evaluando a partir de modelos contrastados experimentalmente como CORMIX, útiles en la fase de diseño del vertido, pero limitados en la fase de explotación. Por otro lado, los modelos de tipo determinista todavía no tienen la capacidad necesaria para resolver el complejo

problema del transporte y difusión de salmueras en el medio marino. En parte, este lento desarrollo está motivado por la escasez de datos experimentales en el medio natural.

En cuanto a la integración en un único modelo del estudio de campo cercano y campo lejano, solamente existen a día de hoy algoritmos que acoplan modelos. Como ejemplo, destaca el trabajo realizado por Gerhard Jirka y Tobias Bleninger del *Institute for Hydromechanics* (Karlsruhe, Alemania), que han generado un algoritmo que acopla los programas CORMIX, para campo cercano, y Delft3D, para campo lejano. No obstante, es inminente la necesidad de un modelo que acople ambos campos en una única aplicación u otras herramientas que solventen las limitaciones de las existentes.

Dada la incertidumbre asociada a las predicciones de los modelos actuales de modelización, y a lo limitado de uso durante la fase de explotación, en este trabajo se expondrá un método que basándose en herramientas de inteligencia artificial, más concretamente en redes neuronales y lógica difusa, será capaz de aprender a partir de valores pasados, y con ello, predecir valores futuros de salinidad.

A continuación se repasarán las herramientas de modelización de vertidos de salmuera, tanto los modelos utilizados actualmente, como las técnicas basadas en inteligencia artificial en las cuales se basa el modelo que será propuesto en esta tesina, y a las cuales se hará referencia frecuentemente en próximos capítulos.

3.1. Modelos físicos

Los modelos físicos, o modelos reducidos, son una técnica para resolver problemas de ingeniería consistente en el ensayo de una réplica del problema a escala reducida. La réplica reducida es lo que se denomina modelo, frente a la realidad, que llamamos prototipo.

La base de los modelos reducidos es la teoría de la semejanza (semejanza geométrica, cinemática y dinámica). La semejanza entre prototipo y modelo se obtiene cuando las posibles relaciones entre las variables geométricas, dinámicas y cinemáticas existentes entre ambos se mantienen constantes.

Entre las ventajas que presentan este tipo de modelos se encuentra el hecho de que permiten un control sobre las variables involucradas, y las condiciones de contorno. Por otra parte, presentan una relativa facilidad de medida y facilitan la comprensión de los procesos. Además, y en cuestiones de calibrado de modelos matemáticos, constituyen la representación del mundo físico real.

Como inconvenientes se pueden citar la lentitud y coste de la generación de alternativas, el hecho de que la semejanza hidráulica es siempre parcial, así como el control de las condiciones de contorno, o las limitaciones debidas a la necesidad de espacio o dificultad de generación y medida.

En el campo de la modelización física de salmueras, cabe destacar los trabajos que desde el año 2000 se han venido realizando en el CEDEX (Ruiz Mateo et al., 2008). Estos incluyen, medidas en la naturaleza sobre vertidos de desaladoras en funcionamiento, y ensayos en laboratorio mediante modelos reducidos en 2 y 3 dimensiones.

De los resultados de los ensayos realizados en modelos reducidos en 2D y 3D y de las medidas realizadas en la naturaleza se puede extraer el siguiente cuadro orientativo respecto a las diluciones que se pueden conseguir en el campo cercano según la forma del vertido:

VALORES APROXIMADOS DE LAS DILUCIONES EN EL CAMPO CERCANO EN DIFERENTES FORMAS DE VERTIDO⁽¹⁾⁽²⁾	
FORMA DEL VERTIDO	DILUCIÓN
Descarga enterrada en una playa de bolos	2,5
A un torrente, cerca de la desembocadura	4
Chorro libre sobre la escollera de un dique	6
Chorro libre horizontal en el lecho marino	12
Chorro libre en un acantilado	18
Difusor sumergido con 2 bocas/elevador	24
Difusor sumergido con 1 boca/elevador	30

(1) Sin oleaje
(2) En función de los detalles del diseño pueden variar entre la mitad y el doble

Tabla 3.1. Diluciones en campo cercano obtenidas a partir de ensayos modelos 2D y 3D y medidas realizadas en la naturaleza. Fuente: Ruiz Mateo, 2007.

3.2. Modelos de ecuaciones integradas

Este tipo de modelos resuelve las ecuaciones hidrodinámicas integradas a lo largo de la sección transversal con el objetivo de estimar la anchura de la capa de mezcla y el desplazamiento transversal. Las ecuaciones integradas se resuelven mediante métodos numéricos sencillos. Las condiciones de contorno de estos modelos incluyen hipótesis como la no existencia de flujo a través de las paredes o que la influencia de los contornos es despreciable. Debido a estas hipótesis, el medio receptor tiene que asumirse como ilimitado, sin posibilidad de simular la interacción de efluente con los contornos. Por ello, con este tipo de modelo, no es posible simular el efecto de la batimetría, la presencia de obstáculos, de canales submarinos, la interfaz entre capas, etc. En Palomar et. al, 2008 se citan como ejemplos de este tipo de modelos las formulaciones de Alavian (1986, 1996) y García (1996).

3.3. Modelos hidrodinámicos

Este tipo de modelos resuelven numéricamente las ecuaciones de gobierno (estado, continuidad, conservación de la cantidad de movimiento y la ecuación de transporte o de advección-difusión). Como principal ventaja de estos modelos se puede destacar que asumen muchas menos hipótesis simplificadoras que los modelos de integración, lo que permite simular fenómenos más complejos y obtener una simulación más realista del comportamiento del flujo. Sin embargo, requieren de un elevado número de datos de partida y presentan una dificultad de aplicación e interpretación de los resultados. Su principal inconveniente se debe al error de difusión numérica que se produce en la resolución de las ecuaciones en las zonas frontera, con importantes gradientes de magnitud. En campo lejano, donde los valores de dilución son pequeños, estos errores, en la zona de interfaz (picnoclina) entre capas, son del orden de magnitud del fenómeno de difusión del fluido, lo que deriva en resultados no fiables, con tamaños de pluma y grado de dilución sobreestimados. Con celdas suficientemente pequeñas se reduciría el error numérico; sin embargo, la localización de las zonas frontera no es conocida a priori, y si se emplean estas celdas en todo el ámbito del modelo, resultaría inviable a nivel computacional (García et al., 2009).

Seguidamente, y a modo de presentación de algunas de las ventajas e inconvenientes de este tipo de modelos, se procederá a realizar una descripción CORMIX, uno de los más empleados para la simulación del comportamiento de la salmuera en el medio marino.

CORMIX

Cornell Mixing Zone Expert System (CORMIX) es una plataforma de simuladores para el análisis, la predicción y el diseño de sistemas de descarga de sustancias tóxicas o convencionales sobre masas de agua. Fue desarrollado conjuntamente por la USEPA (United States Environmental Protection Agency) y la Universidad de Cornell durante el periodo de 1985-1995. Sus fórmulas, rigurosamente verificadas mediante ensayos de laboratorio y campo, y la aplicación a multitud de casos reales desde hace más de diez años, confirman la fiabilidad de la herramienta. En la actualidad es admitido como estándar por la EPA (U.S. Environmental Protection Agency) y solicitado en los estudios de vertidos por la EEA (European Environmental Agency).

Se trata de una herramienta que permite predecir las características de la geometría y la dilución de la zona de mezcla inicial para poder juzgar la calidad del agua tras el vertido. Además, en teoría, es capaz de estimar el comportamiento de la pluma hipersalina a grandes distancias del punto de vertido, en la región de campo lejano.

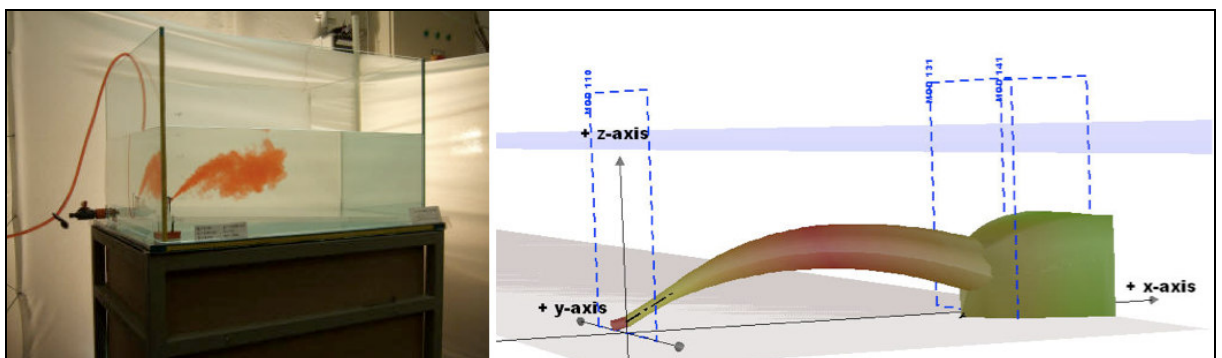


Figura 3.1. Izq: Experimento de laboratorio para simular el vertido de salmuera procedente de una planta RO. Derch: Modelado con CORMIX.

El sistema de clasificación experto CORMIX está basado en unas reglas de entrada y procesamiento de datos e incluye tres subsistemas según el tipo de dispositivo de vertido a simular: CORMIX1 para el análisis de descargas sumergidas mediante chorro

individual, CORMIX2 para el análisis de descargas sumergidas mediante tramo difusor de chorros múltiples, y CORMIX3 para el análisis de descargas superficiales.

En esencia, CORMIX es un modelo basado en el análisis dimensional del fenómeno, que deriva en fórmulas de tipo escala de longitud para la caracterización del flujo, dotado de un avanzado sistema experto de clasificación de flujos basados en algoritmos de inteligencia artificial. El sistema CORMIX consta de varias posibles combinaciones de módulos, cuya secuencia de aplicación es determinar mediante un primer análisis y clasificación del problema. De esta manera, tras la entrada de datos realiza un cálculo elemental de varios parámetros físicos incluyendo los flujos, escalas de longitud y relaciones adimensionales. En base a estos parámetros un sistema experto realiza una clasificación del flujo, aplica las fórmulas de cálculo semiempíricas que corresponda y proporciona una breve descripción de sus principales características.

CORMIX asume simplificaciones muy significativas en el comportamiento del vertido que hay que tener en cuenta en la interpretación de los resultados obtenidos. Por ello, es de especial interés realizar la validación del modelo comparando los resultados de la simulación con las mediciones realizadas in situ. Las simplificaciones más importantes son:

- La esquematización de la sección transversal de la masa de agua receptora, como un canal rectangular no limitado lateralmente por contornos o limitado, que se asume uniforme en la dirección de la corriente. No es posible simular el avance de la pluma de salmuera por efecto de la pendiente de fondo.
- El análisis de las evaluaciones de la zona de mezcla se realiza bajo la suposición de condiciones ambientales y de descarga en estado estacionario. Esta suposición suele ser adecuada cuando los procesos de mezcla son rápidos en relación a la escala de tiempo de las variaciones hidrográficas. En flujos mareales permite realizar simulaciones considerando efectos de reentrada en el comportamiento de la pluma.
- Aunque D-CORMIX permite simular descargas superficiales de vertidos hiperdensos, los resultados obtenidos en la validación demuestran incertidumbres muy significativas.

- Los resultados obtenidos se basan en fórmulas semiempíricas y serán tanto más precisos cuanto más fiables sean los datos de partida. Además, tras una extensa comparación con datos de campo y laboratorio de los resultados de las simulaciones con el modelo CORMIX, ha mostrado que las predicciones de las diluciones y las concentraciones suelen tener una precisión del 50% (desviación estándar), respecto a las medidas reales.

Entre las ventajas fundamentales frente a sus principales competidores destacan:

- Mientras la mayoría de modelos asumen un cuerpo receptor infinito, CORMIX realiza un énfasis especial en la consideración de la posibilidad de considerar contornos, tales como la superficie, el fondo, la existencia de orillas o incluso termoclinas, y de simular la interacción del flujo con los mismo. En base a estas fronteras establece la estabilidad de la descarga, avisando en caso de descargas inestables.
- Considera la influencia de las corrientes ambientales y el ángulo que forman con el chorro y de la estratificación en la columna de agua del medio receptor.
- Simula el comportamiento del efluente en campo lejano, cuando forma una corriente de densidad en el medio receptor que se desplaza por gravedad y que está sometida a procesos de difusión pasiva, produciendo en ocasiones intrusiones campo arriba del punto de vertido. Este tipo de corrientes es ignorado en el resto de modelos.
- CORMIX realiza una rigurosa clasificación del vertido basado en algoritmos de inteligencia artificial, de amplio consenso entre la comunidad científica.
- Es capaz de simular adhesión del flujo al fondo, tanto las provocadas por flujos turbulentos (uniones “coanda” en la literatura inglesa) como uniones de carácter más débil.
- Es capaz de analizar la eficacia de numerosas configuraciones de dispositivo de vertido, según el número de boquillas, su distanciamiento, alineación, etc., considerando además la posible interacción entre chorros.

3.4. Sistemas basados en inteligencia artificial

3.4.1. Redes neuronales

Las redes neuronales artificiales (*artificial neural networks*), son sistemas que imitan (esquemáticamente) la estructura hardware (neuronal) del cerebro para tratar de reproducir algunas de sus capacidades. Mediante un estilo de computación paralelo, distribuido y adaptativo, son capaces de aprender a partir de ejemplos, presentando un cierto comportamiento "inteligente" (Martín, 2006).

El origen de las redes neuronales data de 1940, cuando McCulloch & Pitts (1943), demostraron que las redes de neuronas artificiales podían, en principio, computar cualquier función aritmética. La primera aplicación práctica de las redes neuronales fue presentada por Rosenblatt a finales del 1950. En su libro, publicado en 1962, *Principles of Neurodynamics*, introdujo un algoritmo de aprendizaje con el que se podían modificar los pesos, y demostró la capacidad para el reconocimiento de patrones con una red perceptron.

Sin embargo, el verdadero desarrollo de las redes neuronales llegó en 1980, cuando Rumelhart & McClelland (1986) desarrollaron el método de aprendizaje más influyente hasta el momento, conocido como algoritmo de retropropagación (*backpropagation algorithm*). Durante esta misma época, surgieron nuevos tipos de redes neuronales con características dinámicas, como la red de Hopfield (Hopfield, 1982) y las redes autoorganizativas de Kohonen (Kohonen, 1982), nuevos desarrollos que pusieron en auge el campo de las redes neuronales.

Para entender el creciente interés por las redes neuronales artificiales, compararemos las diferencias entre la arquitectura de los actuales computadores y el cerebro, cuya estructura intentan imitar las redes neuronales. Los computadores son máquinas de Von Neumann, en esencia una máquina de procesamiento (hardware) que actúa ejecutando en serie (una tras otra) una secuencia de instrucciones o programa (software), que almacena en su memoria. Este tipo de estructuras, necesitan, para resolver un determinado problema, la creación de un algoritmo, que se materializará en forma de programa escrito en un determinado lenguaje de programación que el computador

entienda, y un microprocesador que deberá ejecutar dicho programa. El proceso descrito, corresponde a un enfoque descendente (desde arriba) hacia la inteligencia.

Por el contrario, el cerebro no es una máquina de Von Neumann, sino que está compuesto por millones de procesadores elementales o neuronas, ampliamente interconectadas conformando redes de neuronas. La neurona es en realidad un pequeño procesador, lento, sencillo y poco fiable (a diferencia de los microprocesadores de los computadores), sin embargo, en el cerebro cohabitan unos cien mil millones de neuronas operando en paralelo. De esta forma, aunque individualmente las neuronas sean capaces de realizar procesamientos muy simples, ampliamente interconectadas a través de las sinapsis y trabajando en paralelo, pueden desarrollar una actividad global de procesamiento enorme.

Por otra parte, las neuronas no deben ser programadas, éstas aprenden a partir de las señales que reciben del entorno, y operan siguiendo un esquema también muy diferente al de los computadores convencionales. En este punto nos encontramos por primera vez con la idea de autoorganización. En una red de neuronas no existe un componente que gobierne el sistema, sino que las neuronas se influyen a través de sinapsis excitatorias e inhibitoras, lo que causa una compleja dinámica de activaciones y desactivaciones entre ellas. Las neuronas, en definitiva, se autoorganizan, aprendiendo del entorno y adaptándose a él, y de esta autoorganización emergen ricas propiedades de procesamiento.

Por tanto, la idea de partida de las redes neuronales artificiales es que para ejecutar aquel tipo de tareas que más eficazmente resuelve el cerebro, puede resultar interesante copiar su estructura hardware, creando un sistema compuesto por múltiples neuronas ampliamente interconectadas, y estudiar si a partir de su autoorganización pueden reproducirse sus capacidades. Esta ruptura con el funcionalismo de la estructura de cómputo clásica aparece también en otros campos, que se suelen englobar dentro de la denominada *computación emergente* (IEEE 99) o *soft computing* (Jang, 1997).

3.4.1.1. *Modelo general de neurona artificial*

Las redes neuronales artificiales constan de una serie de dispositivos simples de cálculo, llamados neuronas, que se encuentran conectadas entre sí por conexiones que tienen asignado un determinado valor numérico denominado peso. Cada neurona, a partir de un vector de entrada procedente del exterior, o de otras neuronas, proporciona una única respuesta o salida. Los elementos que constituyen la neurona de etiqueta i son:

- Conjunto de entradas, $x_j(t)$
- Pesos sinápticos de la neurona i , w_{ij} que representan la intensidad de interacción entre cada neurona presináptica j y la neurona postsináptica i .
- Regla de propagación $\sigma(w_{ij}, x_j(t))$, que proporciona el valor del potencial postsináptico $h_i(t) = \sigma(w_{ij}, x_j(t))$, de la neurona i en función de sus pesos y entradas.
- Función de activación $f_i(a_i(t-1), h(t))$, que proporciona el estado de activación actual $a_i(t-1)$ y de su potencial postsináptico actual.
- Función de salida $F_i(a_i(t))$, que proporciona la salida actual $y_i(t) = F_i(a_i(t))$ de la neurona i en función de su estado de activación.

De este modo, la operación de la neurona i puede expresarse como:

$$y_i(t) = F_i(f_i[a_i(t-1), \sigma_i(w_{ij}, x_j(t))]) \quad \text{Ec.3.1.}$$

3.4.1.2. *Arquitecturas de redes neuronales*

Se denomina arquitectura a la topología, estructura o patrón conexas de una red neuronal. Atendiendo a múltiples conceptos, pueden establecerse diferentes tipos de arquitecturas neuronales. Así, en relación a su estructura en capas, podemos hablar de redes monocapa y de redes multicapa. Las redes monocapa son aquellas compuestas por una única capa de neuronas, mientras que en las redes multicapa, las neuronas se organizan en varias capas.

Asimismo, atendiendo al flujo de datos en la red neuronal, podemos hablar de redes unidireccionales (*feedforward*) y redes recurrentes (*feedback*). En las primeras, la información circula en un único sentido, desde las neuronas de entrada hacia las de salida. En las redes recurrentes sin embargo, la información puede circular entre las capas en cualquier sentido, incluido el de salida-entrada.

Por último, encontramos las redes autoasociativas y heteroasociativas. Si una red se entrena para que ante la presentación de un determinado patrón de entradas responda con otro diferente, se dice que la red es heteroasociativa, mientras que si una red es entrenada para que asocie un patrón de entrada consigo mismo, se dice que la red es autoasociativa. El interés del último tipo de redes (autoasociativas), como es el caso de la de Hopfield (Hopfield, 1982), reside en que ante la presentación del patrón $A' = A + ruido$, su respuesta sea el patrón original A , eliminando así el ruido presente en la señal de entrada.

3.4.1.3. Tipos de aprendizaje.

Descritas la estructura y las diferentes arquitecturas, el siguiente elemento diferenciador dentro de los múltiples tipos de redes neuronales es el tipo de aprendizaje que se aplica durante el entrenamiento.

En el contexto de las redes neuronales puede definirse el aprendizaje como el proceso por el que se produce el ajuste de los parámetros libres de la red a partir de un proceso de estimulación por el entorno que rodea a la red. El tipo de aprendizaje viene dado por la forma en la que dichos parámetros son adaptados.

El proceso de aprendizaje es usualmente iterativo, actualizándose los pesos sinápticos una y otra vez, hasta que la red neuronal alcanza el rendimiento deseado.

Los dos tipos de aprendizaje son el **supervisado** y el **no supervisado**. Ambas modalidades pretenden estimar funciones de entrada/salida multivariable o densidades de probabilidad, pero mientras que en el aprendizaje supervisado se proporciona cierta información sobre estas funciones, en el autoorganizado no se proporciona información alguna. Las reglas de aprendizaje supervisadas suelen ser computacionalmente más complejas, pero también más exactos sus resultados.

Además de las dos formas básicas citadas, pueden distinguirse muchas otras, entre las que destacan el aprendizaje híbrido y el reforzado. Los cuatro tipos citados pueden definirse de la forma siguiente:

- **Aprendizaje supervisado.** Sea $E(W)$ un funcional que representa el error esperado de la operación de la red, expresado en función de sus pesos sinápticos W . En el aprendizaje supervisado se pretende estimar una cierta función multivariable desconocida $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (la que representa la red neuronal) a partir de muestras (x,y) ($x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$) tomadas aleatoriamente, por medio de la minimización iterativa de $E(W)$ mediante una aproximación estocástica. Hablando en términos menos formales, en el aprendizaje supervisado se presenta a la red un conjunto de patrones, junto con la salida deseada u objetivo, e iterativamente ésta ajusta sus pesos hasta que su salida tiende a ser la deseada, utilizando para ello información detallada del error que comete en cada paso. De este modo, la red es capaz de estimar relaciones entrada/salida sin necesidad de proponer una cierta forma funcional de partida.
- **Aprendizaje no supervisado o autoorganizado.** El aprendizaje no supervisado se puede describir genéricamente como la estimación de la función de densidad de probabilidad $p(x)$ que describe la distribución de patrones x pertenecientes al espacio de entrada \mathbb{R}^n a partir de muestras (ejemplos). En este tipo de aprendizaje se presentan a la red multitud de patrones sin adjuntar la respuesta que deseamos. La red, por medio de la regla de aprendizaje, estima $p(x)$, a partir de lo cual pueden reconocerse regularidades en el conjunto de entradas, extraer rasgos, o agrupar patrones según similitud (clustering). Un ejemplo típico de modelo que emplea este tipo de aprendizaje es el de los mapas autoorganizados.
- **Aprendizaje híbrido.** En este caso, coexisten en la red los dos tipos básicos de aprendizaje, el supervisado y el no supervisado, los cuales tienen lugar normalmente en distintas capas de neuronas. El modelo de contra-propagación y las redes neuronales de base radial (RBF) son ejemplos de redes que hacen uso de este tipo de aprendizaje.

- **Aprendizaje reforzado** (*reinforcement learning*). Se sitúa a medio camino entre el supervisado y el autoorganizado. Como en el primero de los citados, se emplea información sobre el error cometido, pero en este caso existe una única señal de error, que representa un índice global del rendimiento de la red (solamente le indicamos lo bien o lo mal que está actuando, pero sin proporcionar más detalles). Como en el caso del no supervisado, no se suministra explícitamente la salida deseada.

En el proceso de entrenamiento es importante distinguir entre el nivel de error alcanzado al final de la fase de aprendizaje para el conjunto de datos de entrenamiento, y el error que la red ya entrenada comente ante patrones no utilizados en el aprendizaje, lo cual mide la capacidad de generalización de la red. Interesa más una buena generalización que un error muy pequeño en el entrenamiento, pues ello indicará que la red ha capturado correctamente el *mapping* subyacente de los datos y que no se producido el fenómeno conocido como sobre-ajuste u *overfitting*, que se traduce en un ajuste perfecto a los valores de entrenamiento y una pérdida de la capacidad de generalización, de forma que el empleo de cualquier vector de entradas diferente al utilizado durante el entrenamiento puede generar resultados pobres durante la predicción.

La explicación de este fenómeno es la siguiente. Al principio la red se adapta progresivamente al conjunto de aprendizaje, acomodándose al problema y mejorando la generalización. Sin embargo, en un momento dado el sistema se ajusta demasiado a las particularidades de los patrones empleados durante el entrenamiento, aprendiendo incluso el ruido en ellos presente, por lo que crece el error que comete ante patrones diferentes a los empleados durante el entrenamiento (error de generalización). En este momento la red no ajusta correctamente el *mapping*, sino que simplemente está memorizando los patrones del conjunto de aprendizaje. Idealmente, dada una arquitectura de red neuronal, esta debería entrenarse hasta un punto óptimo en el que el error de generalización es mínimo. El procedimiento consistente en entrenar y validar a la vez para detenerse en el punto óptimo se denomina validación cruzada (*cross validation*), y es ampliamente utilizado en la fase de desarrollo de las redes neuronales.

No obstante, el procedimiento puede ser engañoso, ya que pueden presentarse varios mínimos para el conjunto de validación, debiendo detener el aprendizaje en el punto óptimo de mínimo error de generalización, y no detenerse en el primer mínimo de validación que aparezca.

De lo expuesto hasta el momento, se deduce que según el tipo de modelo de neurona que se utilice, de la arquitectura o topología de las conexiones y del algoritmo de aprendizaje, surgirán distintos modelos de redes neuronales.

Los dos conceptos que más caracterizan a un modelo neuronal son el tipo de aprendizaje y la arquitectura de la red, y por ello, muchas clasificaciones de redes neuronales se basan en estos aspectos. Un ejemplo es la clasificación propuesta por Martín, 2006, que ha sido la adoptada en este documento.

De esta manera, en primer lugar, se puede realizar una distinción en cuanto al tipo de aprendizaje, por lo que aparece una primera clasificación en modelos supervisados, no supervisados, de aprendizaje híbrido y modelos de aprendizaje reforzado. A su vez, y dentro de cada uno de los grandes grupos, tendremos en cuenta el tipo de topología de la red, por lo que se distinguirá además entre redes realimentadas y redes unidireccionales (no realimentadas).

Como se ha podido comprobar en líneas anteriores, dentro de las redes neuronales existe una gran multitud de modelos y variantes, cuya descripción detallada va más allá del propósito de este documento. Es por ello que a continuación se describirá uno de los modelos más conocidos y estudiados, y que se encuentra entre los más empleados en la predicción de valores futuros de series temporales, el perceptrón multicapa (MLP).

3.4.1.4. Perceptrón multicapa (MLP)

El perceptrón multicapa o MLP (*Multi-Layer Perceptron*) se encuentra dentro del grupo de las redes unidireccionales con aprendizaje supervisado. Dicho aprendizaje suele realizarse mediante el algoritmo denominado retropropagación de errores, BP (*backpropagation*) o alguna de sus variantes o derivados.

Una red neuronal MLP con una única capa oculta es capaz de aproximar hasta el nivel deseado cualquier función continua en un intervalo, por lo que si en lugar de una sola capa, la red neuronal presenta varias capas ocultas, se convierte en un aproximador universal de funciones.

El diseño de una red neuronal MLP lleva asociado un proceso de decisión acerca del número de capas ocultas y del número de neuronas presentes en cada capa. Este proceso conlleva un compromiso entre rendimiento como aproximador de funciones y rendimiento computacional, ya que a mayor número de capas y neuronas mayor será la capacidad de la red de ajustar cualquier función, pero también más elevado será el esfuerzo computacional necesario para entrenarla. Además, se corre el riesgo de sobreentrenar a la red, de forma que tal como se ha comentado, ésta se ajusta perfectamente a los valores que se le proporcionan para el entrenamiento, pero pierde la capacidad de generalización más allá de estos valores.

Tal como ya se ha indicado, el algoritmo de aprendizaje más empleado para el entrenamiento de las redes MLP es el de retropropagación del error (BP en adelante). Esta técnica de aprendizaje consiste, básicamente, en actualizar los pesos de las conexiones entre neuronas proporcionalmente a su contribución a una cierta medida de error hasta obtener un resultado satisfactorio. Existen dos versiones de esta técnica, la conocida como versión *online* y la versión *offline* (o *batch*). En la primera de las versiones, los pesos sinápticos se actualizan tras la presentación de cada patrón o vector de entrada, mientras que en la versión *batch* se lleva a cabo una fase de ejecución para todos y cada uno de los patrones de entrenamiento, se calcula la variación en los pesos debida a cada patrón, se acumulan, y solamente entonces se procede a la actualización de pesos.

El algoritmo BP constituye un método de gran generalidad, lo que presenta ventajas e inconvenientes. Su principal ventaja es que se puede aplicar a multitud de problemas diferentes, proporcionando con frecuencia buenas soluciones con no demasiado tiempo de desarrollo, mientras que un inconveniente del BP es su lentitud de convergencia, precio a pagar por disponer de un método general de ajuste funcional que no requiere (en principio) información apriorística. Sin embargo, se debe tener en cuenta que el BP no requiere tanto esfuerzo computacional como el que sería necesario si se tratasen de

obtener los pesos de la red mediante evaluación directa de las derivadas. Otro inconveniente del BP es que puede incurrir en el denominado sobreajuste (*overfitting*), fenómeno directamente relacionado con la capacidad de generalización de la red a partir de los ejemplos presentados. Por otra parte, debe tenerse en cuenta que el algoritmo BP no garantiza alcanzar el mínimo global de la función error, tan sólo un mínimo local, por lo que el proceso de aprendizaje puede estancarse en uno de estos mínimos locales.

Para evitar el sobreajuste o sobreentrenamiento, existen varias técnicas, como la limitación del número de iteraciones realizadas durante el entrenamiento, la validación cruzada o la adición de ruido blanco sobre las variables de entrada. De todas estas técnicas, la validación cruzada es la más extendida. El método consiste en dividir los datos disponibles en dos grupos: entrenamiento y validación. Los datos del primer grupo se utilizan para ajustar los parámetros de la red a través del proceso de aprendizaje, mientras que los datos del grupo de validación cruzada se presentan a la red tras cada iteración del algoritmo. Así, el proceso de entrenamiento se interrumpirá cuando haya transcurrido un número de iteraciones sin reducción del error de los datos de validación.

En cuanto a su aplicación para predicción de series temporales, la red neuronal MLP es válida por su capacidad de aproximar funciones no lineales, permitiendo incluso diseñar modelos autorregresivos no lineales para los cuales sólo es necesario utilizar como variable de entrada los valores pasados de la variable a predecir (Fernández, 2007).

3.4.2. Lógica difusa

Si las redes neuronales artificiales emulan la estructura del cerebro para reproducir sus capacidades, los sistemas basados en lógica difusa, cuyas bases teóricas fueron enunciadas en 1965 por Lofti A. Zadeh, tratan de emular la forma en que el cerebro razona y piensa.

La denominada lógica borrosa (*fuzzy logic*) permite tratar información imprecisa, como *estatura media*, *presión alta* o *volumen pequeño*, en términos de conjuntos borrosos o difusos. Estos conjuntos difusos se combinan en reglas para definir acciones, como por ejemplo, SI la temperatura es alta ENTONCES el volumen es pequeño. De esta forma,

los sistemas de control basados en lógica difusa combinan unas variables de entrada (definidas en términos de conjuntos difusos), por medio de grupos de reglas que producen uno o varios valores de salida.

En términos más rigurosos, la teoría de conjuntos borrosos parte de la teoría clásica de conjuntos, añadiendo una función de pertenencia al conjunto, definida ésta como un número real entre 0 y 1. Así, se introduce el concepto de conjunto o subconjunto borroso asociado a un determinado valor lingüístico, definido por una palabra, adjetivo o etiqueta lingüística F , que toma valores dentro del universo de discurso U , o los límites entre los que está definido. Para cada conjunto o subconjunto difuso se define una función de pertenencia o inclusión $\mu_A(t)$, que indica el grado en que la variable t está incluida en el concepto representado por la etiqueta F . Esto representa la generalización del concepto clásico de conjunto, en el que la función de pertenencia toma solamente los valores 0 o 1; por el de conjunto difuso, en el que la función puede tomar también valores intermedios.

Según lo visto hasta ahora, se puede afirmar que las funciones de pertenencia son los elementos diferenciadores de la teoría de conjuntos borrosos. Estas funciones (*membership function*), consisten en un conjunto de pares ordenados $F = \{(u, \mu_F(u)) / u \in U\}$ si la variable es discreta, o una función continua si no lo es. Como ya se ha comentado, el valor de $\mu_F(u)$ indica el grado en que el valor u de la variable U está incluida en el concepto representado por la etiqueta F . Las funciones de pertenencia más usuales son las de tipo trapezoidal, triangular, exponencial, gaussiana, sigmoideal y campana.

3.4.2.5. Estructura de un sistema de inferencia difusa (FIS)

Básicamente, un sistema de inferencia difusa está compuesto por cinco bloques (Figura 3.2):

- Un sistema de reglas Si-Entonces
- Una base de datos que define las funciones de pertenencia de los conjuntos difusos empleados en las reglas difusas.

- Una unidad de toma de decisiones que ejecuta las operaciones de inferencia en las reglas.
- Una interfaz de borrosificación (*fuzzyfication*) que transforma los inputs en grados de coincidencia con los valores de las reglas lingüísticas.
- Una interfaz de desborrosificación (*defuzzyfication*) que transforma los resultados difusos en nítidos outputs.

Así, el razonamiento difuso llevado a cabo por los sistemas de inferencia difusa comienza con la comparación de las variables de entrada con las funciones de pertenencia en la parte de las premisas para obtener los valores de pertenencia (o medidas de compatibilidad) de cada una de las etiquetas lingüísticas (borrosificación o *fuzzyfication*).

El siguiente paso será la combinación (a través de un operador T-norm específico, normalmente multiplicación o mínimo) de los valores de pertenencia en la parte de las premisas para obtener el peso de cada una de las reglas, con lo que se generará el consecuente de cada regla en función de su peso. Por último, se produce la agregación de los consecuentes cualificados para producir la salida. (desborrosificación o *defuzzyfication*).

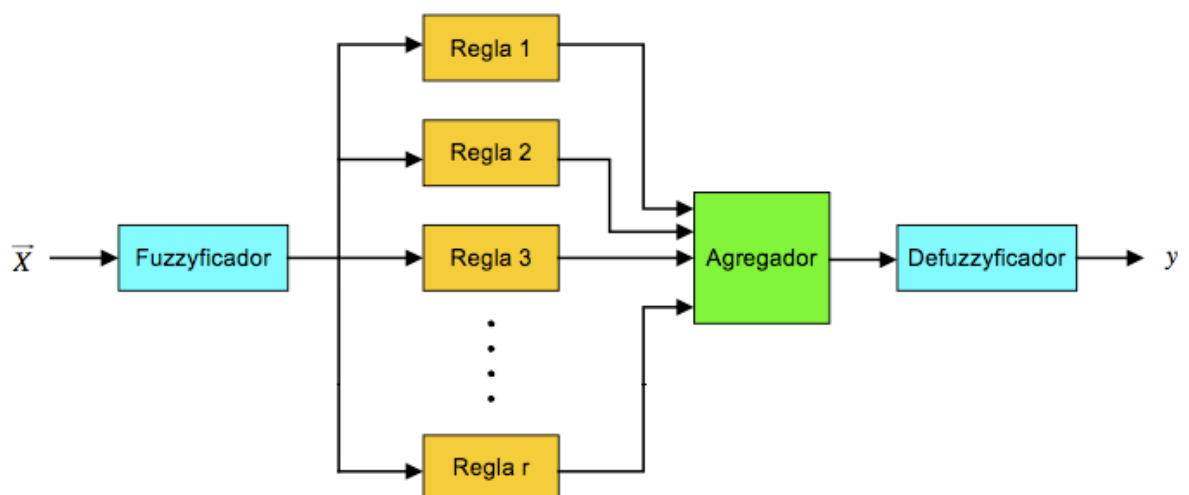


Figura 3.2. Estructura de un sistema de inferencia difusa

La generación de la base de reglas borrosa puede llevarse a cabo empleando dos formatos, el borroso puro o de tipo Mandami y el de tipo Takagi-Sugeno. Este último es uno de los sistemas más utilizados en la actualidad y su característica principal reside en la formulación de la parte consecuente de las reglas, que es una combinación lineal (un polinomio) de las variables de entrada.

En función del grado del polinomio de la parte consecuente de las reglas se puede distinguir entre sistemas de orden cero (es decir, el consecuente es una constante numérica), sistemas de primer orden (polinomios de primer grado), o sistemas de orden superior (polinomios de segundo grado o superior).

Así, las reglas de un sistema de inferencia con dos inputs x_1 y x_2 , y un solo output, y , con una base de reglas del tipo Takagi-Sugeno de primer orden, vendrían representadas por las siguientes ecuaciones:

$$\text{SI } x_1 \text{ es } A_1 \text{ e } x_2 \text{ es } B_1, \text{ ENTONCES } y_1 = p_1x_1 + q_1x_2 + r_1$$

$$\text{SI } x_1 \text{ es } A_2 \text{ e } x_2 \text{ es } B_2, \text{ ENTONCES } y_2 = p_2x_1 + q_2x_2 + r_2$$

donde p_1 , p_2 , q_1 y q_2 son constantes numéricas.

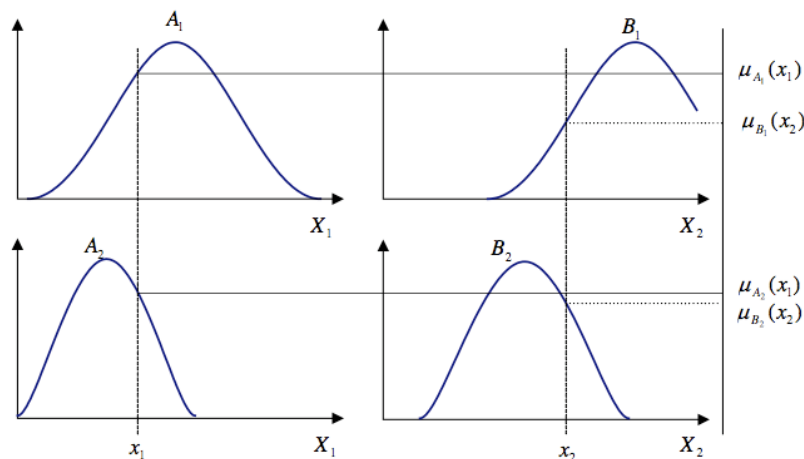


Figura 3.3. Funciones de pertenencia de un sistema de inferencia difusa con dos inputs x_1 y x_2

Y puesto que la salida de cada regla es numérica, el elemento desborrosificador no sería necesario, sino que la salida del sistema de inferencia difusa será la media ponderada de la salida de cada una de las reglas, tal como se expresa en la ecuación 3.2:

$$y_s = \frac{\sum_{i=1}^N w_i y_i}{\sum_{i=1}^N w_i} \quad \text{Ec.3.2.}$$

3.4.3. Sistemas neuro-difusos

Los denominados sistemas neuro-difusos provienen de la combinación de lógica difusa y redes neuronales, y tratan de aunar la capacidad de aprendizaje de éstas últimas, con la facilidad de los sistemas borrosos para trabajar con información incierta y su traducirla a reglas.

La combinación de redes neuronales con sistemas difusos ha resultado en el desarrollo de sistemas artificiales a través de numerosas aproximaciones, tales como GARIC, ANFIS, FUN, NEFCON, FALCON, SONFIN, FINEST, EFuNN, dmEFuNN, diseño evolutivo de sistemas neuro-difusos, y muchos otros. Entre los métodos mencionados, uno de los que optimiza totalmente la fuerza complementaria de las redes neuronales (ANN) y los sistemas de inferencia difusa (FIS) es el sistema adaptativo de inferencia difusa (ANFIS). En comparación con los demás, ANFIS presenta una mayor velocidad de entrenamiento, el algoritmo de aprendizaje más efectivo, y es simple en términos de estructura (Vairappan et al., 2009).

3.4.3.6. Sistema adaptativo de inferencia neuro-difusa (ANFIS)

ANFIS implementa un FIS de tipo Takagi-Sugeno y tiene una arquitectura de cinco capas. La primera capa oculta es para la “borrosificación” de las variables de entrada, mientras que los operadores T-norm son desplegados en la segunda capa oculta para calcular el antecedente de las reglas. La tercera capa oculta normaliza el peso de las reglas, y en la cuarta capa se determinan los consecuentes de las reglas. La capa de salida calcula el input total como el sumatorio de todas las señales entrantes.

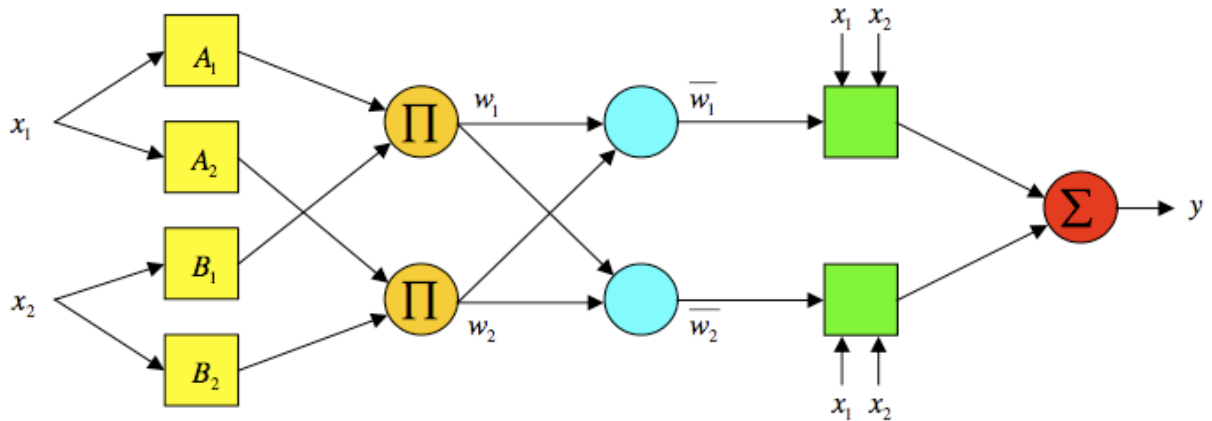


Figura 3.4. Estructura en cinco capas de la arquitectura ANFIS

ANFIS emplea el algoritmo de retropropagación (*backpropagation*) para determinar los parámetros de las premisas (para aprender los parámetros relacionados con las funciones de pertenencia) y la estimación por mínimos cuadrados para determinar los parámetros del consecuente. Un paso en el proceso de aprendizaje tiene dos partes: en la primera de ellas los patrones del input son propagados, y los parámetros del consecuente son estimados mediante un procedimiento de mínimos cuadrados, mientras que los parámetros del grupo de las premisas se asumen fijos para el ciclo actual a lo largo del entrenamiento. En la segunda parte, los patrones son propagados de nuevo y, en este paso, se emplea el algoritmo de retropropagación para modificar los parámetros de las premisas, mientras que los parámetros del consecuente permanecen fijos. Este procedimiento sigue entonces un proceso iterativo.

3.4.3.7. Arquitectura ANFIS

Suponiendo un sistema de inferencia con dos inputs x e y , y un solo output, z , con una base de reglas que contiene dos reglas Si-Entonces del tipo Takagi-Sugeno tenemos:

Regla 1: Si x es A_1 e y es B_1 , entonces $f_1 = p_1x + q_1y + r_1$

Regla 2: Si x es A_2 e y es B_2 , entonces $f_2 = p_2x + q_2y + r_2$

Las funciones de los nodos de este sistema se describen a continuación:

Capa 1: Cada nodo i en esta capa es un nodo cuadrado con una función:

$$O_i^1 = \mu_{A_i}(x) \quad \text{Ec.3.3.}$$

donde x es el input del nodo i , y A_i es la etiqueta lingüística (pequeño, grande, etc.) asociada a esta función de nodo. En otras palabras, O_i^1 es la función de pertenencia de A_i y especifica el grado en el cual x satisface el cuantificador A_i . Normalmente $\mu_{A_i}(x)$ tiene forma de campana, con máximo igual a 1 y mínimo igual a 0, cuya expresión se muestra en la Ec.3.4.

$$\mu_{A_i}(x) = \exp\left\{-\left(\frac{x - c_i}{a_i}\right)^2\right\} \quad \text{Ec.3.4.}$$

donde $\{a_i, b_i, c_i\}$ es el conjunto de parámetros conocido como los parámetros de las premisas.

Capa 2: Cada nodo de esta capa calcula el producto de sus inputs. El output de la capa viene dado por:

$$\omega_i = \mu_{A_i}(x) * \mu_{B_i}(y), i = 1,2. \quad \text{Ec.3.5.}$$

Cada salida del nodo representa el peso de una regla.

Capa 3: Los nodos de esta capa normalizan el peso de las reglas calculando la relación entre el peso de la i -ésima regla y la suma de los pesos de todas las reglas.

$$\overline{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2 \quad \text{Ec.3.6.}$$

Capa 4: Cada nodo i de esta capa es un nodo cuadrado con una función:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad \text{Ec.3.7.}$$

donde \bar{w}_i es el output de la capa 3, y $\{p_i, q_i, r_i\}$ son el conjunto de parámetros.

Capa 5: El nodo de esta capa calcula el output total como la suma de todas las señales entrantes, p.e,

$$O_i^5 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad \text{Ec.3.8.}$$

3.4.3.8. *Algoritmo de aprendizaje híbrido*

El proceso de identificación y optimización de los parámetros en los modelos neuro-difusos puede llevarse a cabo con métodos procedentes de las redes neuronales, como el método del descenso por gradiente en combinación con el proceso de retropropagación. Sin embargo, este método de optimización emplea mucho tiempo hasta alcanzar la convergencia del modelo.

Por otra parte, el método de mínimos cuadrados es una potente herramienta ampliamente utilizada en procesamiento de señales o el campo de la estadística. De hecho, se ha comprobado que este método es esencial e indispensable a la hora de construir modelos matemáticos lineales, y sus conceptos fundamentales pueden extenderse también a los modelos no lineales. De esta forma, el método de descenso por gradiente y la estimación por mínimos cuadrados proporcionan el fundamento matemático más importante a la hora de resolver problemas de modelado neuro-difuso, y su combinación proporciona el algoritmo de aprendizaje híbrido (Jang, 1993), que permite una rápida optimización de los parámetros de la red.

La teoría que se encuentra detrás del aprendizaje híbrido es el ajuste de los parámetros del modelo, que denotados como S , pueden dividirse en dos subgrupos $S1$ y $S2$, donde únicamente los elementos en $S2$ son parámetros lineales. Para unos valores fijados de

S1, los parámetros de S2 se obtienen a partir del método de mínimos cuadrados con el objetivo de minimizar la suma del error cuadrático. Una vez fijados los parámetros en S2, los de S1 son actualizados con el método del descenso de gradiente. La salida del sistema, f se escribe:

$$\begin{aligned}
 f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 = \overline{w_1} f_1 + \overline{w_2} f_2 & \text{Ec.3.9.} \\
 &= (\overline{w_1} x) p_1 + (\overline{w_1} y) q_1 + (\overline{w_1}) r_1 \\
 &+ (\overline{w_2} x) p_2 + (\overline{w_2} y) q_2 + (\overline{w_2}) r_2
 \end{aligned}$$

que es lineal en los parámetros del consecuente ($p_1, q_1, r_1, p_2, q_2, y r_2$).

En resumen, empleando un modelo difuso basado en reglas de tipo Sugeno y combinando la estimación por mínimos cuadrados con el método de descenso por gradiente en la regla de aprendizaje híbrida, la estructura ANFIS es un sistema apropiado para el modelado y la predicción de series temporales no lineales.

3.4.4. Comparación de los tres métodos

Tal como se ha visto, la lógica difusa, las redes neuronales y los sistemas neuro-difusos tienen sus propias características y diferentes implementaciones, pero sin embargo tienen algo en común, y es que todas ellas son válidas para la aproximación de funciones no lineales y el pronóstico de series temporales.

En el caso de los sistemas basados en lógica difusa, la principal ventaja que reportan radica en la facilidad que representa para la mente humana entender el propio modelo. Sin embargo, la limitación a la hora de realizar pronósticos con estos sistemas viene dada por la dificultad de generar reglas difusas apropiadas que sean capaces de generar los valores de salida correctos a partir de los valores de entrada. Por otra parte, los modelos basados en lógica difusa, carecen de la capacidad de aprendizaje de las redes neuronales.

Las redes neuronales por su parte, han sido empleadas en numerosos estudios, comprobándose que son una herramienta efectiva para el pronóstico de series

temporales. En cuanto a las arquitecturas empleadas en este área, una gran parte de los modelos emplean redes feedforward multicapa, tales como MLP, entrenadas empleando la regla de aprendizaje de retropropagación. La mayor dificultad a la que se enfrenta el modelador al emplear redes neuronales es el desarrollo de una red de tamaño adecuado, en lo que al número de nodos de entrada y capas ocultas se refiere, que sea capaz de capturar los patrones subyacentes presentes en el conjunto de datos de entrenamiento. Por otra parte, y a diferencia de los modelos de inferencia difusa, las redes neuronales se comportan como una caja negra, de forma que la extracción de conocimiento a partir de este tipo de modelos es mucho más limitada que en los anteriores.

Los sistemas neuro-difusos permiten obviar las desventajas de las dos técnicas anteriores. Por una parte, y al contrario que los sistemas de inferencia difusa son capaces de aprender, al poder ajustar, por lo general, los parámetros del sistema de forma adaptativa y por otra parte, generan reglas difusas fácilmente interpretables, con lo que el carácter de caja negra de las redes neuronales desaparece en este tipo de modelos. Así, los sistemas neuro-difusos combinan las ventajas que ofrecen las redes neuronales con las de los sistemas de inferencia difusa.

4 METODOLOGÍA

Una vez descrita la problemática del vertido de salmuera y realizada la revisión bibliográfica de las herramientas disponibles para su modelización y predicción de futuros valores de salinidad, en este capítulo se presenta la metodología seguida para el análisis del caso, así como un repaso de las diferentes técnicas disponibles para la construcción de modelos neuro-difusos, que ayudará a la selección de las más adecuadas para el caso de estudio que será presentado en el Capítulo 5.

4.1. Descripción de los agentes del medio físico y de explotación que participan en el modo de fallo principal: superación de los umbrales de salinidad

La ROM 0.0 establece que los factores de proyecto son un conjunto de parámetros, agentes y acciones con los que se definen y se comprueban la seguridad, el servicio y la explotación de una obra y su entorno. Agente es todo lo que puede ejercer o producir en o sobre la obra y su entorno efectos significativos en la fiabilidad, funcionalidad y operatividad. Acción es cualquier efecto que un agente puede producir en la obra y su entorno como resultado de su mutua interferencia.

Entre los parámetros que definen una obra tal como una desalinizadora se encuentran:

- Salinidad y temperatura de las aguas receptoras. La salinidad y temperatura del agua de mar constituyen los **parámetros físico-químicos** que definen el medio receptor. Son responsables de la densidad del mismo y así de la generación de gradientes, lo que en ocasiones supone la existencia de estratificación en el medio receptor. Cuando se da esta situación, y la temperatura es menor por debajo del nivel de la termoclina, la pluma puede separarse del fondo y desplazarse a una profundidad intermedia de la columna de agua, ya que su densidad será intermedia entre la presente por debajo de la termoclina y la de la superficie, lo que puede dar una percepción errónea del grado de dilución alcanzado, ya que los instrumentos de medida situados en el fondo proporcionarán valores menores de salinidad.
- Características del fondo marino. La geomorfología de los fondos es un **parámetro del terreno** que constituye un factor primordial en el control de la trayectoria de

desplazamiento de la pluma y en su dilución. Debido a su elevada densidad, este tipo de vertidos tiende a hundirse hacia el fondo, donde se desplaza a favor de la pendiente negativa. Factores como la pendiente y la rugosidad del fondo afectan a la velocidad a la cual se desplaza la pluma hipersalina, que a su vez es responsable del grado de dilución de la misma. Así, las mayores pendientes generan una mayor dilución, mientras que un elevado coeficiente de rugosidad incrementa la fricción de la pluma hipersalina con el fondo, disminuyendo su velocidad y así, el grado de dilución (Kämpf et al., 2009).

Los **agentes del medio físico** determinantes en el comportamiento de las plumas hipersalinas son, las corrientes, la estratificación del medio y el oleaje. Estas variables determinan tanto la capacidad de mezcla del medio receptor como la trayectoria de la pluma.

- Corrientes. Las corrientes de fondo, generadas por diferentes forzamientos (viento, mareas, oleaje, etc.) influyen, junto con la pendiente del fondo, en el desplazamiento de la pluma de salmuera. En función de la dirección y sentido de la corriente, ésta favorecerá en mayor o menor grado la dilución de la pluma hipersalina, siendo esta generalmente mayor en presencia de corrientes transversales (Roberts et al., 1997).

Las corrientes inducidas por las oscilaciones mareales inducen un movimiento oscilatorio del concentrado de salmuera, favoreciendo el mezclado del efluente con el agua receptora. Los flujos mareales desaparecen dos veces por cada ciclo de marea. Estos momentos de baja actividad favorecen la acumulación de la salmuera cerca del punto de descarga, así como pulsos subsecuentes de plumas hiperdensas a favor de pendiente.

Las corrientes inducidas por el oleaje dominan en la zona de rompientes, donde se establecen corrientes paralelas a la línea de costa en la dirección del gradiente de energía del oleaje. La magnitud de este tipo de corrientes se ve incrementada con la altura de ola y la oblicuidad del oleaje respecto a la línea de costa.

Las corrientes inducidas por el viento superficial, aunque en general son pequeñas comparadas con las debidas a oleaje y mareas, pueden ser importantes en la región comprendida entre la zona de rotura y el límite de la plataforma continental.

- Oleaje. Además de ser responsable de la generación de corrientes en la zona de rotura, el movimiento oscilatorio inducido por el oleaje aumenta la turbulencia, favoreciendo la dilución en campo cercano. Por otra parte, observaciones procedentes de una campaña de campo llevada a cabo por SIDMAR entre diciembre de 2008 y febrero de 2009 muestran una disminución de la salinidad en el fondo tras un evento de tormenta ($H_s > 1\text{m}$) de nueve horas de duración, medido en Enero de 2009. El descenso de la salinidad de fondo registrado fue de 1.3 psu, en un punto situado a unos 500 m del punto de descarga, fuera de la zona de surf. Estas observaciones muestran que el oleaje, o al menos los eventos de alta energía, es posible que influyan en la dilución de la pluma hipersalina también fuera de la zona de rotura.
- Variaciones de nivel del mar. Además de ser responsables de la generación de corrientes periódicas, las oscilaciones del nivel del mar, entendiendo estas como un resultado combinado de la marea astronómica, la marea meteorológica y el oleaje, suponen variaciones de la superficie libre que modifican la altura de la columna de agua de forma local y así la lámina de agua disponible para la dilución de la salmuera.
- Viento. La acción del viento genera un desplazamiento aleatorio de la superficie libre del mar que se conoce como mar de viento u oleaje de tipo SEA, y que presenta generalmente ondas muy peraltadas con periodos y longitudes de onda pequeños, aunque en una amplia gama de frecuencias. El viento produce la mezcla de la capa más superficial hasta unos 10-20 m de profundidad, constituyendo por tanto un elemento que puede incrementar la capacidad de mezcla de la columna de agua.

En cuanto a los **agentes de explotación** de las plantas desalinizadoras, los factores a tener en cuenta son tanto los relativos a las características de efluente (salinidad, temperatura y caudal), como a las del agua empleada para la dilución de éste. La

salinidad y temperatura que presenta el efluente de salmuera al ser vertido al medio, determinarán la flotabilidad de la pluma hipersalina. Tal como ya se ha visto, las plantas desalinizadoras que emplean la técnica de ósmosis inversa generan aguas de rechazo con una salinidad media de 67 psu y una temperatura igual a la del medio, características que determinan la elevada densidad de estos vertidos y que los dotan de una flotabilidad negativa.

De esta forma, tenemos un agente de uso y explotación, el vertido hipersalino resultante del proceso desalinizador, y unos parámetros y agentes que interaccionan con este vertido y condicionan el comportamiento del mismo en el medio receptor. Estos parámetros y agentes son por tanto, los factores clave a tener en cuenta durante el desarrollo de la ecuación de verificación, constituida por ANFIS, con la que se determinará si se produce el modo de fallo principal (superación de los umbrales críticos de salinidad establecidos para las especies sensibles presentes en el área de estudio).

4.2. Desarrollo del modelo predictivo basado en ANFIS

La predicción de la salinidad en áreas afectadas por vertidos hipersalinos es un tema muy importante en lo que a protección del medio marino se refiere. La alerta temprana de eventos de superación de los umbrales de tolerancia marcados para especies como *Posidonia oceanica* puede ayudar a prevenir y mitigar los efectos negativos del vertido de salmuera sobre dicho ecosistema.

El principal objetivo de los pronósticos de salinidad es generar predicciones fiables acerca de valores futuros desconocidos y poder así proporcionar alertas con un tiempo de antelación suficiente, que permita a los responsables de las plantas desalinizadoras tomar decisiones acerca del protocolo de gestión del vertido a aplicar (aumento de la dilución, número de bastidores en funcionamiento, etc.), en función de las condiciones meteorológicas.

Así, un modelo de predicción junto con la adquisición en tiempo real de datos de variables océano-meteorológicas, puede convertirse en una herramienta estándar para propósitos de predicción de valores de salinidad. En la actualidad, la predicción se

realiza a través de modelos numéricos, con gran número de variables implicadas y una gran incertidumbre, derivada de las limitaciones que estos presentan. Por otra parte, y tal como ya se ha comentado, estos modelos se utilizan más bien para el diseño del vertido o la caracterización apriorística del efluente de salmuera, y no durante la fase de explotación de la planta, cuando las excedencias sobre los umbrales de salinidad y otros impactos sobre el medio son detectados a posteriori, aplicándose por tanto medidas con carácter corrector. Sin embargo, una herramienta de predicción fiable permitiría que estas medidas pasasen a tener un carácter preventivo, evitándose así, tanto el efecto de las superaciones de los límites críticos de especies como *Posidonia oceanica*, como los posibles efectos de la persistencia de los excesos de salinidad sobre el medio.

Modernos métodos de inteligencia artificial tales como los sistemas difusos y las redes neuronales pueden ser utilizados para la predicción de valores futuros de series temporales. Estos métodos proporcionan soluciones rápidas, baratas y lo suficientemente buenas, que además permiten tratar datos con ruido y de sistemas dinámicos, algo especialmente importante cuando las relaciones físicas subyacentes son complejas o no totalmente conocidas. El objetivo de la simulación realizada con estas técnicas es emplear el modelo ajustado para generar una secuencia de observaciones estocásticamente equivalentes que probablemente se den en el futuro.

Las reglas del sistema difuso, que servirán para describir la evolución de las series temporales de salinidad, pueden ser construidas a partir de observaciones de estados previos o de otros agentes influyentes tales como el oleaje, el viento o las condiciones de explotación de la planta desalinizadora. Este tipo de modelo puede ser empleado para la simulación y predicción de observaciones futuras desconocidas a priori. Además, con este tipo de modelado los datos imprecisos también pueden ser tenidos en cuenta, lo cual es una tarea difícil en los modelos tradicionales. El rendimiento del sistema difuso puede mejorarse empleando un sistema adaptativo tal como ANFIS, que actualiza los parámetros de las funciones de pertenencia aplicando una combinación de mínimos cuadrados y de retropropagación del error.

4.2.1. Selección de las variables de entrada

Una de las fases más importantes y laboriosas del proceso de diseño de un sistema de predicción basado en redes neuronales es la selección de las variables de entrada. Normalmente, no todas las variables consideradas como entradas potenciales al modelo serán igualmente informativas, ya que algunas pueden estar correlacionadas entre sí, proporcionando información redundante, ser fuentes de ruido, o tener una pequeña relación con la variable de salida que está siendo modelada. Sin embargo, pocos son los artículos científicos que incluyen una metodología de selección de variables como parte importante de su estudio (Maier et al., 2000), lo que puede conducir a ciertos defectos en el diseño, ya que a medida que se incrementan las variables de entrada, la complejidad computacional y los requerimientos de memoria del modelo aumentan, el aprendizaje se vuelve más complicado y el número de mínimos locales presentes en la superficie de error crece.

Además, la inclusión de más variables explicativas que las estrictamente imprescindibles, genera la necesidad de disponer de una serie de datos más extensa, debido a lo que se conoce como maldición de la dimensionalidad (*curse of dimensionality*), que implica que el número de datos disponible debe crecer exponencialmente con el número de variables de entrada.

Sin un conocimiento apriorístico de la influencia de las variables candidatas a ser las entradas al modelo, podríamos intentar construir tantos modelos como fueran posibles, considerando en cada uno de ellos diferentes combinaciones de dichas variables. De esta forma, podríamos determinar empíricamente el modelo con la mejor combinación de entradas, siendo éste el que presentando un error de predicción menor, también haya requerido menor tiempo de ajuste o entrenamiento. Para d entradas potenciales al modelo, existen $2^d - 1$ conjuntos de entradas posibles, por lo que para valores bajos de d , es factible generar todos los modelos y comprobar de forma empírica cual es el más adecuado. Sin embargo, para valores mayores de d , como es normalmente el caso de los problemas complejos que se tratan con redes neuronales, el número de combinaciones posibles crece exponencialmente, con lo que un algoritmo de selección eficiente es algo necesario.

Normalmente, el primer conjunto de variables candidatas para ser empleadas como variables de entrada, se selecciona a partir del conocimiento previo del sistema por parte del modelador. Sin embargo, y pese a que una selección a partir del conocimiento previo del sistema, es un paso necesario para definir un conjunto de variables candidatas, empleado en muchos de los trabajos realizados con redes neuronales, constituye un sistema dependiente del conocimiento del experto, y por tanto, es muy subjetivo y dependiente del caso. Así, la mejor aproximación para la determinación de las variables de entrada, incluye una combinación de conocimiento previo y aproximaciones analíticas (Fernando et al., 1998; Maier et al., 1997; Maier et al., 1998).

Entre las técnicas analíticas más empleadas para la selección de las variables adecuadas para la modelación del sistema en estudio se encuentran los métodos basados en el análisis de correlaciones y autocorrelaciones entre las variables de entrada, o en el análisis de componentes principales. Sin embargo, estas técnicas únicamente son capaces de detectar relaciones lineales entre dos variables, siendo relativamente poco adecuadas para sistemas no lineales.

Otro enfoque clásico para la selección de variables es la aproximación heurística, que consiste en la selección ordenada de variables. Este método se utiliza normalmente para evitar la consideración de todos los subconjuntos de variables. Los dos métodos estándar dentro de la aproximación heurística son la selección "hacia delante" y "hacia atrás". La selección "hacia delante" es la más comúnmente utilizada, y comienza con la búsqueda de la mejor variable de entrada individual (sólo una variable por modelo) y su selección para el modelo final. En el paso siguiente, se construyen, con la variable anterior y otra variable más, todos los modelos posibles con dos entradas. La variable correspondiente al modelo con menor error conforma la segunda variable seleccionada. El proceso continúa construyendo todos los modelos posibles con tres entradas, y escogiendo la tercera variable de acuerdo al criterio seguido con las dos primeras. El proceso finaliza cuando no se obtiene mejora en los resultados como consecuencia de añadir una nueva variable de entrada.

Por otro lado, la eliminación "hacia atrás" comienza con el conjunto total de posibles variables, y de forma secuencial va eliminando aquellas variables cuya exclusión disminuye el rendimiento del modelo en menor grado. La mayor desventaja de este tipo

de aproximaciones es que están basadas en prueba-error, y por tanto, no existe garantía de que sean capaces de encontrar el mejor subconjunto global de variables de entrada.

Otra alternativa se basa en los análisis de sensibilidad, que constituyen un método muy común de extraer información a partir de una RN entrenada. En primer lugar, se construye un modelo utilizando como entradas todas las variables disponibles. Una vez ajustado el modelo se prueba su funcionamiento haciendo que el valor de todas las entradas coincida con su valor medio respectivo, excepto para una de las entradas, en la que su valor varía ligeramente en torno a su valor medio. Las variables de entrada que con su variación en torno a la media producen mayores variaciones en la salida, son aquellas que tienen un mayor peso relativo en el valor de la misma. Aquellas otras entradas cuya variación respecto a su valor medio apenas tiene influencia en la variación del valor de salida, denotando una baja sensibilidad, podrían eliminarse como posibles variables de entrada.

También es posible utilizar combinaciones de los métodos anteriores. Por ejemplo, Scheleiter (1999) empleó varios métodos para la determinación de las variables de entrada de un modelo de calidad de aguas, bioindicación y dinámica de poblaciones lóxicas en corrientes, algunos de los cuales eran la correlación de Pearson, el análisis de regresión hacia delante y el análisis de sensibilidad.

4.2.2. División de los datos

La práctica común consiste en dividir los datos disponibles en dos subconjuntos, el conjunto de entrenamiento, utilizado para construir el modelo, y un conjunto independiente de datos de validación, para estimar el rendimiento del mismo. Normalmente, las redes neuronales no son capaces de extrapolar por encima del rango de los datos empleados durante el entrenamiento (Minns et al., 1996; Hettiarachchi et al., 2005), con lo que cabe esperar predicciones pobres cuando el set de datos de validación contiene valores fuera del rango de aquellos empleados durante el entrenamiento. Por este motivo, la división de los datos merece especial atención, siendo necesario que ambos conjuntos sean representativos de la misma población, con unas propiedades estadísticas similares (media, mediana, desviación típica..) y valores dentro de un mismo rango.

Tal como ya se ha visto al repasar la teoría de redes neuronales, el grupo de datos de validación se emplea para comprobar la capacidad de generalización del modelo, asegurando que no se ha producido sobreentrenamiento, fenómeno responsable de que el modelo se ajuste a la perfección a los datos de entrenamiento pero proporcione pobres resultados cuando se le presenta cualquier otro vector de entradas diferente al empleado durante el entrenamiento. Normalmente, se utilizan dos tercios de los datos para el entrenamiento del modelo y un tercio para validación (Hammerstrom, 1993).

Una modificación de la división propuesta anteriormente se produce cuando se emplea el método de validación cruzada (Stone, 1974) para determinar cuando finalizar el entrenamiento y comprobar la capacidad de generalización del modelo. En este caso, los datos deben ser divididos en tres conjuntos: entrenamiento, validación y test. Al igual que antes, el set de datos de entrenamiento se emplea para ajustar el modelo, y el grupo empleado para determinar cuando se debe finalizar el entrenamiento y evitar el fenómeno de sobre-ajuste o sobreentrenamiento se denomina conjunto de datos de validación. El tercer grupo de datos es el grupo de datos de test, y es empleado para comprobar la capacidad de generalización de la red ya entrenada. Este debe ser un grupo de datos independiente, pues no debe ser utilizado como parte del proceso de entrenamiento. Con el objetivo de encontrar la proporción óptima de datos en cada uno de los conjuntos (entrenamiento, validación y test), Shanin et al., 2004b investigaron el impacto que, la proporción de datos incluidos en cada uno de los diferentes grupos de datos, tenía en el rendimiento de una red neuronal. La conclusión a la que llegaron es, que no existe una relación clara entre la proporción de datos empleada para entrenamiento, validación y test y los resultados del modelo. Sin embargo, comprobaron que los mejores resultados se obtenían cuando el 20% de los datos se empleaba para validación o test (comprobación de la capacidad de generalización del modelo ya entrenado) y el resto era dividido en 70% para entrenamiento y 20% para validación (control del fenómeno de sobre-entrenamiento).

4.2.3. Transformación de las variables de entrada

Un paso a considerar durante el desarrollo de un modelo basado en redes neuronales es la selección de un método apropiado de transformación de los datos. En general, las

variables de entrada al modelo presentan unos rangos y características muy diferentes entre sí, con lo que en ocasiones, puede ser de utilidad la aplicación de alguna técnica de normalización para que todas las variables puedan recibir igual atención durante el proceso de entrenamiento. Por otra parte, recientemente se ha señalado que, puesto que el error cuadrático medio (RMSE) es comúnmente utilizado como la función objetivo al entrenar una red neuronal, es necesario que los datos se encuentren normalmente distribuidos para asegurar resultados óptimos, ya que esta medida corresponde a una estimación de máxima probabilidad únicamente bajo la hipótesis de normalidad (Fortin et al., 1997). Sin embargo, este hecho no ha sido confirmado con pruebas empíricas, sino que éstas han demostrado que existen modelos que se ajustan igualmente, sin importar el hecho de que los datos estén normalizados o no (Faraway y Chatfield, 1998).

Además de las técnicas de normalización, existen otros métodos de transformación de datos, como la estandarización estadística o la aplicación de funciones matemáticas. La mejora de la respuesta de las redes neuronales con la aplicación de algunas de estas técnicas fue estudiada por Bowden et al. (2003), que encontraron que, los modelos desarrollados empleando técnicas de transformación lineales, proporcionaban mejores resultados que aquellos desarrollados empleando transformaciones logarítmicas, de eliminación de la estacionalidad, o de normalización. Estos resultados parecen ser además consistentes con los encontrados por Faraway y Chatefield (1998), cuyos modelos no mostraban mejoras significativas con la aplicación de una transformación logarítmica. Además, la capacidad de predicción de la red se veía mermada al eliminar la estacionalidad de los datos.

Estos hallazgos refuerzan la creencia popular de que al utilizar redes neuronales (en particular aquellas que emplean el método de retropropagación del error, que ajustan los datos sin asumir ninguna forma funcional), no es necesario que los datos se distribuyan de alguna forma en particular.

Entre las transformaciones lineales, la más sencilla consiste en escalar todas las variables de entrada entre 0 y 1. Esta transformación viene dada por la ecuación 4.1:

$$x_{j,norm} = \frac{x_j}{\max(x_k)_{k=1\dots N}} \quad \text{Ec.4.1.}$$

Otra técnica de transformación consiste en ajustar el valor mínimo de la serie a 0 y el máximo a 1. Este cálculo viene dado por:

$$x_{j,norm} = \frac{x_j - \min(x_k)_{k=1\dots N}}{\max(x_k)_{k=1\dots N} - \min(x_k)_{k=1\dots N}} \quad \text{Ec.4.2.}$$

Además del escalado entre 0 y 1, también es común ajustar los datos entre -1 y 1, cuando la función de activación de la red neuronal es hiperbólica. Sin embargo, es posible que, el conjunto de datos de entrenamiento, no incluya todos los valores posibles que las variables pueden tomar, con lo que puede ser recomendable dejar un cierto margen, y ajustar los datos entre +0.1 y +0.9 o -0.9 y +0.9. La primera de estas transformaciones, vendrá dada por:

$$x_{j,norm} = 0.8 \frac{x_j - \min(x_k)_{k=1\dots N}}{\max(x_k)_{k=1\dots N} - \min(x_k)_{k=1\dots N}} + 0.1 \quad \text{Ec.4.3.}$$

4.2.4. Calibración del modelo

Una vez han sido seleccionadas y transformadas las variables de entrada al modelo, el siguiente paso será la calibración del mismo. Este proceso es equivalente a la fase de estimación de parámetros de los modelos estadísticos convencionales (Maier et al., 2000).

En un modelo neuro-difuso tal como ANFIS, el proceso de calibración supone la identificación de la estructura y la estimación de parámetros del modelo. La identificación de la estructura consiste, en resolver el número de funciones de pertenencia necesarias para definir correctamente cada una de las variables de entrada, así como las iteraciones necesarias para alcanzar una determinada función objetivo, generalmente la raíz del error cuadrático medio (RMSE). Las principales ventajas de

utilizar el RMSE como medida del error que debe ser minimizado durante el entrenamiento son: que se calcula de forma sencilla y que penaliza los errores altos.

El procedimiento comienza con la selección de uno de los métodos disponibles para la partición del espacio de las variables de entrada. Estos métodos son la "partición en rejilla", que realiza una partición de los datos generando una serie de reglas equiespaciadas, o el agrupamiento substractivo, que calcula los grupos de datos relacionados entre sí, y selecciona los lugares del espacio de entrada con mayor densidad de puntos, como centros de los grupos o clusters. El primer método se aplica en sistemas con puntos distribuidos por todo el espacio, mientras que el agrupamiento substractivo es recomendable en el caso de que existan zonas con alta densidad de puntos o cuando el número de reglas generadas con el otro método resulte excesivo.

En caso de haber seleccionado el método de partición en rejilla, antes de generar el FIS (Sistema de Inferencia Difusa), será necesario elegir el tipo de función de pertenencia para las variables de entrada. La elección de la forma de dichas funciones, así como el número de las mismas, se justifica según el error cuadrático medio obtenido y el número de parámetros no lineales resultantes, de forma que es conveniente probar todos los tipos posibles y elaborar una tabla comparativa. Si por el contrario, el método seleccionado es el agrupamiento substractivo, el parámetro a ajustar por el modelador es el llamado radio de vecindad, que especifica el rango de influencia del centro del cluster, en cada una de las dimensiones de los datos. Un valor pequeño del radio de vecindad puede generar un número excesivo de clusters y desembocar en problemas de sobreentrenamiento, mientras que con un valor demasiado alto se generarán pocos clusters, pudiendo resultar un modelo muy poco ajustado, incapaz de modelar el proceso adecuadamente. Según lo indicado en el manual de Fuzzy Logic de MATLAB, un valor entre 0.2 y 0.5 suele ser lo más apropiado.

El siguiente paso será el ajuste de los parámetros, tanto los del antecedente (constantes que caracterizan las funciones de pertenencia) como los del consecuente (parámetros lineales de la salida del modelo). Para ello, es posible emplear el método de retropropagación del error (*backpropagation*), o el método de aprendizaje híbrido (Jang, 1995), métodos ya descritos en el Capítulo 3.

Una vez identificados los parámetros del consecuente, la suma de errores cuadráticos procede de la diferencia entre la salida de la red y la salida deseada, presentada en los pares de entrenamiento. Cuando se alcanza un error estable y razonable, se valida la predicción frente a los datos destinados a la tarea de validación cruzada, para asegurar que no se está produciendo sobreentrenamiento. El número de parámetros resultantes también se emplea como criterio para elegir la estructura, ya que el número total no debería exceder la sexta parte del número de datos procedentes del archivo principal.

4.2.5. Validación y fiabilidad de las predicciones

Una vez completada la fase de calibración, el rendimiento de la red ya entrenada debe ser validado, empleando un conjunto de datos independiente y a partir de alguna de las técnicas disponibles para este propósito. Cabe destacar la importancia de que los datos de validación no hayan sido empleados como parte del entrenamiento de ninguna forma. Si la diferencia en el error obtenido con los datos de validación es notablemente diferente que el obtenido con los datos empleados durante el entrenamiento, es muy probable que los dos conjuntos de datos no sean representativos de la misma población, o que el modelo haya sido sobreentrenado (Masters, 1993). Un resultado pobre durante esta etapa puede deberse también a un pre-procesado incorrecto de los datos, a la falta del mismo, a una mala selección de las funciones de pertenencia, y en definitiva, a cualquier defecto en el tratamiento de los datos o en la identificación de la estructura y los parámetros de la red. A este respecto, existen herramientas para el diagnóstico de la idoneidad de modelos estadísticos, que también pueden ser utilizadas en las redes neuronales. Por ejemplo, Hsu et al. (1995), Lek et al. (1996) o Marcé et al., estudiaron la normalidad, correlación y tendencias de los residuos, con el objetivo de evaluar la adecuación del modelo seleccionado.

5 APLICACIÓN DE LA PROPUESTA A PLANTAS DESALADORAS REALES: PLANTAS I Y II DEL CANAL DE ALICANTE

En este capítulo se presenta el caso de estudio, en el cual se analizará la capacidad del modelo neuro-difuso ANFIS, para la predicción de valores futuros de salinidad y así la idoneidad para su integración como parte del proyecto ASDECO. En primer lugar se describe el caso concreto y peculiaridades, para finalmente presentar los resultados computacionales obtenidos con el modelo desarrollado a partir de las técnicas y directrices expuestas en el Capítulo 4 de metodología.

5.1. Introducción.

La planta desalinizadora que servirá como "Planta Piloto" para este estudio, es la del Canal de Alicante, considerando como tal las plantas Alicante I, Ampliación de Alicante I y Alicante II, que se estudian conjuntamente ya que realizan el vertido de sus aguas de rechazo de esta misma forma, en la Cala de los Borrachos. El motivo principal para la selección de esta desalinizadora, es que se trata de la planta piloto del proyecto ASDECO, marco de desarrollo de esta tesina. Asimismo, el nombrado proyecto seleccionó esta planta básicamente porque dispone de una importante fuente de información procedente de los planes de vigilancia que la Universidad de Alicante había realizado en la zona, y que permitió disponer de datos de partida con los que empezar a trabajar.

La desaladora del **Canal de Alicante** inaugurada en 2003, proporciona anualmente 50.000 m³ de agua potable a la red de la Mancomunidad de los Canales del Taibilla, un caudal que abastece a una población aproximada de 600.000 personas que residen en Alicante, Elche, Santa Pola y San Vicente de Raspeig, municipios adscritos. Esta desaladora, ubicada en el paraje de Aguamarga, en el término municipal de Alicante, emplea la técnica de ósmosis inversa.

La planta, diseñada para producir 50.000 m³ de agua potable al día, capta el agua del mar a través de una red de 18 pozos costeros, desde donde se eleva a las instalaciones. Para realizar el proceso de ósmosis inversa, se cuenta con siete bastidores de

membranas y otros tantos grupos de alta presión para suministro y posterior recuperación de energía, que se encargan de separar la sal del agua. El caudal ya desalado, se somete a un postratamiento para que cumpla todos los requisitos establecidos por la normativa vigente para las aguas de consumo humano. Finalmente, una vez el agua es apta para su consumo humano, el caudal se impulsa a través de una conducción de 22 kilómetros hasta un depósito regulador ubicado en las proximidades de Elche, para su posterior entrega a los canales del Taibilla.

Por otro lado, el agua de rechazo que resulta del proceso, es conducida a un punto de la costa situado a un kilómetro de la desaladora, conocido como la cala de los Borrachos, donde se vierte al mar (vertido que hasta la construcción de la ampliación de Alicante I se realizaba de forma directa, sin ningún tipo de dilución).

La desaladora de **Alicante I (ampliación)** está puesta en servicio desde 2005 y se incluye dentro de la actuación “Ampliación de la desaladora de la Mancomunidad de los Canales de Taibilla en Alicante”, contemplada entre las Prioritarias y Urgentes de la Ley 11/2005 de 22 de junio.

La ampliación de la desaladora del Canal de Alicante I responde al concepto +Agua para un desarrollo+sostenible del Programa A.G.U.A., ya que incluirá todas las obras e instalaciones precisas para incrementar la producción de 18-24 hm³/año, y contribuirá a garantizar, junto con el resto de las actuaciones de desalación previstas, el abastecimiento actual y futuro de una población estable superior a 2.200.000 habitantes, que en época estival supera los 3.000.000 de habitantes. La ampliación propuesta es la toma adicional de agua mediante una batería de 11 pozos situados en la zona de servidumbre de protección del dominio público marítimo-terrestre. La producción de agua desalada se consigue igualmente, mediante la técnica de ósmosis inversa.

A unos 180 metros al Norte de la Cala de los Borrachos, se encuentran las obras de captación de agua de mar y su impulsión hasta la actual obra de vertido. El objetivo de esta captación de agua es conseguir una dilución, no inferior a 2 partes de agua de mar y 1 parte del agua de rechazo, previa al vertido de esta última.

Esta ampliación supone un incremento de la capacidad de producción de la desaladora actual de los 50.000 m³/día hasta los 65.000 m³/día de agua dulce apta para el consumo humano.

Por último, la nueva desalinizadora del Canal de Alicante (**Alicante II**) se sitúa en las inmediaciones de la actual, en el término municipal de Alicante e incluye las obras e instalaciones precisas para aportar al sistema hidráulico de la Mancomunidad de los Canales del Taibilla un volumen anual de 24 hm³ de agua desalinizada.

La toma de agua de mar se realiza mediante 65 drenes oblicuos de una longitud media de 25 metros, ejecutados desde un túnel paralelo a la línea de costa, conectados a un pozo de bombeo hasta la planta desalinizadora. El vertido del agua de rechazo (salmuera) se realiza mediante conexión a la conducción de vertido de la desalinizadora Alicante I.

5.2. Descripción del área de estudio.

El complejo de desaladoras descrito se encuentra ubicado en el paraje de Agua Amarga, en el término municipal de Alicante, junto a la carretera N-332. Dichas instalaciones realizan el vertido de sus aguas de rechazo en la Cala de los Borrachos, que forma parte de la playa de Agua Amarga, ubicada dentro del LIC (Lugar de Interés Comunitario) «Illa de Tabarca» (Figura 5.1), que incluye diversos hábitats prioritarios y especies en peligro de extinción.

La batimetría del área de estudio se muestra en la Figura 5.1. La pendiente de fondo es 1/70 desde la línea de costa hasta aproximadamente 10 m de profundidad, punto a partir del cual se hace más pronunciada (1/300). Esta pendiente se mantiene hasta los 16 m de profundidad, punto en el que se alcanza la pradera de *Posidonia oceanica*. El gradiente de los contornos batimétricos se encuentra orientado hacia el Sur-Este.

Pese a la existencia del LIC de Illa Tabarca, las presiones a las que se ve sometida la zona son variadas, y entre ellas cabe destacar el vertido de aguas residuales tratadas de la EDAR de Rincón de León, el vertido del matadero, la influencia del puerto de Alicante y su reciente ampliación, el fondeo de grandes barcos, la pesca, y obviamente

el vertido de salmuera de la desaladora del Canal de Alicante. De hecho, según indica la DIA del proyecto de "nueva desaladora del Canal de Alicante" (B.O.E. nº155, 2005), la confluencia de todos estos factores implica que sólo a partir de 1500 - 2000 m de la costa y lejos de las zonas urbanizadas, se encuentren aguas de calidad.

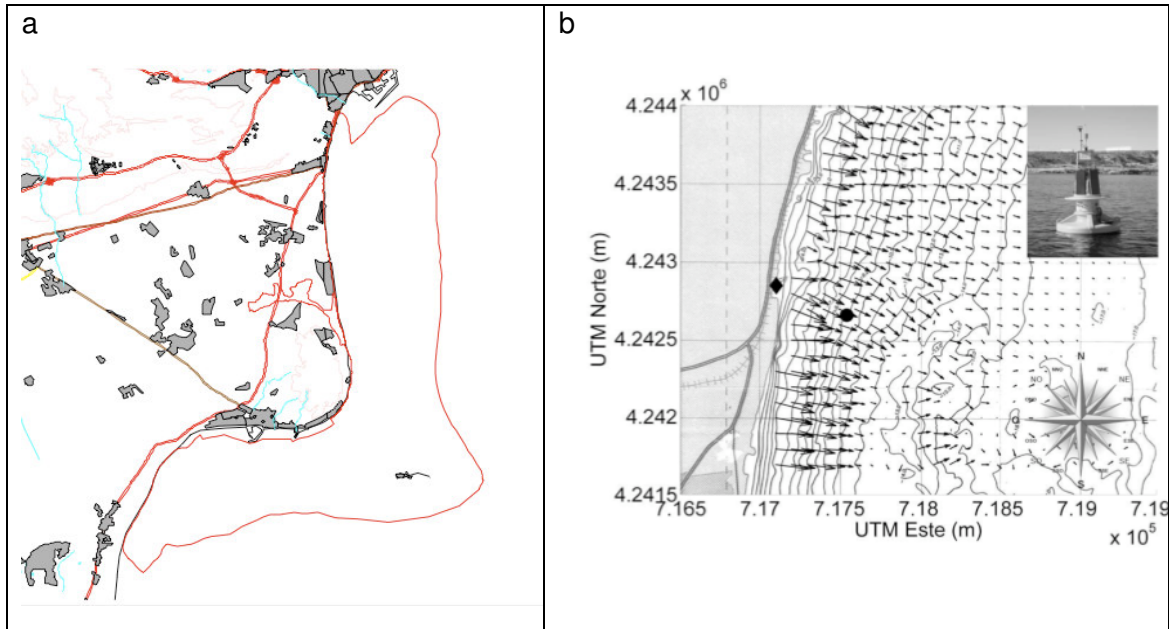


Figura 5.1. a) Plano del LIC. La línea roja marca el límite del LIC. b) Mapa de contornos batimétricos y situación del punto de descarga (v).

5.2.2. Áreas y ecosistemas protegidos

El vertido de las plantas desalinizadoras en estudio se realiza en el ámbito marítimo del **LIC de Illa de Tabarca** (ES5213024). Este presenta una superficie de 14.557 Ha, de las cuales 269 Ha son terreno continental y 14.288 Ha corresponden a superficie marina, e incluye los municipios de Santa Pola, Elx y Alicante, así como el frente litoral de todos ellos.

El LIC es un área básicamente marina, que alberga las que probablemente sean las praderas de *Posidonia oceanica* mejor conservadas del litoral valenciano. La isla de Tabarca, además, alberga diversos hábitats marinos de gran interés.

El hábitat mejor representado son las praderas de *Posidonia oceanica*. En cuanto a especies, las de mayor relevancia están también asociadas al medio marino o lagunar,

con especial mención a aves como *Hydrobates pelagicus*, incluida en el anexo I de la Directiva Aves (*Directiva 79/409/CEE*), o cetáceos como *Tursiops truncatus*, presente en el anexo II de la Directiva Hábitats (*Directiva 92/43/CEE*), especies que sin embargo no se ven afectadas por el efecto del vertido de salmuera.

En cuanto a los **ecosistemas** presentes en el área, el más relevante por la figura de protección bajo la que se encuentra, así como por su relevancia ecológica y fragilidad, son las praderas de *Posidonia oceanica*. La vulnerabilidad de los ecosistemas de *Posidonia oceanica*, se ha puesto de manifiesto con la regresión de la pradera observada a lo largo del periodo durante el cual se ha llevado a cabo la Vigilancia Ambiental de la zona. Los numerosos impactos que sufre la pradera en la zona (vertido de aguas residuales tratadas de la EDAR de Rincón de León, el vertido del matadero, el vertido de la desaladora de Alicante, la influencia del puerto de Alicante y su reciente ampliación, el fondeo de grandes barcos y la pesca de arrastre en la zona más profunda) parecen indicar que la regresión observada, probablemente se deba al efecto acumulado de todos ellos. Por otra parte, la distribución de los impactos observados, junto con la estabilidad de las praderas durante los dos primeros años de funcionamiento de la desaladora, indican que el principal impacto que está acelerando la regresión de la pradera de *Posidonia oceanica* ha sido la ampliación hacia el Sur del puerto de Alicante (Universidad de Alicante, PVA 2007).

5.3. Características del efluente de salmuera y del sistema de vertido

El efluente de salmuera que se descarga en la Cala de los Borrachos, integra los rechazos de las plantas Alicante I y Alicante II. Estos son, la salmuera o rechazo de la osmosis inversa, el agua de lavado de los filtros de arena y los vertidos de limpieza de membranas y drenaje de suelos. El primero es un vertido continuo, mientras que los demás son esporádicos. Además, el efluente contendrá trazas de los aditivos químicos empleados en los procesos de pretratamiento, entre los que se encuentran: hipoclorito sódico (desinfección bacteriológica del agua de captación), cloruro férrico (coagulante), ácido sulfúrico (reducción del pH), hexametáfosfato sódico (se emplea para incrementar la solubilidad de algunas sales y prevenir así la precipitación de las mismas), y bisulfito sódico (eliminación de oxidantes que degradan las membranas).

El rechazo del proceso de osmosis inversa es agua de mar con un contenido en sales directamente proporcional a la tasa de conversión de la planta (45%), que para el caudal de producción previsto para las dos desaladoras supone un vertido de 158.000 m³/día con una salinidad de 68 psu.

26/02/04		
Salinidad	‰	67.7
Conductividad	(mS/cm)	94.3
pH		7.31
CO ₃ ²⁻	(mg/l)	0
HCO ₃ ⁻	(mg/l)	265
PSS	(mg/l)	<1
Pdisueltas	(g/l)	78.6
Cl ⁻	(mg/l)	42639
SO ₄ ²⁻	(mg/l)	5745.0
HPO ₄ ²⁻	(mg/l)	<0.1
NO ₃ ⁻	(mg/l)	1.4
NO ₂ ⁻	(mg/l)	<0.01
K ⁺	(mg/l)	850
Ca ²⁺	(mg/l)	1170
Na ⁺	(mg/l)	21500
Mg ²⁺	(mg/l)	2880.0

Tabla 5.1. Composición del vertido en un muestreo realizado el 26/02/2004. U. Alicante. PVA, 2004.

Las aguas de rechazo descritas no se vierten directamente al medio, sino que se diluyen previamente en diferentes proporciones, de las cuales 2:1 es la indicada en las DIA de las plantas desalinizadoras.

Sobre el proceso de dilución de esta desaladora, cabe destacar que la zona de donde se toma el agua de dilución está afectada por el vertido, de modo que tanto la salinidad de dilución como la salinidad del rechazo varían en el tiempo.

El **sistema de vertido** de los rechazos procedentes de la planta desaladora está formado por varios elementos, que van desde instalaciones para tratamientos previos, hasta la obra de vertido al mar. Estos son:

- Planta de tratamiento de los vertidos procedentes del lavado de filtros y membranas de ósmosis inversa. La instalación está formada por los siguientes elementos:
 - Balsa de regulación y decantación donde se recogen directamente los vertidos procedentes del lavado de filtros y de membranas.

- Sistema de floculación - decantación para el tratamiento del residuo procedente de la balsa anterior.
 - Tratamiento de sólidos sedimentables.
 - Depósito de neutralización de los vertidos procedentes del lavado de membranas.
- Conducción de Poliéster Reforzado con Fibra de Vidrio (PRFV) para un caudal de rechazo de 1,875 m³/s y un diámetro de 1800 mm, que discurre desde la planta de tratamiento del agua de lavado hasta el punto de vertido, situado a 1190 metros en la proximidades de la Cala de los Borrachos. En esta conducción, el flujo se consigue por gravedad, no siendo necesario ningún bombeo.
 - Obra de toma de agua de mar para dilución del vertido salino, consistente en un recinto protegido por escolleras laterales y frontal, y por una cámara de bombeo dotada con tres grupos de motobombas de 2 m³/s de caudal unitario, suficiente para diluir la salmuera en la proporción indicada en la DIA (1:2) y también disponer de un margen de caudal adicional para incrementar la dilución hasta 1:6 en caso preciso. La conexión de la toma con el aliviadero de Cala de los Borrachos se realiza mediante una hinca de 1800 mm de diámetro interior.
 - Obra de vertido en la Cala de los Borrachos, consistente en una arqueta que provoca un vertido sumergido con difusión mediante escollera.

5.4. Caracterización del clima marítimo.

Se define clima marítimo como la caracterización del oleaje en periodos largos de tiempo, o la descripción estadística de la variación en el dominio del tiempo de los estados de mar en un emplazamiento dado, y puede considerarse definido a partir de la estadística unidimensional y bidimensional de los parámetros geométrico-estadísticos y espectrales representativos del estado de mar en la zona considerada. (ROM 0.3.91, Anejo I)

Llegar a la caracterización y previsión del oleaje en aguas profundas es imprescindible para poder abordar cualquier estudio y proyecto de ingeniería marítima (ROM 0.3.91, Anejo I). En el caso de estudio, la caracterización del clima marítimo es básica para conocer la influencia del mismo en el comportamiento del vertido de salmuera.

5.4.1. Oleaje

5.4.1.1. Bases de datos de oleaje

Existen diferentes tipos de fuentes de datos de las cuales extraer la información de oleaje necesaria para realizar la caracterización del clima marítimo. En general, es posible utilizar datos visuales provenientes de las observaciones en barcos de ruta; datos medidos en boyas y datos generados mediante técnicas de reanálisis. A continuación se describen brevemente las características de cada uno de ellos.

- **Datos visuales**

El Centro de Estudios de Puertos y Costas (CEPYC), cuenta con un Banco de Datos Visuales, cuya información es suministrada por el National Climatic Data Center de Asheville (Carolina del Norte, USA). Esta información procede de datos tomados por observadores entrenados desde barcos de tráfico marítimo comercial. Cada observación contiene la siguiente información:

- Longitud y latitud en el punto de observación.
- Fecha y hora del momento de la observación.
- Presión atmosférica y temperatura del aire.
- Velocidad y dirección del viento.
- Altura de ola, período y dirección del oleaje SWELL.
- Altura de ola y período del oleaje SEA. (Se asume que tiene la misma dirección que el viento).

- **Datos instrumentales**

La Dirección General de Puertos y Costas, posee un Banco de Datos Oceanográficos que fue creado para establecer una base de datos de acceso público, que sirviera de apoyo al desarrollo de nuevos proyectos de ingeniería civil en el entorno portuario y costero. Este Banco de Datos dispone de dos tipos de medidas instrumentales de oleaje: la red de boyas de aguas profundas (REDEXT) y la red de boyas costeras (REDCOS).

La Red Exterior está compuesta por boyas de tipo Wavescan y SeaWatch. Las boyas de tipo SeaWatch miden tanto oleaje como parámetros atmosféricos y oceanográficos, mientras que las boyas de tipo Wavescan solo miden oleaje y variables atmosféricas.

Estas boyas se caracterizan por estar fondeadas lejos de la línea de costa a gran profundidad (mas de 200 metros de profundidad), por lo que las medidas de oleaje de estos sensores no están perturbadas por efectos locales. Por ello, cada boya proporciona observaciones representativas de grandes zonas litorales.

La Red Costera está compuesta por boyas escalares de tipo Waverider (Datawell) y boyas direccionales de tipo Triaxys (Axis). Estas boyas se caracterizan por estar ubicadas en las proximidades de instalaciones portuarias, estando fondeadas a menos de 100 m. de profundidad. En la mayoría de los casos, las medidas están perturbadas tanto por el perfil de costa, como por los efectos del fondo sobre el oleaje, motivo por el cual los datos de estas boyas son representativos, solamente, de condiciones locales.

Adicionalmente, se puede encontrar información en el documento de Recomendaciones de Obras Marítimas ROM 0.3-91-Oleaje, Atlas de Clima del Litoral Español, que ofrece información de los regímenes medios y extremales provenientes de registros instrumentales de boyas, así como de la estructura espectral escalar básica de temporales.

- **Datos obtenidos con técnicas de reanálisis**

Las bases de datos generadas mediante técnicas de reanálisis están formadas por series temporales de parámetros de viento y oleaje procedentes de modelado numérico. Son por tanto, datos simulados y no proceden de medidas directas de la naturaleza. Entre

estas bases de datos, destacan los proyectos WANA e HIPOCAS, ambas incluidas en el Banco de Datos de Puertos del Estado.

Las series WANA proceden del sistema de predicción del estado de la mar que Puertos del Estado ha desarrollado en colaboración con el INM. Para cada instante el modelo proporciona campos de viento y presión consistentes con la evolución anterior de los parámetros modelados y con las observaciones realizadas. Para generar los campos de oleaje se utiliza el modelo numérico WAM, un modelo espectral de tercera generación que resuelve la ecuación de balance de energía sin establecer ninguna hipótesis a priori sobre la forma del espectro de oleaje. Este modelo trabaja en el Atlántico con una resolución de 0.25 grados (30 Km.), y en el Mediterráneo con una resolución de 0.125 grados (15 Km.). Las series comienzan en 1996 y desde ese año la base de datos se actualiza diariamente. La cadencia de la serie es de un registro cada 3 horas.

Por otra parte, el conjunto de datos SIMAR-44, obtenidos por Puertos del Estado en el marco del proyecto europeo HIPOCAS, proviene de modelado numérico de alta resolución de atmósfera, nivel del mar y oleaje que cubre todo el entorno litoral español.

Para generar los campos de oleaje se emplea el modelo numérico WAM, con una malla de espaciamiento variable con una resolución de 15'Lat.*15'Lon. para el borde Este de la malla y de 7.5'Lat.*7.5'Lon. (aprox. 12.5Km*12.5Km) para el resto del área modelada, en el caso del área mediterránea. Las series temporales almacenadas en el Banco de Datos cubren el periodo 1958-2001 con un dato cada 3 horas.

5.4.1.2. Selección de la base de datos

Una vez descritas las fuentes de información de oleaje existentes, es necesario un proceso de evaluación que permita seleccionar la base de datos más adecuada para la caracterización del oleaje en la zona de interés.

Los datos visuales son buenos cualitativamente, incluyen información direccional y, como ventaja fundamental se encuentra la larga duración de sus registros. Sin embargo, tienden a presentar una gran subjetividad, ya que se ven influidos por la apreciación y experiencia del observador, y los valores extremos no están bien recogidos.

En cuanto a las series instrumentales provenientes de los datos tomados por la boya de la REDCOS de Alicante, cabe destacar que son escalares hasta el año 2007, con lo que la serie potencialmente útil se reduce al periodo entre dicho año y la actualidad, lo que supone una longitud insuficiente para el estudio a realizar.

En cuanto a los datos procedentes de técnicas de reanálisis, ambos son adecuados para este estudio, ya que incluyen información sobre oleaje direccional, así como de campos de viento. Sin embargo, y por no considerarse necesaria una serie más larga para este estudio, se ha utilizado la serie WANA, ya que ésta incluye el periodo de estudio.

5.4.1.3. Resultados

La caracterización del **régimen medio** en profundidades indefinidas permitirá conocer las condiciones medias anuales del oleaje, con lo que será posible evaluar si las condiciones de clima marítimo de la zona son favorables o desfavorables para la dilución de la salmuera.

Para dicha caracterización se han obtenido en primer lugar, los descriptores estadísticos para la altura de ola significativa, H_s y el periodo de pico, T_p . Tras este primer paso, se determina la función de distribución de mejor ajuste para cada una de las variables, que en este caso ha resultado ser la función log-normal, tanto para la altura de ola significativa, como para el periodo.

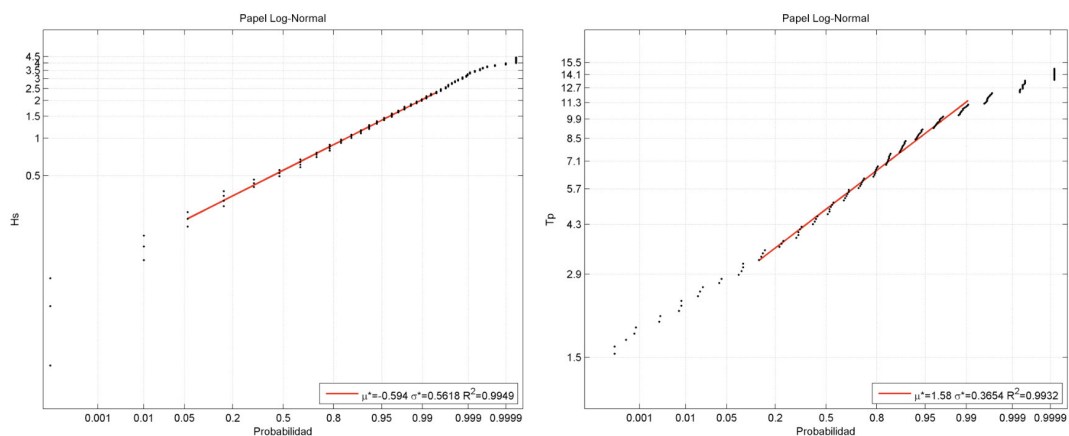


Figura 5.2. Funciones de distribución de mejor ajuste de a) H_s y b) T_p

La media en una distribución log-normal es e^{μ} , con lo que el valor medio de la altura significativa en profundidades indefinidas será de 0.55 m, es decir que el clima marítimo de régimen medio no es muy energético en esta zona. En cuanto al periodo, el valor medio en profundidades indefinidas es de 4.85 s.

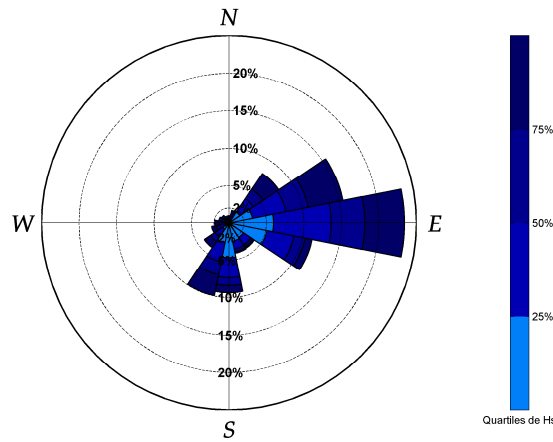


Figura 5.3. Rosa de oleaje de la altura significativa

En las representaciones de la rosa de oleaje para H_s , se observa como la dirección preferente del oleaje es E, dirección en la que se acumula algo más del 20% de la probabilidad. Sin embargo, la dirección más probable no es la más energética, ya que si se presta atención a la escala de color de las barras, se puede concluir que las mayores alturas de ola proceden del sector NE. Esta información se puede extraer también de la tabla de estadísticos básicos, que se muestra a continuación:

TABLA ESTADÍSTICOS BASICOS

Variable medida: H_s

direcciones(°)	prob.direccion	$H_{s_{50\%}}$	$H_{s_{90\%}}$	$H_{s_{99\%}}$	$H_{s_{12}}$
N	0.0097	0.6000	1.2000	1.7720	2.5911
NNE	0.0170	0.6000	1.1000	1.9860	2.6952
NE	0.0798	0.7000	1.3000	2.5000	3.6233
ENE	0.1549	0.7000	1.3000	2.5000	3.5000
E	0.2339	0.6000	1.1000	2.1000	3.4000
ESE	0.1129	0.5000	0.8000	1.4000	2.0000
SE	0.0405	0.4000	0.7000	1.3000	1.7674
SSE	0.0412	0.4000	0.7000	1.3420	2.3000
S	0.0935	0.5000	0.9000	1.6000	2.5000
SSW	0.0991	0.6000	1.3000	2.0000	2.5000
SW	0.0396	0.6000	1.2000	1.8000	2.2000
WSW	0.0236	0.6000	1.2000	2.0000	2.3237
W	0.0189	0.7000	1.3000	2.0000	2.4000
WNW	0.0139	0.7000	1.4000	2.1000	2.4000
NW	0.0121	0.7000	1.4000	2.0760	2.6253
NNW	0.0093	0.6000	1.2000	2.7000	2.8000

Tabla 5.2. Tabla de estadísticos básicos de la variable H_s

La Tabla 5.2 muestra que la dirección de oleaje predominante proviene del Este. Un gran porcentaje de tiempo, la altura de ola significativa es menor de 0.6 metros, y menos del 10% del tiempo la altura de ola es mayor de 1.2 metros. El periodo de pico para la dirección predominante se encuentra entre 5.2 y 11.2 metros .

Para la caracterización del **régimen extremal**, se ha aplicado el modelo POT (excedencias sobre un umbral), que presenta todos aquellos valores que superan un valor umbral, que corresponde en este caso al cuartil del 99%. El estudio del régimen extremal del oleaje en la zona de estudio permite caracterizar aquellos eventos extremos que se dan pocas veces al año pero que, por su magnitud y consecuencias, son de gran importancia, al ser los que favorecen en mayor medida el inicio de la dilución y mezcla.

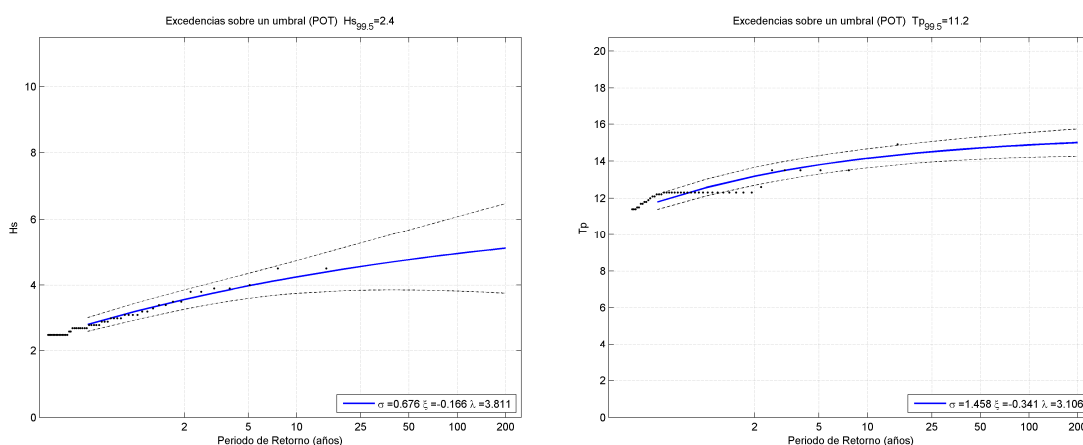


Figura 5.4. Excedencias sobre el umbral del cuartil 99% de Hs y Tp

La altura de ola que es superada un 1% de las veces es aquella que tiene una probabilidad del 99% de no ser superada, y esa altura de ola corresponde, según el ajuste del modelo POT a 2.4 m.

La caracterización de los **regímenes de viento** de la zona se realizó también a partir de los datos de la serie WANA. Los resultados para el **régimen medio** muestran un ajuste a una función de distribución Gumbel de máximos con las mayores probabilidades de ocurrencia del NE y ENE y valores medios de velocidad entre 4-5 m/s.

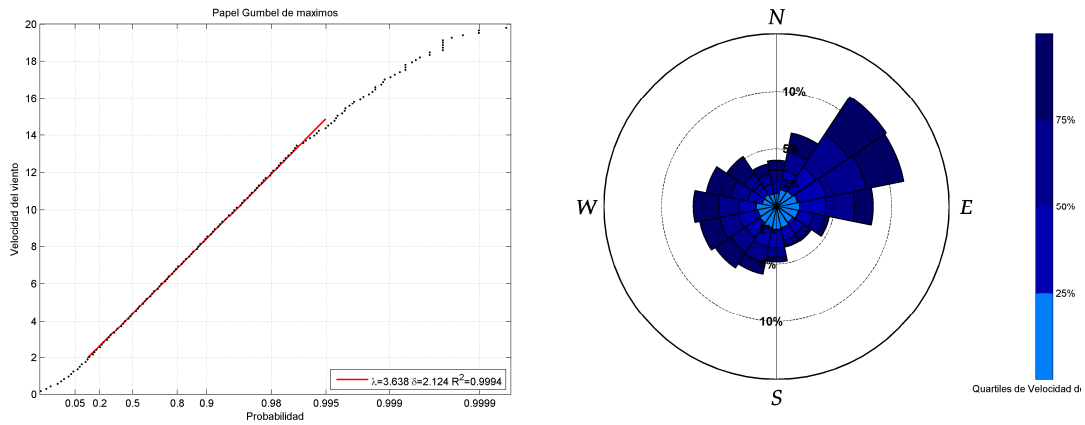


Figura 5.5. a) Función de distribución de mejor ajuste para la velocidad del viento. b) Direcciones de procedencia del viento.

En cuanto al **régimen extremal** de esta variable, la velocidad superada un 1% de las veces es 14.5 m/s.

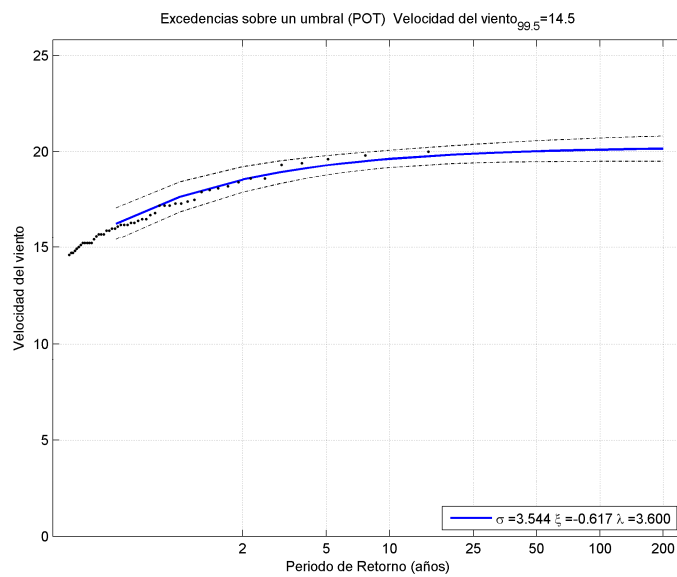


Figura 5.6. Excedencias sobre un umbral de la velocidad del viento

5.4.2. Nivel del mar

5.4.2.4. Bases de datos de nivel del mar

Existen dos redes de medida del nivel del mar en las costas españolas, REDMAR, red de mareógrafos del Ente Público Puertos del Estado (EPPE) y la red de mareógrafos del Instituto Nacional de Geografía.

- **Mareógrafos REDMAR**

El conjunto de datos REDMAR está formado por las medidas procedentes de la Red de Mareógrafos de Puertos del Estado. Esta red estaba compuesta inicialmente por mareógrafos acústicos SONAR. Posteriormente (2001), se instalaron algunos sensores de presión Aanderaa y desde 2006 toda la red se está renovando y ampliando con sensores de radar Miros.

De manera general, los mareógrafos de la REDMAR proporcionan datos brutos cada 5 minutos que son enviados a una estación receptora ubicada en la Autoridad Portuaria. Posteriormente, mediante correo electrónico dicha información llega a Puertos del Estado con una periodicidad que depende de cada una de las estaciones (6h, 1h ó 1 minuto). Cada hora, los datos que van llegando son sometidos a un control de calidad automático.

En los nuevos mareógrafos de radar los datos tienen una frecuencia de muestreo de 0.5 segundos. La información también es transmitida a una estación receptora ubicada en la Autoridad Portuaria en la que los datos de nivel del mar son promediados cada minuto y los parámetros de oleaje cada 20 minutos. Toda esta información se transfiere a Puertos del Estado por Internet, y cada hora, los datos que van llegando son sometidos a un control de calidad automático.

- **Mareógrafos del IGN y del IEO**

Las redes de mareógrafos del Instituto Geográfico Nacional (IGN) y del Instituto Español de Oceanografía (IEO) están constituidas por mareógrafos de flotador. Este

tipo de mareógrafos realizan medidas directas basándose en el movimiento ascendente y descendente del flotador.

5.4.2.5. Selección de la base de datos

En este caso, al contrario que para el oleaje, la elección de la base de datos es directa, puesto que el área de estudio no se encuentra dentro de la red de mareógrafos REDMAR, con lo que únicamente se dispone de los datos proporcionados por el mareógrafo del IGN. En este caso se empleó una serie de datos proveniente del mareógrafo Alicante 2, con datos desde 2007 hasta 2009.

5.4.2.6. Resultados

A partir de la serie completa de registros del mareógrafo del IGN y de las componentes mareales, se ha calculado la **marea astronómica**. Puesto que el nivel del mar registrado por el mareógrafo incluye tanto la marea meteorológica, como la marea astronómica, la resta de la componente astronómica calculada proporciona el **residuo meteorológico** (Figura 5.7).

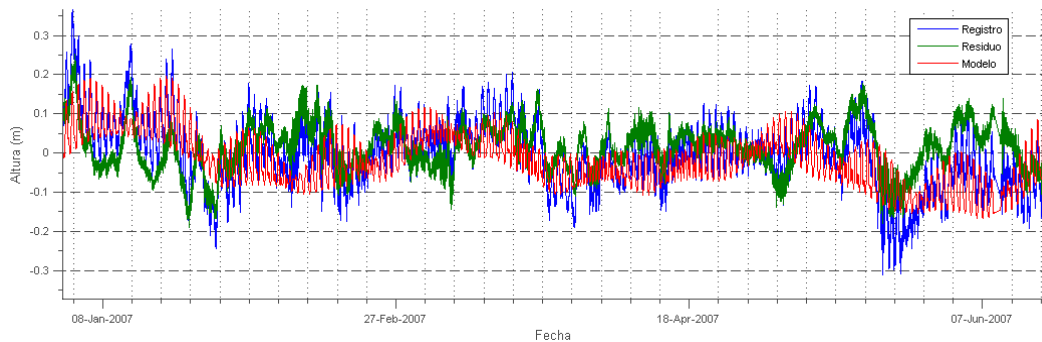


Figura 5.7. Series de nivel del mar (azul), marea astronómica (rojo) y residuo meteorológico (verde)

Tal como se puede apreciar en el análisis realizado, el rango de marea en Alicante es centimétrico, llegando a un máximo del orden de los 30 cm durante la mareas vivas. Igualmente, la influencia de las borrascas y anticiclones en el nivel del mar es del mismo orden.

5.5. Resultados computacionales

5.5.1. Datos disponibles

Los datos disponibles para el desarrollo del sistema de predicción son, las series temporales registradas por el sistema de monitorización en tiempo real del proyecto ASDECO, y los datos referentes a las condiciones de explotación de las plantas desalinizadoras del Canal de Alicante, procedentes de la documentación del proyecto ASDECO.

El sistema de monitorización, implantado por TECNOMA y SIDMAR en el área de estudio en 2008, sigue un esquema formado por tres estaciones de medida, una en tierra y dos en el mar. En la arqueta donde se integran los rechazos de Alicante I y II se monitorizan pH, salinidad y temperatura del agua de rechazo. En el campo cercano, se sitúa la estación SADO I, donde se monitorizan los perfiles verticales de velocidades, salinidades, temperatura, presión y turbidez así como el oleaje, altura de la columna de agua, viento y presión atmosférica. En límite de la pradera de *Posidonia oceanica* se sitúa la estación SADO II, donde se miden salinidad, temperatura y presión en el fondo.

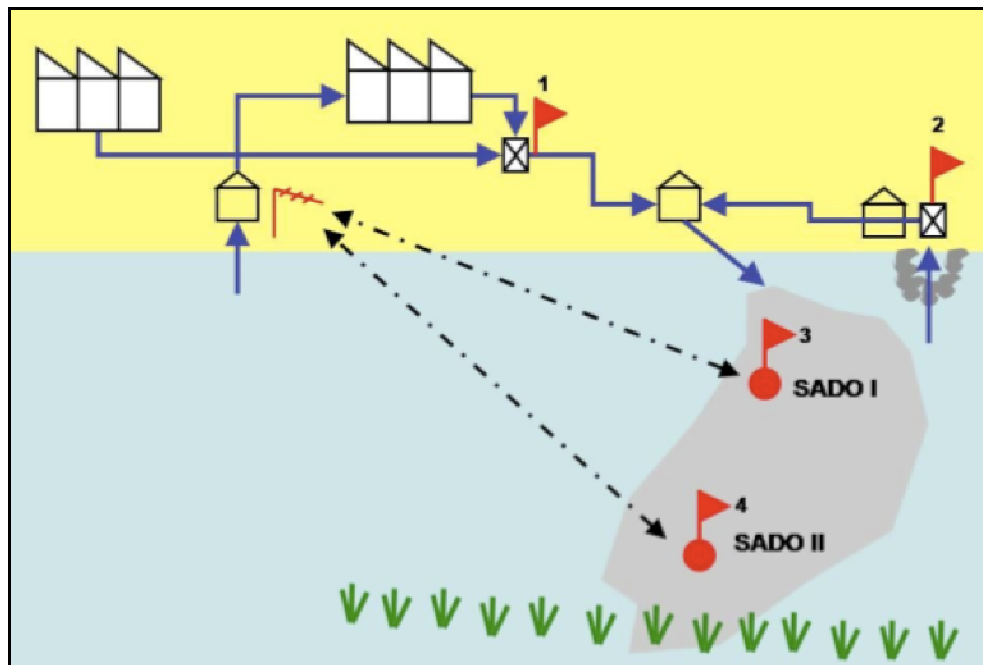


Figura 5.8. Esquema lógico del sistema ASDECO. Documentación del proyecto, 2008.

La adquisición de datos por parte de los sensores instalados en la estación SADO I se realiza cada 60 minutos, siendo la duración de la medida de 10 minutos, con un offset (tiempo transcurrido entre el inicio del intervalo y comienzo de la medida) de 0 minutos.

Todos los sensores emplean la configuración indicada salvo el sensor de oleaje TRIAXYS, que mide durante un intervalo de 20 minutos. Todos los muestreos se realizan a 1Hz de frecuencia salvo el oleaje (2Hz) y corrientes (>2Hz).

En cuanto a las estaciones situadas en campo lejano (SADO II) y en el efluente, el intervalo de tiempo entre medidas es de 10 minutos, y la duración de las mismas de 3 minutos con un offset de 0 minutos al igual que en el caso anterior. Todos los muestreos se realizan a 1Hz de frecuencia.

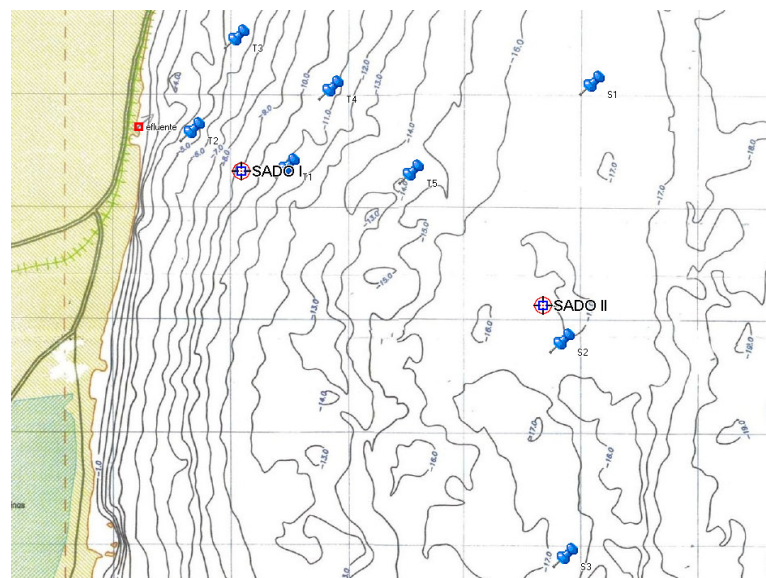


Figura 5.9. Posición de las estaciones de medida SADO I y SADO II

Puesto que los datos a considerar como posibles entradas a la red neuronal proceden de diferentes sensores, nos enfrentamos al hecho de que no siempre existe simultaneidad entre medidas. Así, y tras estudiar las series temporales disponibles y buscar la coincidencia en el tiempo de variables que a priori se suponen relevantes para el modelo (oleaje -altura de ola significativa, máxima y dirección-, salinidad y caudal del efluente, velocidad y dirección del viento..), se encuentra que el periodo más adecuado para la calibración de ANFIS, con registros coetáneos de todos los sensores y sin vacíos

importantes de datos está constituido por 1034 datos, que corresponden a medidas cada hora durante el periodo comprendido entre diciembre de 2009 y enero de 2010.

Las series temporales empleadas en la generación del modelo serán expuestas en el punto 5.5.3, tras realizar la selección de las variables de entrada al modelo.

5.5.2. Selección de variables

En el Capítulo 3 se presentaron algunas de las metodologías más empleadas para la selección de las variables de entrada de los modelos basados en redes neuronales, estableciéndose como recomendables, aquellas que combinan el conocimiento previo del sistema a modelar con alguna de las técnicas analíticas disponibles.

Siguiendo este criterio, se realizó una selección inicial de variables candidatas a partir del conocimiento actual de los factores relevantes en el comportamiento de la salmuera. De esta forma, según lo expuesto en el Capítulo 4, las variables que a priori influirán en los valores de salinidad alcanzados en el área de estudio son, los factores de explotación de la planta, de los que se seleccionan el caudal de rechazo y la salinidad del mismo, y los agentes y parámetros del medio, tales como oleaje, nivel del mar, viento, corrientes o la propia salinidad del medio.

Antes de comenzar el análisis, se decidió recoger las condiciones de explotación de la planta en un solo parámetro. En primer lugar se planteó la alternativa de emplear la carga de sales incremental [Ec.5.1], esto es, los kg/s que vierte la desalinizadora, y que incrementan la concentración que hay en el mar:

$$\text{Carga de sales incremental} = Q_{\text{rechazo}} (C_{\text{rechazo}} - C_{\text{dilución}}) \quad \text{Ec.5.1.}$$

donde:

Q_{rechazo} : caudal de salmuera de la desalinizadora en m^3/s .

C_{rechazo} : concentración de la salmuera de la desalinizadora en kg/m^3 .

$C_{\text{dilución}}$: concentración del agua empleada en la dilución de la salmuera en kg/m^3 .

Si la salinidad de la salmuera fuera 38 g/l la carga incremental sería cero y no produciría efectos en el mar. Ahora bien, como la carga incremental sólo depende del caudal de rechazo, en este parámetro no se tiene en cuenta la dilución, lo que no es correcto. Había que introducir por tanto el término de concentración de sales de salida. De esta forma, se empleó como parámetro de sales el producto de la carga incremental (que incluye la sal vertida) por la concentración de sales de salida (que incluye la dilución), que se denominó parámetro de sales:

$$\text{Parámetro de sales} = Q_{\text{rechazo}} \cdot C_{\text{rechazo}} (C_{\text{rechazo}} - C_{\text{dilución}}) \quad \text{Ec.5.2.}$$

Sin embargo, al estudiar el caso concreto, se llega a la conclusión de que si no se dispone de un CTD en el punto de la toma de dilución, es innecesario tener en cuenta el incremento de salinidad que supone la salmuera con respecto al agua de dilución, ya que si esta salinidad se supone constante, al igual que lo es la concentración del rechazo de la desalinizadora, estaremos, simplemente, multiplicando por una constante. Así, se ha decidido utilizar un parámetro que es resultado de multiplicar el caudal de rechazo (m^3/s) por la concentración del vertido (kg/m^3), siendo las unidades del parámetro resultante kg/s .

$$\text{Parámetro de sales} = Q_{\text{rechazo}} \cdot C_{\text{rechazo}} \quad \text{Ec.5.3.}$$

Tras el cálculo del parámetro de sales como indicativo de las condiciones de explotación de la planta, y partir de la primera aproximación de las variables de entrada realizada a partir del conocimiento previo del sistema, se aplicó la técnica de selección "hacia delante" o *forward selection* (Jang, 1997).

```

Selecting input 2 ...
ANFIS model 12: Sal_vertido Vel_Wind --> trn=0.2227, chk=0.3530
ANFIS model 13: Sal_vertido Direc_Wind --> trn=0.2244, chk=0.3302
ANFIS model 14: Sal_vertido Hs --> trn=0.1903, chk=0.3809
ANFIS model 15: Sal_vertido Hmax --> trn=0.1981, chk=0.3731
ANFIS model 16: Sal_vertido Wave_dir --> trn=0.2201, chk=0.4570
ANFIS model 17: Sal_vertido Qrechazo --> trn=0.2301, chk=1.0583
ANFIS model 18: Sal_vertido Parametro_sales --> trn=0.2051, chk=1.1373
ANFIS model 19: Sal_vertido Parametro_sales_2 --> trn=0.2041, chk=0.4046
ANFIS model 20: Sal_vertido nm-valor medio) --> trn=0.2102, chk=0.3467
ANFIS model 21: Sal_vertido nm-nivel de referencia --> trn=0.2251, chk=0.3119
Currently selected inputs: Hmax Sal_vertido

Selecting input 3 ...
ANFIS model 22: Hmax Sal_vertido Vel_Wind --> trn=0.1868, chk=0.5175
ANFIS model 23: Hmax Sal_vertido Direc_Wind --> trn=0.1809, chk=0.3687
ANFIS model 24: Hmax Sal_vertido Hs --> trn=0.1852, chk=0.3736
ANFIS model 25: Hmax Sal_vertido Wave_dir --> trn=0.1673, chk=0.5607
ANFIS model 26: Hmax Sal_vertido Qrechazo --> trn=0.1913, chk=0.5827
ANFIS model 27: Hmax Sal_vertido Parametro_sales --> trn=0.1781, chk=1.0107
ANFIS model 28: Hmax Sal_vertido Parametro_sales_2 --> trn=0.1752, chk=0.8668
ANFIS model 29: Hmax Sal_vertido nm-valor medio) --> trn=0.1729, chk=0.3167
ANFIS model 30: Hmax Sal_vertido nm-nivel de referencia --> trn=0.1861, chk=0.3912
Currently selected inputs: Hmax Wave_dir Sal_vertido

Selecting input 4 ...
ANFIS model 31: Hmax Wave_dir Sal_vertido Vel_Wind --> trn=0.1496, chk=0.9481
ANFIS model 32: Hmax Wave_dir Sal_vertido Direc_Wind --> trn=0.1525, chk=0.5344
ANFIS model 33: Hmax Wave_dir Sal_vertido Hs --> trn=0.1412, chk=0.7128
ANFIS model 34: Hmax Wave_dir Sal_vertido Qrechazo --> trn=0.1445, chk=8.0018
ANFIS model 35: Hmax Wave_dir Sal_vertido Parametro_sales --> trn=0.1334, chk=4.5237
ANFIS model 36: Hmax Wave_dir Sal_vertido Parametro_sales_2 --> trn=0.1362, chk=2.2105
ANFIS model 37: Hmax Wave_dir Sal_vertido nm-valor medio) --> trn=0.1412, chk=0.7639
ANFIS model 38: Hmax Wave_dir Sal_vertido nm-nivel de referencia --> trn=0.1512, chk=0.6866
Currently selected inputs: Hmax Wave_dir Sal_vertido Parametro_sales

```

Figura 5.10. Resultados del método selección "hacia delante".

De la aplicación del método *forward selection* se obtiene que el modelo que resulta en un menor RSME es aquel que incluye como variables la altura de ola máxima, el parámetro de sales, la dirección del oleaje y la salinidad del vertido. Las variables parámetro de sales y salinidad del vertido se encuentran correlacionadas entre sí, por lo que la introducción de ambas en el modelo, supondría introducir información redundante e innecesaria, con lo que las variables seleccionadas como más importantes son la altura de ola máxima, el parámetro de sales y la dirección del oleaje.

Teniendo en cuenta que el efecto del oleaje, cuando este es representativo de un estado energético alto, sobre la salinidad no es inmediato, y que estados anteriores influyen en la respuesta actual de la salinidad, se ha decidido incluir un regresor de la altura de ola máxima. Basándose en la información obtenida a partir de los resultados procedentes de análisis de correlación cruzada y por prueba error, el orden del vector regresor ha sido

establecido en 15, con lo que la altura de ola máxima en $t-15$ será otra de las variables de entrada del modelo.

Para la variable a predecir, la salinidad, se realizaron estudios de auto-correlación que permiten detectar fluctuaciones a corto plazo de la variable y tendencias, e identificar el valor o valores pasados de la misma que tengan una influencia mayor sobre el valor actual.

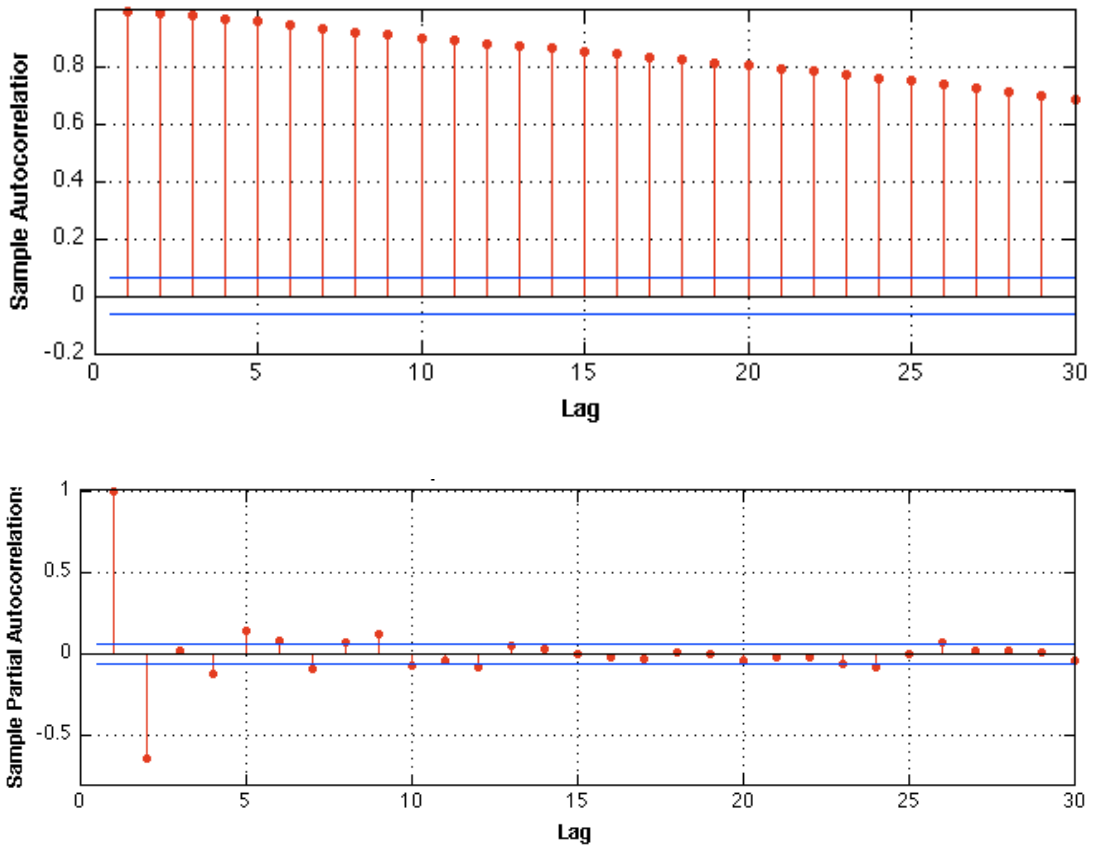


Figura 5.11. Funciones de a) Auto-correlación, ACF (el eje negativo no se muestra puesto que los resultados son simétricos a ambos lados) y b) Auto-correlación parcial de la salinidad.

Los valores de la función de auto-correlación (ACF) muestran picos en los valores anteriores (*lag*) 1 y 2, con valores subsiguientes que no decaen rápidamente, lo que se corresponde con un patrón típico de un proceso auto-regresivo. Por otra parte, en la representación de la función de auto-correlación parcial, se observan dos valores notablemente distintos de cero, que corresponden a los *lag* uno y dos, y que indica que el proceso auto-regresivo es de orden 2. Se seleccionan por tanto los valores de salinidad en $t-1$ y $t-2$ como entradas para el modelo a construir.

5.5.3. Tratamiento de la información

Una vez seleccionada la serie temporal y las variables de entrada al modelo y previamente a su introducción en el mismo, se llevó a cabo un minucioso estudio de las variables elegidas (Altura de ola máxima, Parámetro de sales, Salinidad (t-1), Salinidad (t-2)). Este proceso tiene como objetivo la determinación de las propiedades estadísticas descriptivas de cada una de las variables que serán utilizadas como entradas al modelo, la detección y corrección de *outliers*, y la generación de una serie temporal sin vacíos importantes de datos, para lo cual se empleó la interpolación lineal entre los datos conocidos a ambos extremos de cada hueco de datos.

Los valores estadísticos descriptivos de las variables de entrada, así como las funciones de distribución a las que se ajustan se muestran a continuación.

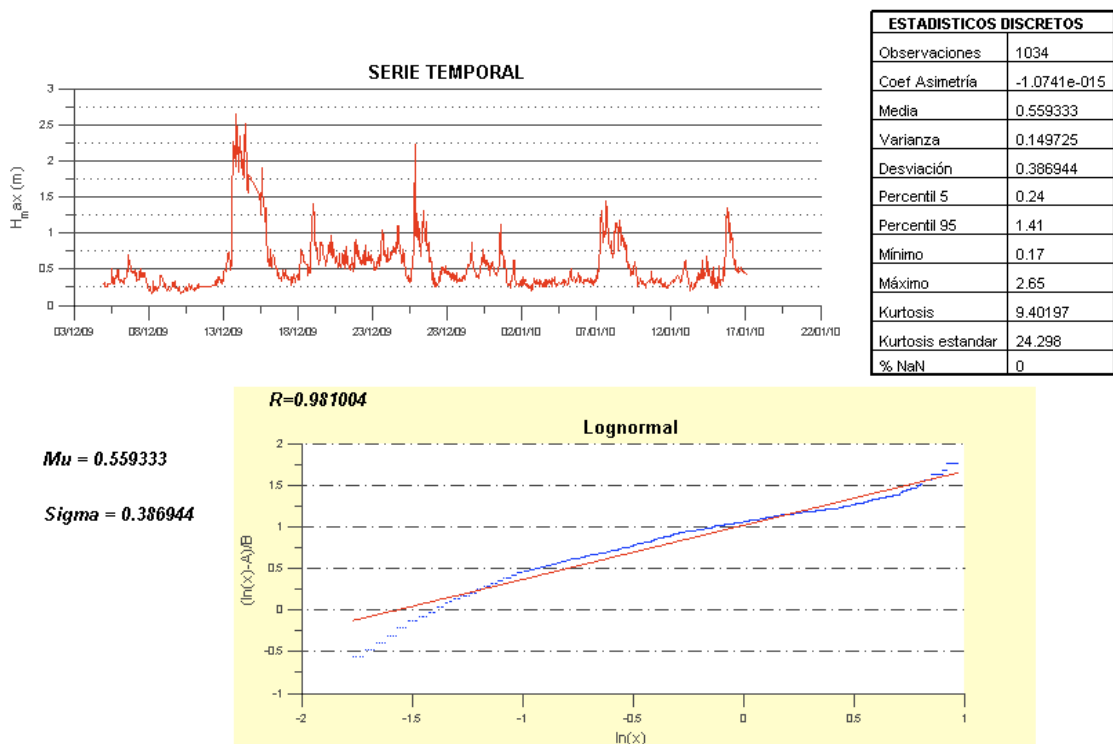


Figura 5.12. Serie temporal de altura máxima de ola empleada. En la tabla se muestran los estadísticos discretos descriptivos de la variable y en la figura inferior la función de distribución de mejor ajuste.

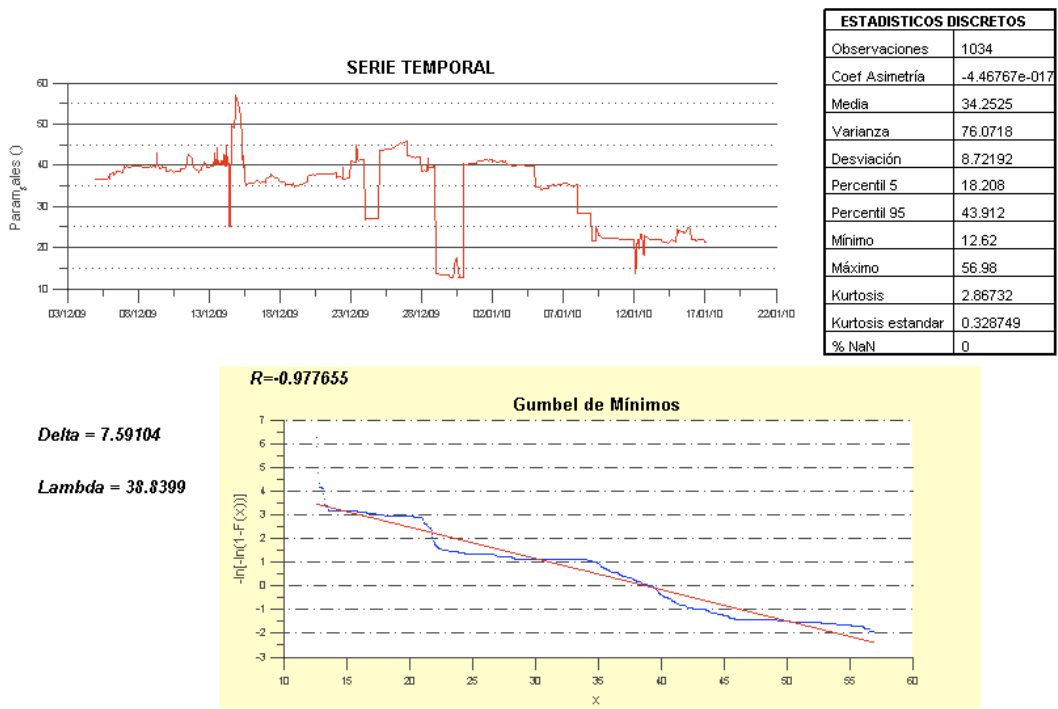


Figura 5.13. Serie temporal parámetro de sales empleada. En la tabla se muestran los estadísticos discretos descriptivos de la variable y en la figura inferior la función de distribución de mejor ajuste.

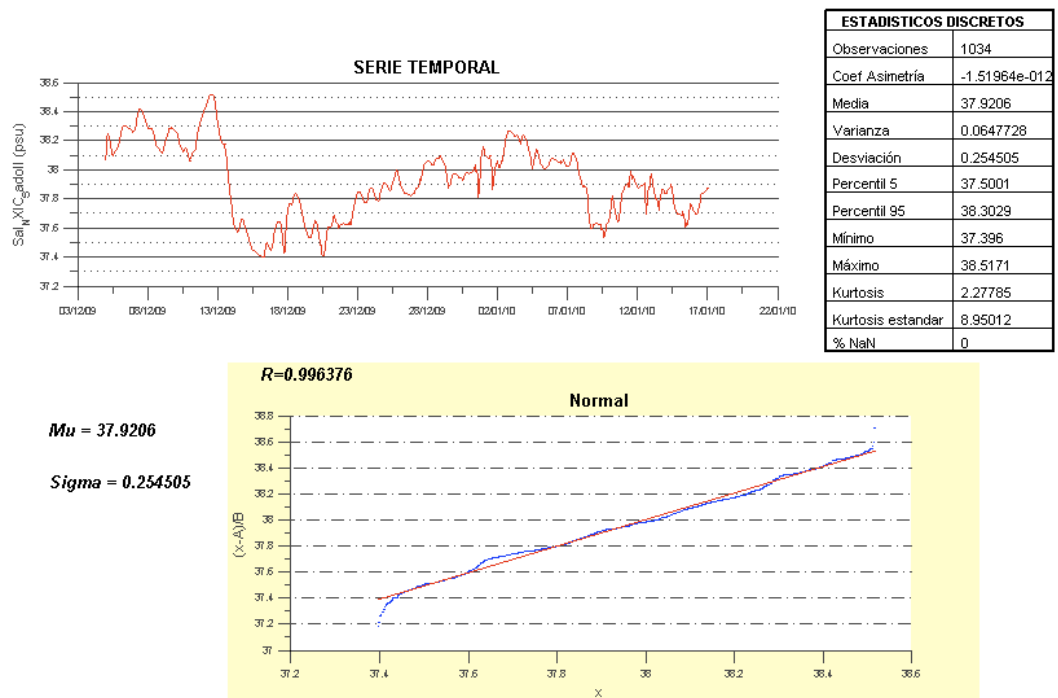


Figura 5.14. Serie temporal de salinidad empleada. En la tabla se muestran los estadísticos discretos descriptivos de la variable y en la figura inferior la función de distribución de mejor ajuste.

En cuanto a la detección de outliers, existen técnicas estadísticas tales como el test Q de Dixon que permiten su detección, sin embargo, y puesto que la extensión de la serie temporal empleada lo permitía, la inspección se realizó de forma visual, a través del examen la representación gráfica de las series de cada variable, así como de sus valores. Las series de salinidad y altura de ola no presentaron ningún valor sospechoso de ser un *outlier*, mientras que para la serie de parámetro de sales fue necesario examinar las series de caudal de rechazo y salinidad de vertido conjuntamente, para ver si las fuertes caídas del valor de parámetro de sales se correspondían con descensos en el caudal de vertido o si por el contrario se debían a fallos instrumentales en la medición de la salinidad. Tras este análisis se eliminaron algunos puntos y la serie resultante es la que se muestra en la figura 5.13. Cabe destacar que la medida del caudal de rechazo es diaria, y que por ello en la representación se pueden observar zonas de valor prácticamente constante, y escalones bruscos, que corresponderían a la diferencia de caudal de un día a otro.

El siguiente pre-tratamiento aplicado a los datos fue el suavizado de la serie temporal de la variable a predecir. Para ello, se empleó una técnica de media móvil ponderada, que multiplica los valores situados dentro de ventana temporal por una cierta cantidad, especificada en un vector de pesos, que en este caso es [0.2, 0.2, 0.2, 0.4, 0.6, 0.6, 0.8, 1]. Como se puede observar, este método asigna una mayor importancia a los valores más recientes.

Por último, y pese a que según lo visto en el Capítulo 3, la normalización de las variables de entrada aún no ha sido demostrada como paso imprescindible en el proceso de generación de modelos basados en redes neuronales, se procedió a realizar dicha transformación de las entradas con el objeto de decidir mediante prueba-error qué grupo de datos proporcionaba mejores resultados, si aquel con variables normalizadas o el que no presenta transformación alguna.

De las posibles transformaciones existentes, algunas de las cuales fueron presentadas en el Capítulo 3, se ha optado por normalizar los valores de todas las entradas entre +0.1 y +0.9. Esta normalización ha sido seleccionada debido a que las funciones de pertenencia de los modelos de inferencia difusa están escaladas entre los valores 0 y 1. El hecho de no escalar las variables a los límites de las funciones de pertenencia se debe a la

recomendación de Maier et al., 2000, que, aunque con respecto a las funciones de transferencia de las redes neuronales, afirman que este tipo de escalado podría generar tamaños extremadamente pequeños de actualización de los pesos, lo que podría desembocar en puntos de bajo rendimiento. Así, se ha extrapolado esta idea a las funciones de pertenencia de los sistemas de inferencia difusa y se ha aplicado la transformación indicada, que viene dada por la ecuación 4.3.

5.5.4. Calibración

Tal como se apuntó en el Capítulo 3, los objetivos de la calibración de un modelo neuro-difuso como ANFIS son la identificación de la estructura y la estimación de parámetros del modelo.

Para la consecución de parte de la primera de las tareas se hizo uso de la interfaz gráfica Fuzzy Logic Toolbox de MATLAB. En primer lugar se prepararon los datos, dividiendo la serie disponible en dos grupos diferentes, uno para el entrenamiento y otro para evitar el sobreajuste de la red. Además se reservó otro grupo de datos para la posterior comprobación de la capacidad de generalización de la red ya entrenada. El primer grupo, o grupo de entrenamiento contiene el 70% de los datos, y el grupo de test el 30% restante. Los datos fueron ordenados aleatoriamente antes de su división.

Se consideraron distintos tipos de modelos hasta un total de cuatro. Estos modelos fueron:

- Modelo con partición del espacio en rejilla con funciones gaussianas. Para este modelo se consideró el mismo número de funciones de pertenencia para todas las entradas, con lo que el número de reglas fue igual a este número elevado al número de entradas. Puesto que el número de funciones de pertenencia se fijó en dos y las entradas en las cinco ya comentadas, el número de reglas fue $2^5 = 32$.
- Modelo con partición del espacio de entrada por agrupamiento substractivo. En este caso las funciones de pertenencia también son gaussianas, y el número de reglas coincide con el número de clusters generados.

- Los dos modelos restantes son iguales que los anteriores pero con las variables de entrada normalizadas.

La tabla 5.3 recoge los valores estadísticos de los datos de cada uno de los tres grupos.

	Entrenamiento			Validación			Test		
	Hmax	Parámetro sales	Salinidad	Hmax	Parámetro sales	Salinidad	Hmax	Parámetro sales	Salinidad
Media	0.56	34.15	37.92	0.56	34.45	37.92	0.58	39.12	38.62
Mínimo	0.17	12.62	37.40	0.17	12.65	37.41	0.2	35.01	36.55
Máximo	2.65	56.98	38.52	2.47	56.22	38.51	2.42	55.83	38.56
Desviación estándar	0.39	8.72	0.26	0.39	8.74	0.25	0.34	3.02	0.42

Tabla 5.3. Valores estadísticos de los tres grupos de datos

Una vez divididos los datos se procedió a resolver la incógnita de cuál es el tipo y el número de funciones de pertenencia (en el caso de utilizar la partición en rejilla), o el número de clusters si el método para la partición del espacio de entrada empleado es el agrupamiento substractivo, necesarios para definir adecuadamente cada una de las variables de entrada.

Tal como se vio en el Capítulo 4, el procedimiento comienza con la selección de uno de los dos métodos disponibles para la partición del espacio de entrada. En caso de haber seleccionado el método de partición en rejilla, antes de generar el FIS será necesario elegir el tipo de función de pertenencia para las variables de entrada. La elección de la forma de dichas funciones, así como el número de las mismas se justifica según el error cuadrático medio obtenido y el número de parámetros no lineales resultantes, de forma que se probaron todos los tipos posibles y se elaboró una tabla comparativa. (Tabla 5.4).

También se hicieron pruebas empleando el método de agrupación substractiva (*substrative clustering*) para la partición del espacio de entrada. En este caso el parámetro a ajustar es el radio de vecindad, que, tal como se vio en el Capítulo 4, definirá el número de clusters generados y en consecuencia, el número de reglas. Un buen ajuste de este parámetro es muy importante, pues un valor demasiado pequeño generará gran número de clusters, incrementando la posibilidad de que se produzca sobreentrenamiento, mientras que un valor demasiado grande generará muy pocos clusters, resultando un modelo muy poco ajustado que probablemente no sea capaz de modelar el proceso de forma apropiada.

El siguiente paso será el ajuste de los parámetros, tanto los del antecedente (constantes que caracterizan las funciones de pertenencia) como los del consecuente (parámetros lineales de la salida del modelo). Para ello se selecciona el método de aprendizaje híbrido (Jang, 1995) descrito en el Capítulo 3.

Una vez identificados los parámetros del consecuente, la suma de errores cuadráticos procede de la diferencia entre la salida de la red y la salida deseada presentada en los pares de entrenamiento. Cuando se alcanza un error estable y razonable se valida la predicción frente a los datos destinados a este propósito para asegurar que no se está produciendo sobreentrenamiento. El modelo con menor RMSE, comprobado que no esté sobreentrenado y cuyo número de parámetros sea adecuado (éstos no deberían exceder a sexta parte del número integral de datos de entrenamiento) será el seleccionado.

Función de pertenencia	RMSE
Triangular	0.014
Trapezoidal	0.051
Gumbel	0.016
Gaussiana	0.012
Gaussiana compuesta	0.028
pi	0.067
DSigmoidal	0.48
PSigmoidal	0.48

Tabla 5.4. RMSE obtenido de la aplicación de los diferentes tipos de funciones de pertenencia

Radio de vecindad	Squash factor	Radio de aceptación	Radio de no aceptación	RMSE
0.5	1.25	0.25	0.15	0.012
0.4	1.25	0.25	0.15	0.012
0.3	1.25	0.25	0.15	0.012

Tabla 5.5. RMSE obtenido con el método de agrupamiento substractivo para diferentes radios de influencia

Tal como ya se había anunciado se seleccionaron, para próximas pruebas, los modelos basados en funciones de pertenencia gaussianas (RMSE 0.012) y el modelo basado en la técnica de agrupamiento substractivo, obtenido con un radio de vecindad de 0.4.

Tras este proceso, ya tenemos una primera aproximación del modelo definitivo, o más bien del tipo y número de funciones de pertenencia, y del radio de vecindad. Sin embargo, para una mejor estimación de los parámetros y para comprobar el rendimiento del modelo, se emplearán las funciones creadas por Marcé, que integran todo el proceso de estimación de la estructura y de los parámetros de la red neuronal e incluyen una serie de cálculos estadísticos que serán los que permitan determinar la validez del modelo. Estos algoritmos han sido completados además con funciones de representación de datos que permitirán la emisión de juicios basados en la inspección visual.

De los cuatro modelos entrenados, todos mostraron un buen rendimiento con ajustes 0.99 entre los datos medidos y los simulados por ANFIS y errores sobre 0.01 y 0.03 en términos de RMSE. En el caso de los modelos que emplearon el agrupamiento substractivo como método de partición del espacio de entrada, aquel que presentaba las variables sin normalizar presentó una correlación ligeramente superior al normalizado. Los otros dos modelos, cuya partición del espacio de entrada fue realizada a partir de la partición en rejilla, mostraron resultados idénticos. A la vista de la escasa mejora del rendimiento del modelo con la normalización de las variables de entrada se decidió seleccionar los modelos con las variables sin transformar.

El resto del proceso se realizó para los modelos obtenidos con el agrupamiento substractivo y la partición en rejilla, sin embargo, a continuación se mostrarán los resultados obtenidos con el último, ya que fue el que presentó un mejor rendimiento

durante los pronósticos "un paso hacia adelante", que serán presentados en el apartado 5.5.6.

El modelo resultante, fue generado asignando dos funciones de pertenencia (MFs) a cada una de las entradas, por lo que el número de reglas difusas es $2^5 = 32$. Las funciones de pertenencia empleadas corresponden a funciones Gaussianas, que presentan dos parámetros, un centro p y un ancho q (Figura 5.15). El modelo difuso en la capa 4 es de tipo Sugeno y orden cero, con lo que las salidas de cada regla corresponden a un valor numérico constante.

Con los ajustes descritos, la estructura ANFIS generada contiene 20 parámetros no lineales (parámetros de las premisas), que corresponden a el número de inputs (5) x número de MFs (2) x número de parámetros de cada MF (2); y 32 lineales, que corresponden al número de reglas.

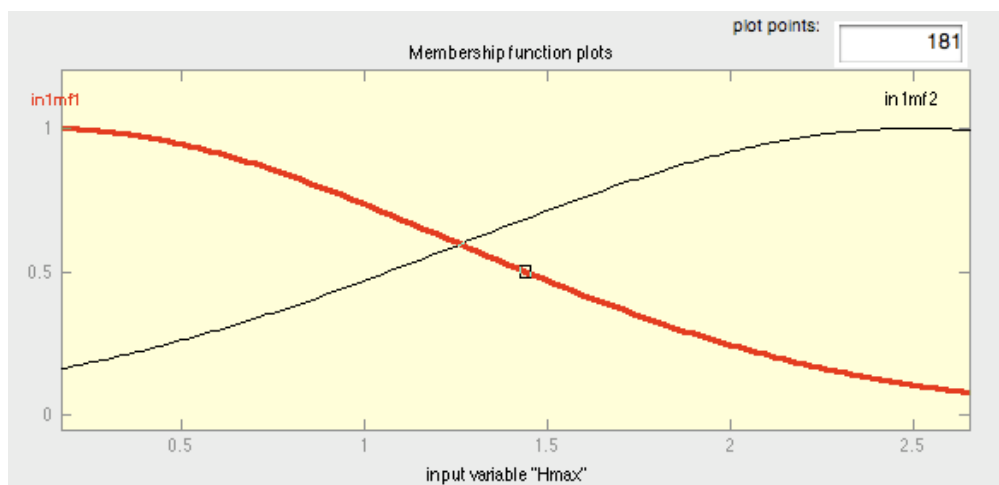


Figura 5.15. Funciones de pertenencia de la variable Hmax. La campana de gauss marcada en rojo corresponde a la primera función de pertenencia, con centro $p_1=0.1217$.

La estructura del modelo ANFIS descrito se muestra en la figura 5.16, donde las entradas alimentan al modelo desde la izquierda y son propagadas a través de cinco capas hasta que la salida es generada en el extremo derecho.

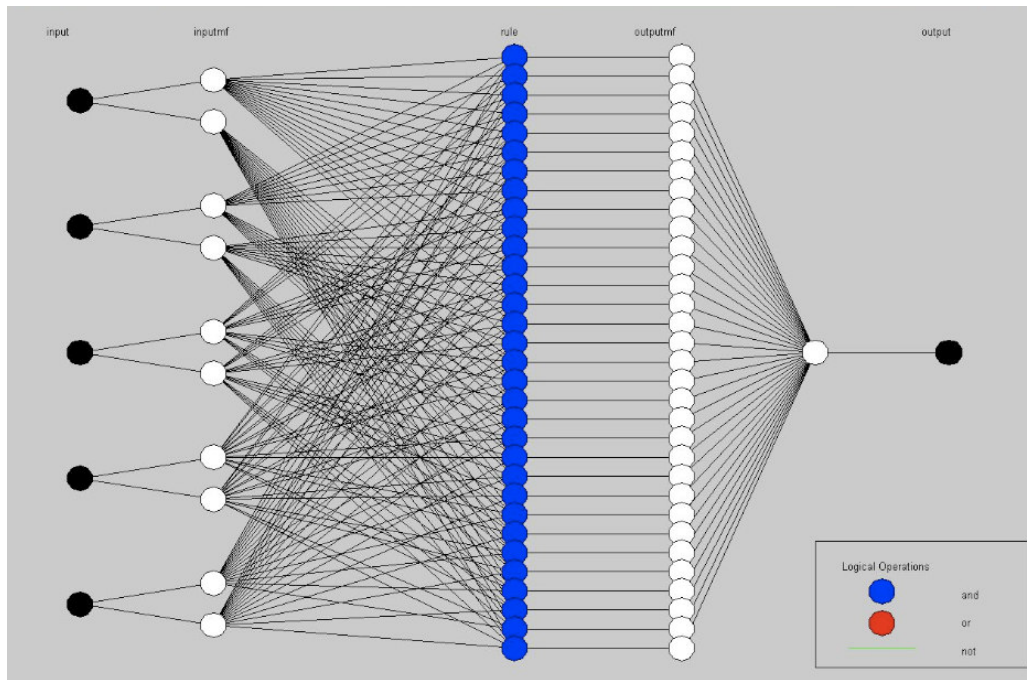


Figura 5.16. Estructura del modelo ANFIS generado.

Los círculos azules de la figura anterior corresponden a las reglas generadas por el modelo a partir de la combinación de las funciones de pertenencia de cada entrada y el ajuste de sus parámetros. Estas reglas son del tipo:

SI H_{\max} es A_i , $H_{\max}(t-15)$ es B_i , Parámetro sales es C_i , $S(t-1)$ es D_i , y $S(t-2)$ es E_i ,
 ENTONCES $S(t) = O_i$

donde A_i , B_i , C_i , D_i y E_i , $i = 1, 2$, son las funciones de pertenencia.

Tal como se expuso en el Capítulo 3, los parámetros lineales se calculan en el paso hacia delante, mientras que, una vez se obtiene el valor predicho por el modelo, se calcula la desviación del valor real y es entonces cuando este se propaga de nuevo en el paso hacia atrás, en el cual se produce la actualización de los parámetros del consecuente (parámetros de las funciones de pertenencia) empleando el método de descenso de gradiente, cuyo objetivo es disminuir la función error, definida por la ecuación:

$$E = \frac{1}{2} [y(t) - \hat{y}(t)]^2 \quad \text{Ec.5.4.}$$

donde $y(t)$ es el valor real, e $\hat{y}(t)$ es el valor estimado por ANFIS.

Una vez el modelo ha ajustado sus parámetros, y tal que se ha expuesto en líneas anteriores, disponemos de dos tipos de pruebas para verificar el rendimiento del mismo:

- La emisión de juicios informados basados en la inspección visual de las gráficas generadas.
- El análisis estadístico de los resultados y de los residuos del modelo, que permitirá corroborar los juicios y detectar inconsistencias que no son fácilmente visibles.

A continuación se muestran los resultados obtenidos en cada uno de los dos tipos de pruebas descritas con el modelo seleccionado.

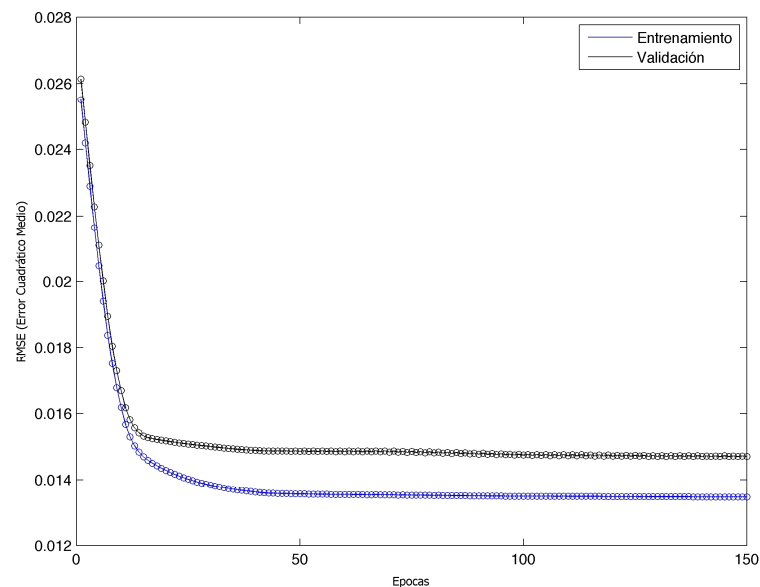


Figura 5.17. Curvas de error RMSE.

La figura anterior corresponde al error alcanzado por ANFIS en cada una de las épocas de entrenamiento, y es indicativa de que no se ha producido el fenómeno de sobreentrenamiento, ya que cuando esto sucede el error del grupo de datos de test comienza a ascender. Además se puede observar que el RMSE no disminuye al

aumentar las iteraciones, con lo que se puede afirmar que la función ha alcanzado su mínimo global.

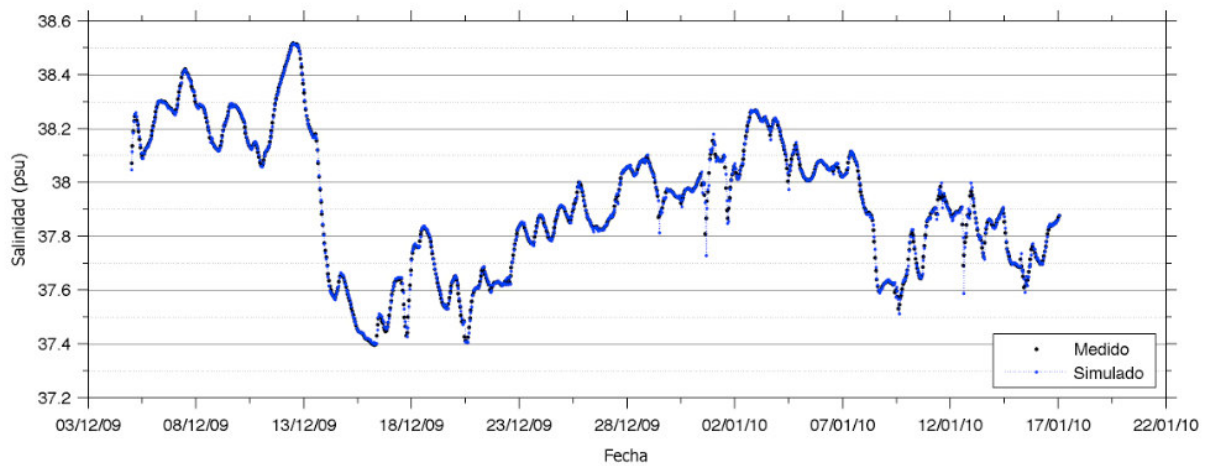


Figura 5.18. Series temporales de salinidad medida y simulada por ANFIS

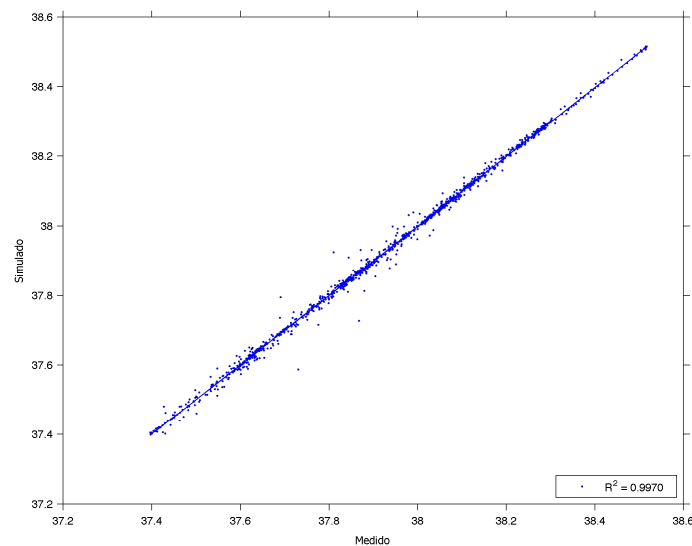


Figura 5.19. Recta de regresión entre valores medidos y simulados por ANFIS.

La figura 5.18 muestra la serie temporal medida junto con la simulada a partir de grupo de datos de entrenamiento. En ella es posible detectar la pequeña sobreestimación que ANFIS hace de algunos de los mínimos de salinidad de la serie en ambos modelos, sin embargo los valores más altos se ajustan a la perfección, siendo las dos series prácticamente indistinguibles.

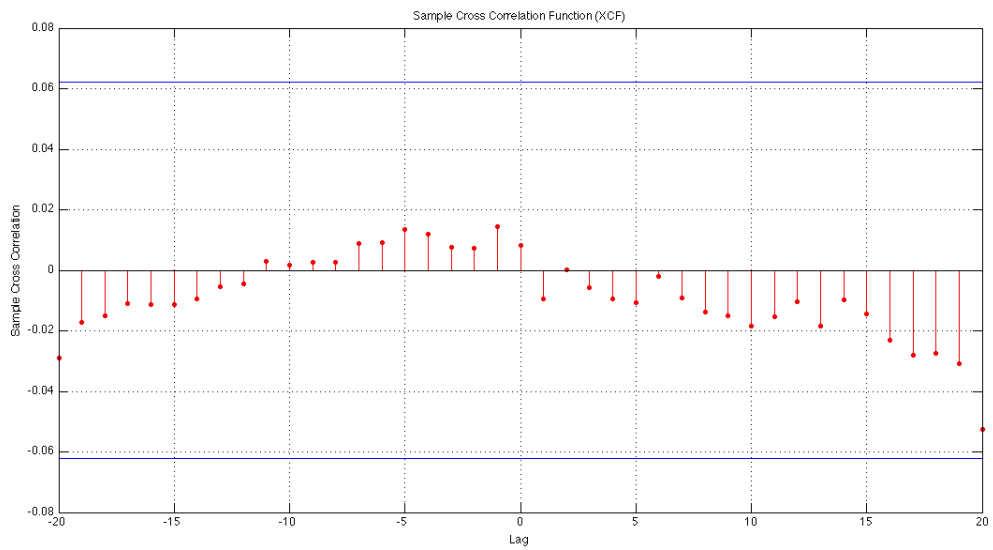
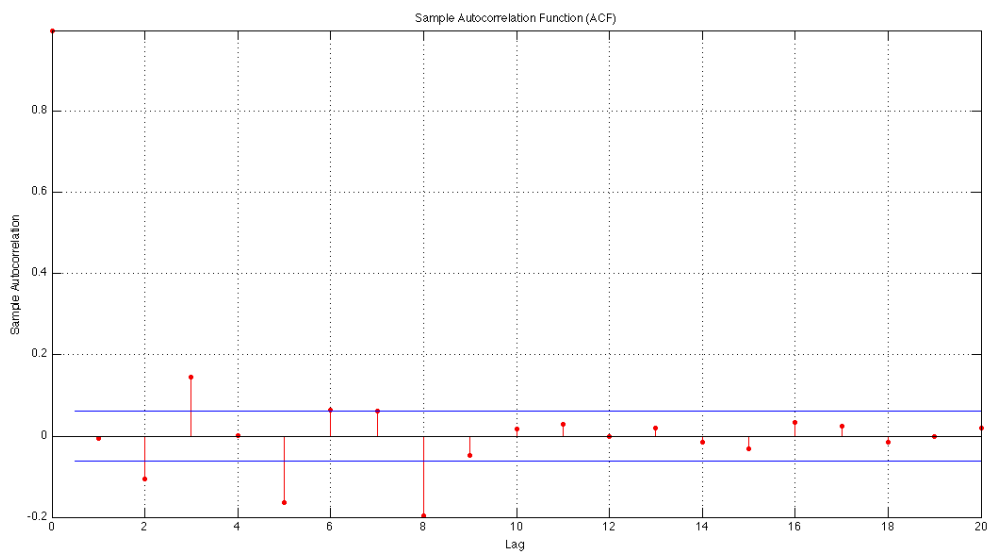
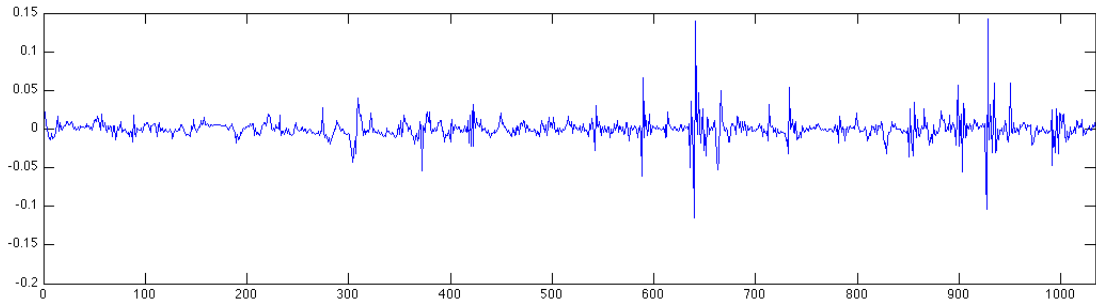
La figura 5.19 corresponde a la recta de regresión de los datos medidos frente a los simulados por ANFIS y, al igual que la representación de las series temporales, muestra como los mayores valores de salinidad son los mejor ajustados por el modelo. Además, como leyenda se incluye el valor del coeficiente de correlación entre datos observados y simulados (0.9970), que indica que la práctica totalidad de la varianza de la serie temporal puede ser explicada por los modelos. Dicho coeficiente fue calculado con la siguiente expresión,

$$R^2 = 1 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad \text{Ec.5.5.}$$

donde N es el número de datos, O_i el i-ésimo dato observado, \bar{O} es la media de los datos observados y P_i es el i-ésimo dato predicho por ANFIS. Este coeficiente proporciona una medida de la correlación entre las salidas del modelo y los datos observados o medidos, pero también es sensible a las diferencias entre las medias y varianzas de ambos grupos de datos. Además, no es sensible al tamaño de la serie de datos. El valor calculado puede variar entre $-\infty$, en el peor de los casos y 1 si la correlación es perfecta.

Finalmente, y para terminar de evaluar el rendimiento alcanzado por el modelo se procedió a analizar los residuos del mismo.

Si los residuos no contienen información acerca de residuos pasados o de la dinámica del sistema, es probable que toda la información haya sido extraída del conjunto de datos de entrenamiento y que el modelo aproxime correctamente los resultados al sistema. Para investigar esto, se estudian las funciones residuales de correlación y auto-correlación, para comprobar si estas no están correlacionadas y, por tanto convergen a una distribución Gaussiana con media cero y varianza $1/N$. Para realizar estos análisis se representaron las funciones de correlación cruzada entre las entradas del modelo y los residuos del mismo, así como la función de auto-correlación de los errores de predicción (residuos). Los resultados se muestran en las figuras expuestas a continuación (Figura 5.20).



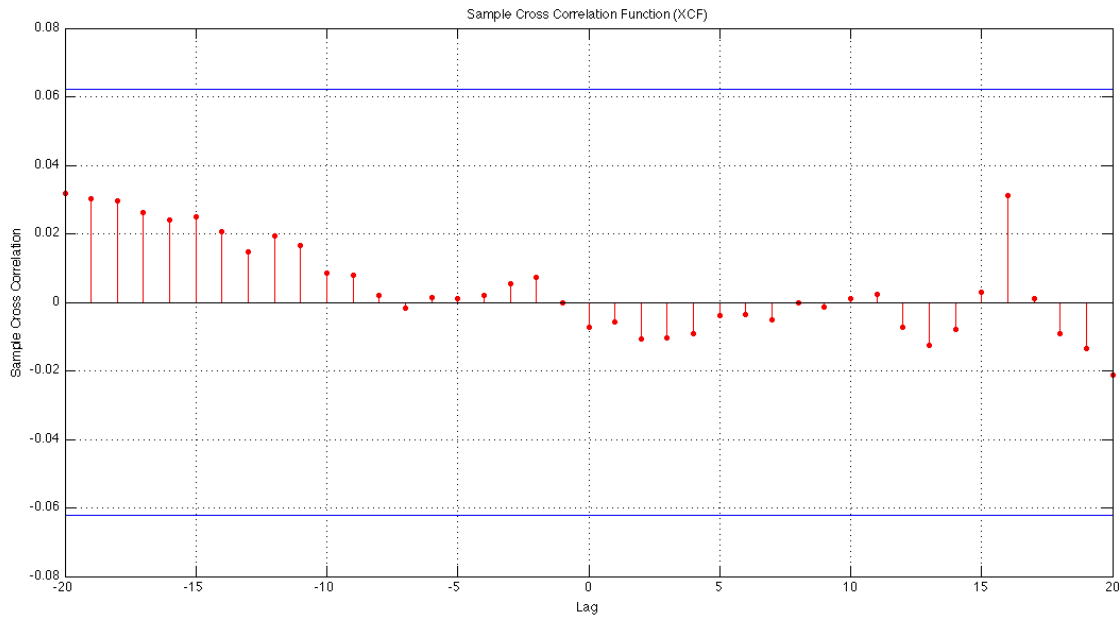


Figura 5.20. a) Residuos del modelo. b) Función de auto-correlación de los residuos. c) Función de correlación cruzada entre la altura de ola máxima y los residuos del modelo. d) Función de correlación cruzada entre el parámetro de sales y los residuos del modelo

La figura 5.20 (a-c) muestra los errores del modelo (residuos) y las funciones de auto-correlación y correlación cruzada de los residuos y de las entradas del modelo. Estas figuras muestran que, para la mayoría de *lag*, las funciones de auto-correlación y correlación cruzada caen dentro del intervalo de confianza del 95%.

5.5.5. Validación

Una vez completada la fase de entrenamiento, el rendimiento de la red ya entrenada debe ser validado empleando un conjunto de datos independiente. Para este propósito se emplea una serie de datos correspondiente al periodo 21/02/2010 - 03/03/2010, que constituye una serie de diez días. De esta forma se cumple el criterio seguido para la división de los datos, ya que el periodo seleccionado correspondería a un 20% del total de datos.

Al presentar el citado periodo a la red ya entrenada, se obtuvieron los resultados mostrados en la Figura 5.21.

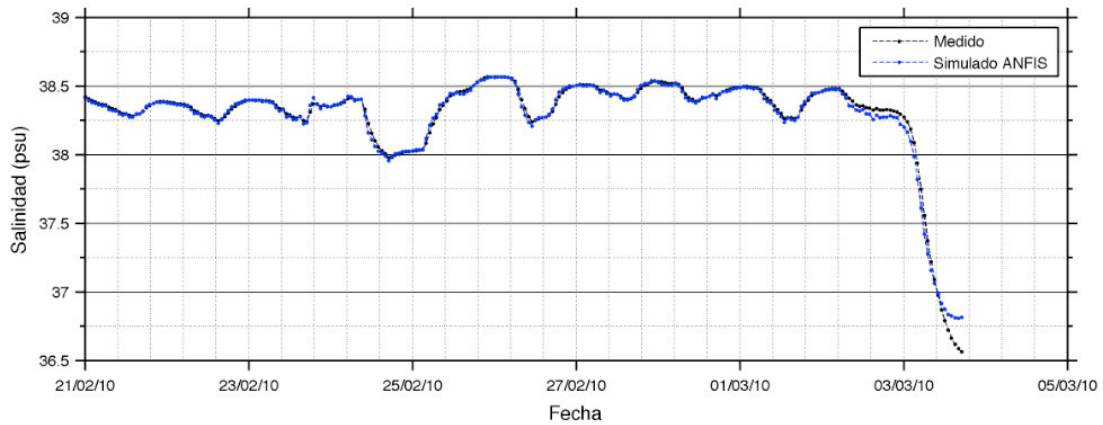


Figura 5.21. Resultados de la validación de la red entrenada con el conjunto de datos de test.

Los resultados de la prueba de testeo de la red ya entrenada y calibrada fueron satisfactorios, tal como se muestra en la figura anterior. Se observa un ajuste perfecto entre datos medidos y simulados, cuyos valores únicamente se separan en la predicción de las últimas horas. Esto se debe a que los valores de salinidad a lo largo de este último intervalo de tiempo se encuentran fuera del rango de valores de salinidad empleado durante el entrenamiento de la red. Como ya se ha mencionado en puntos anteriores del texto, los sistemas basados en redes neuronales presentan la limitación de no poder extrapolar adecuadamente valores más allá que los empleados durante la calibración. Sin embargo, el error cometido no es muy importante, estando en torno a los 0.2 psu.

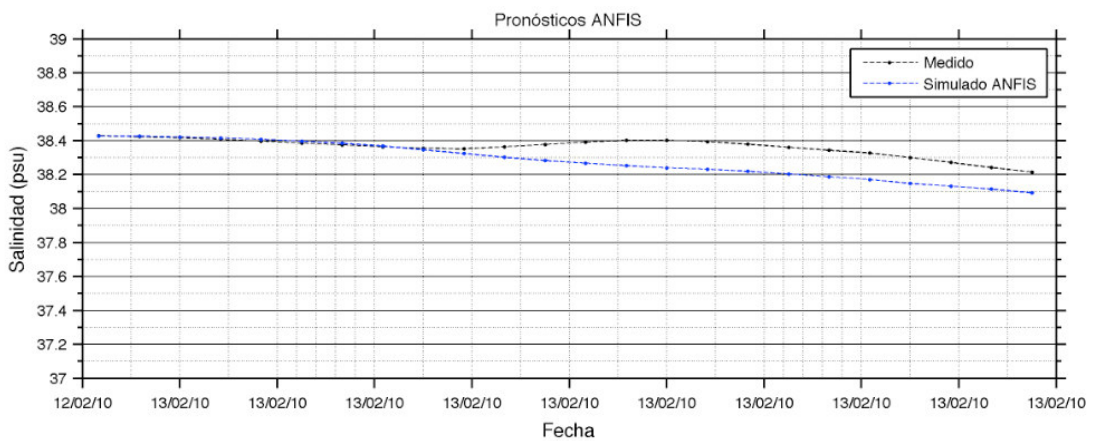
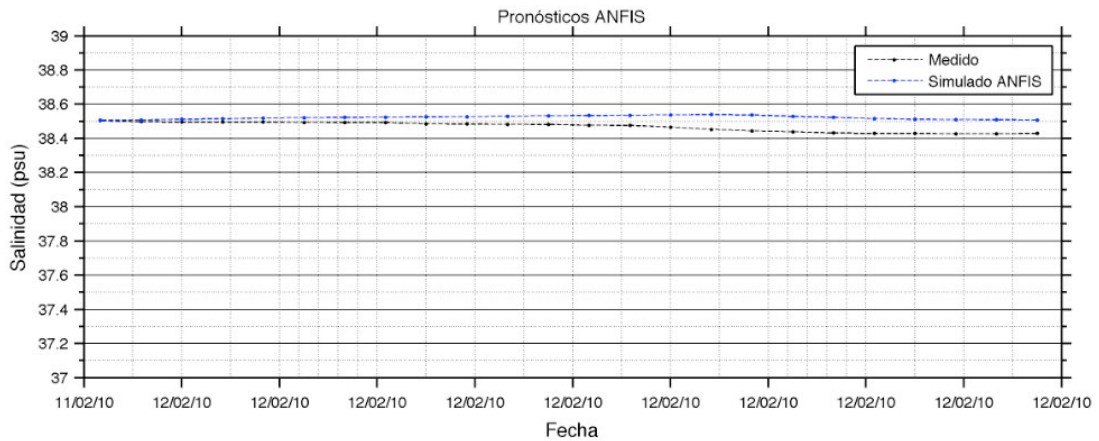
5.5.6. Generación de pronósticos

Como último paso se empleará el modelo calibrado a lo largo del proceso anterior, para generar pronósticos un paso hacia delante (*one-step-ahead*) de la salinidad. Para ello, el valor de salinidad predicho para un tiempo t servirá de entrada a la red para el pronóstico del valor en $t+1$. Para las variables exógenas del modelo y , debido a la carencia de predicciones de oleaje y de registros de SCADA con las estimaciones de los factores de explotación de la planta necesarios para el cálculo del parámetro de sales, se emplean datos medidos por los sensores, de periodos diferentes a los empleados para la calibración y validación de la red.

El objetivo es el pronóstico de la salinidad alcanzada en el área de la estación SadoII, con un horizonte de predicción de 24 horas, tiempo que ya sería suficiente para que los

responsables de la de la planta adaptasen el protocolo de gestión del vertido en caso necesario.

En la Figura 5.22 se muestran los resultados de los pronósticos de salinidad con horizonte de predicción de 24 h, obtenidos para diferentes periodos. Cabe destacar que, debido a la limitación que este tipo de sistemas presentan a la hora de extrapolar más allá de los valores empleados durante el entrenamiento, los periodos seleccionados para llevar a cabo estas pruebas fueron elegidos siguiendo este criterio, de forma que ninguno presenta valores fuera del rango empleado durante la calibración del modelo.



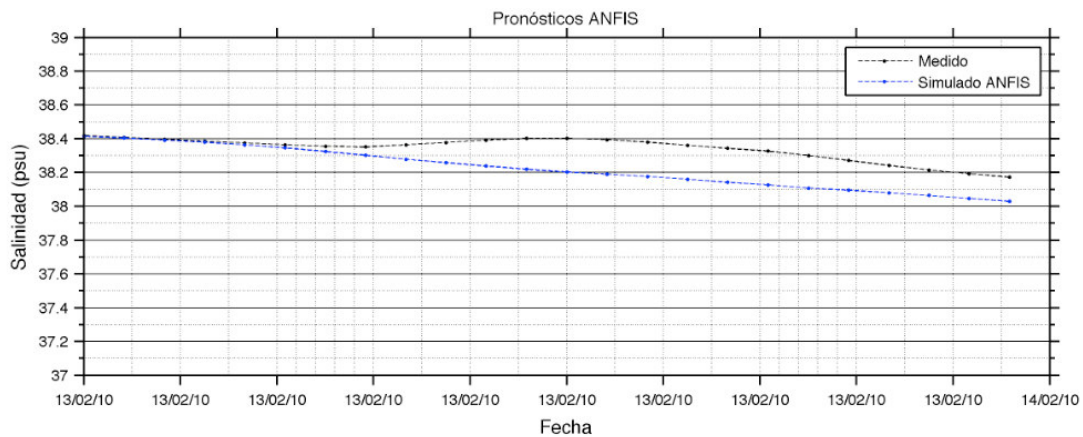


Figura 5.22. Pronósticos de salinidad para horizontes de predicción de 24 h en tres días consecutivos.

La figura anterior muestra los resultados obtenidos al aplicar el método "un paso hacia delante" para la predicción de la salinidad. El primer punto de cada pronóstico se calcula a partir de datos instrumentales de las variables de entrada. Una vez obtenido este primer valor, se calcula el segundo punto, en el cual aún se involucra un dato de salinidad procedente de medidas instrumentales, el valor de salinidad en $t-2$. A partir de la tercera hora, todos los pronósticos siguientes se realizan empleando como entradas los valores predichos en pasos anteriores.

Los resultados obtenidos muestran que el modelo es capaz de seguir la tendencia de la salinidad a lo largo de las 24 horas, si bien no se ajusta completamente a pequeñas variaciones horarias dentro del rango de predicción, algo que se aprecia también en la Figura 5.23, correspondiente a un periodo diferente a los anteriores, y en la que los valores simulados por ANFIS se ajustan a la curvatura de los datos medidos, obviando sin embargo pequeños picos a lo largo del transcurso del día.

Pese a la cierta linealidad de las predicciones, los errores cometidos por el modelo no superan los 0.2 psu, lo que se considera que un buen rendimiento por parte del sistema. Por otra parte, se aprecia que el modelo tiende a subestimar los valores de salinidad, ya que todas las predicciones mostradas, excepto la del primer día, se encuentran por debajo de los valores reales.

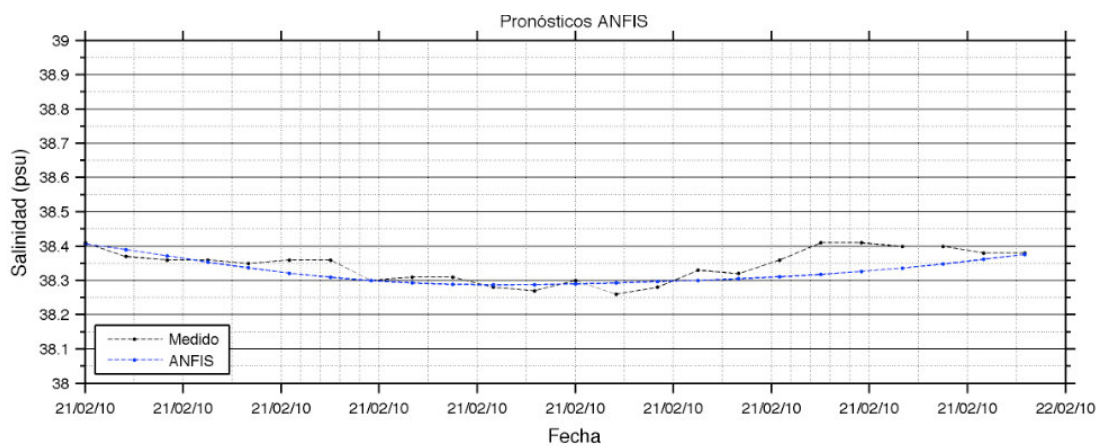


Figura 5.23. Pronósticos de salinidad para horizontes de predicción de 24 h (21/02/10, 0:00h -21/02/10, 23:00)

6 CONCLUSIONES

En este capítulo se presentan las conclusiones de la tesina, en las que se comentará la validez del modelo desarrollado para su posible implementación en un sistema de predicción en tiempo real de la salinidad. El capítulo se completa señalando futuras líneas de investigación que pueden ser llevadas a cabo a partir del modelo propuesto.

6.1. Conclusiones

A lo largo del desarrollo de esta tesina, se ha presentado la tecnología de la desalinización como solución a la escasez de los recursos de agua, planteándose asimismo su problemática ambiental, generada principalmente por el vertido al mar de grandes volúmenes de salmuera, que afectan negativamente a especies estenohalinas presentes en el medio receptor. Una vez planteada la problemática, se han repasado las técnicas actuales de modelación de la salmuera y las técnicas de inteligencia artificial en las que se basa el modelo desarrollado.

Expuesta la problemática, y establecidas las bases teóricas, se procedió a la generación del modelo neuro-difuso basado en la arquitectura ANFIS, principal objetivo de esta tesina. El modelo resultante se construyó empleando cinco variables de entrada, tres de las cuales son características de las condiciones del medio receptor y de explotación de la planta, mientras que las dos restantes están constituidas por vectores regresores de la salinidad. Se emplearon dos funciones de pertenencia gaussianas para describir el espacio de entrada de cada una de las variables, lo que resultó en un sistema neuro-difuso formado por 32 reglas y 52 parámetros a ajustar por ANFIS a partir del método de aprendizaje híbrido.

Los altos valores de correlación obtenidos en las pruebas realizadas, demuestran que el modelo es capaz de reproducir el comportamiento de la salinidad de forma razonable. Por otra parte, los estudios de correlación y auto-correlación de las variables de entrada y los residuos del modelo, indican que éste aproxima correctamente los resultados al sistema de estudio.

Tras el entrenamiento de la red, se empleó el modelo ya calibrado para pronosticar los valores de salinidad dentro de un horizonte de predicción de 24 horas. Para ello, el modelo entrenado fue utilizado de forma recurrente, al emplear valores predichos en tiempos anteriores para alimentar a la red en la predicción siguiente. De esta forma, los errores obtenidos para un tiempo t son propagados en la predicción siguiente, afectando al pronóstico generado en un tiempo $t+1$. Este efecto acumulativo del error en la simulación hace que el horizonte de predicción máximo para este modelo se haya establecido en 24 horas. Aún así, los pronósticos obtenidos para los periodos seleccionados muestran desviaciones de no más de 0.3 psu, lo que indica un buen rendimiento.

Los resultados obtenidos demuestran que un modelo basado en ANFIS tal como el propuesto, puede ser recomendable como alternativa para la predicción de la salinidad en tiempo real, en puntos sensibles del medio, afectados por el vertido de salmuera.

6.2. Futuras líneas de investigación

Son varias las líneas de investigación que pueden ser desarrolladas a partir de los resultados presentados en esta tesina. Algunas de ellas están relacionadas con la mejora del modelo propuesto y otras se podrían centrar en el uso de otro tipo de arquitecturas con el fin de desarrollar el modelo más preciso posible, con horizontes de predicción mayores que el aquí propuesto.

Algunas de estas líneas de investigación se describen seguidamente:

- Se propone el uso de series temporales más largas, ya que la calibración del modelo expuesto en este trabajo se realizó a partir de una serie de datos de poco más de un mes. Los beneficios derivados del uso de series temporales largas son varios. Uno de ellos se encuentra relacionado con una limitación ya comentada en el texto, y que viene dada por la incapacidad de los modelos basados en redes neuronales de extrapolar más allá de los valores empleados durante el entrenamiento. Si se emplea una serie temporal de al menos un año, se estarán recogiendo las variaciones estacionales, tanto de las variables océano-meteorológicas como de las condiciones de explotación de la planta, de forma que es mucho más probable que se recojan los

verdaderos rangos de las variables implicadas, previniendo así que, durante la predicción, aparezcan datos fuera del rango dentro del cual la red ha sido entrenada. Esta recomendación lleva directamente a otra, relacionada con la ampliación del análisis estadístico de los datos, ya que sería recomendable incluir estudios de estacionalidad y presencia de otros patrones en las series, que no se han considerado relevantes para la serie de datos empleada.

- Verificación de la fiabilidad de los pronósticos del modelo con la introducción como variables de entrada de valores procedentes de modelos de predicción de oleaje y de las estimaciones realizadas por los sistemas SCADA de las plantas desalinizadoras. Como ya se indicó anteriormente, los pronósticos realizados en este trabajo emplean datos de altura de ola y parámetro de sales procedentes de los registros de la instrumentación del proyecto ASDECO. Sin embargo, para su posible uso como sistema de predicción en tiempo real, el sistema debe ser alimentado con datos de predicción procedentes de otros modelos. Asimismo, es posible plantear la posibilidad de calibrar un modelo con estos datos, de forma que los errores cometidos por los modelos de predicción sean contemplados durante el entrenamiento. De esta forma, al introducir los datos en el modelo para realizar pronósticos no estaremos introduciendo el error cometido por el modelo de predicción de oleaje por ejemplo, sino que al haber sido este introducido en el entrenamiento, se obtendrán los valores de salinidad adecuados.
- Introducción de nuevas variables al modelo. Tal como se ha visto, la conocida "maldición de la dimensionalidad" limita el número de variables de entrada que pueden ser utilizadas para una serie de datos determinada, ya que la extensión de ésta debe crecer exponencialmente con el número de variables de entrada. Sin embargo, para una serie de datos más larga sería posible introducir variables como la dirección del oleaje o la velocidad y dirección del viento, variables con las cuales se han realizado pruebas que aunque menor que el del modelo presentado, mostraban un buen rendimiento. De esta forma, aumentando el número de datos se podrían estudiar combinaciones de variables diferentes, que incluyan más variables explicativas que las aquí utilizadas.

De lo expuesto anteriormente se extrae que, con una serie de datos más extensa, se presentan numerosas posibilidades de investigación, y que a la vista de los buenos resultados obtenidos con la serie empleada, de extensión limitada, de estas investigaciones pueden resultar modelos eficaces, perfectamente capaces de predecir valores de salinidad de forma fiable.

REFERENCIAS

- [1] ACUAMED, 2009. Problemática de las DIAs de las desaladoras. Jornada Avances tecnológicos en el diseño y control del vertido de salmueras en el medio marino. Valencia.
- [2] Ahmed, M., W.H. Shayya, D. Hoey, A. Mahendran, R. Morris y J. Al-Handaly, 2000. Use of evaporation ponds for brine disposal in desalination plants. *Desalination* 130, 155-168.
- [3] ASDECO, 2007-2009. Informes Técnicos del Proyecto.
- [4] Amara, R. and Salanik, G., 1972. Forecasting: From conjectural art toward science. *Technological Forecasting and Social Change*, 3(3): 415-426.
- [5] Albaina, A., 2007. Tecnologías Sostenibles sobre la gestión del agua: La desalación. Estudio del caso en la zona costera del Mediterráneo. Tesis de master. Universitat Politècnica de Catalunya. Càtedra Unesco de Sostenibilitat.
- [6] Bell, J.D., Harmelin-Vivien, M., 1982. Fish fauna of French Mediterranean *Posidonia oceanica* seagrass meadow. Community structure. *Téthys* 10: 337-347
- [7] Bowden, G.J., Dandy, G.C., Maier, H.R., 2003. Data transformation for neural network models in water resources applications. *Journal of Hydroinformatics* 05.4, 245-258.
- [8] BOE: Boletín Oficial del Estado, 2005. Resolución por la que se formula la declaración de impacto ambiental sobre la evaluación del proyecto de «nueva desaladora del Canal de Alicante». BOE num. 155, de 30 de Junio de 2005.
- [9] BOE: Boletín Oficial del Estado, 2005. Resolución por la que se formula la declaración de impacto ambiental sobre la evaluación del proyecto de «Ampliación de la desaladora del Canal de Alicante». BOE num. 274, de 16 de Noviembre de 2005.

- [10] Cebrian E, Duarte CM (2001) Detrital stocks and dynamics of the seagrass *Posidonia oceanica* (L.) Delile in Spanish Mediterranean. *Aquatic Botany* 70: 295-309
- [11] Chang F., Chang Y., 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources*, 29, 1-10.
- [12] Chau KW., Wu CL., Li YS., 2005. Comparison of several flood forecasting models in Yangtze river. *Journal of Hydrologic Engineering* 10, 485-491.
- [13] CORMIX, 1996. User's manual for Cormix: a hydrodynamic mixing zone model and decision support system for pollutant discharges into surface waters. Office of Science and Technology U.S. Environmental Protection Agency Washington, DC 20460.
- [14] Einav, R., Harussi, K., Perry, D., 2002. The footprint of the desalination processes on the environment. *Desalination* 152 (2002) 141–154
- [15] Faraway, J., Chatfield, C., 1998. Time series forecasting with neural networks: a comparative study using the airline data. *Applied Statistics* 47 (2), 231–250.
- [16] Fernández Jiménez, L. A., 2008. Modelos avanzados para la predicción a corto plazo de la producción eléctrica en parques eólicos. Tesis Doctoral. Universidad de La Rioja.
- [17] Fernández-Torquemada, Y., Sánchez-Lizaso, J.L., González-Correa, J.M., 2005. Preliminary results of the monitoring of the brine discharge produced by the SWRO desalination plant of Alicante (SE Spain). *Desalination* 182, 395-402.
- [18] Fernando, D.A.K., Jayawardena, A.W., 1998. Runoff forecasting using RBF networks with OLS algorithm. *Journal of Hydrologic Engineering* 3 (3), 203–209.
- [19] Fortin, V., Ouarda, T.B.M.J., Bobée, B., 1997. Comment on 'The use of artificial neural networks for the prediction of water quality parameters' by H.R. Maier and G.C. Dandy. *Water Resources Research* 33 (10), 2423–2424.

- [20] Gacia E., Duarte C.M., 2001. Sediment retention by a Mediterranean *Posidonia oceanica* meadow: the balance between deposition and resuspension. *Est Coast Shelf Sci* 52: 505-514
- [21] Gacia E., Invers O., Manzanera M., Ballesteros E., Romero J., 2007. Impact of the brine from a desalination plant on a shallow seagrass (*Posidonia oceanica*) meadow. *Est Coast Shelf Sci* 72: 579-590
- [22] García Gómez, A., Revilla Cortezón, J.A., Losada Rodriguez, I., 2009. Modelos de campo lejano. Bases teóricas. Modelación hidrodinámica. Plan Nacional I+D+i: 045/RN08/3.3. Jornada Avances tecnológicos en el diseño y control del vertido de salmueras en el medio marino. Valencia.
- [23] Gautam, D.K., Holz, K.P., 2001. Rainfall-runoff modelling using adaptative neuro-fuzzy systems. *Journal of Hydroinformatics*, 03.1.
- [24] Hammerstrom, D., 1993. Neural Networks at work. *IEEE Spectrum*, pp. 26-32.
- [25] Hernández J.M., Hernández Mascarell A.B., Navarro M., Martín M., Cortés J.M., 2009. Monitoring and decision support systems for impacts minimization of desalination plant outfall in marine ecosystems. *Desalination for the Environment: Clean Water and Energy*. Baden-Baden (Alemania).
- [26] Hettiarachchi, P., Hall, M.J., Minns, A.W., 2005. The extrapolation of artificial neural networks for the modelling of rainfall – runoff relationships. *Journal of Hydroinformatics*, vol. 7, nº 4, 291-296.
- [27] Hsu, K.L., Gupta, H.V., Sorooshian, S., 1995. Artificial neural network modeling of the rainfall–runoff process. *Water Resources Research* 31 (10), 2517–2530.
- [28] Hodges, Ben R., Paula S. Kulis, and Cedric H. David. *Desalination Brine Discharge Model: Final Report*. Submitted to Texas Water Development Board by Center for Research in Water Resources, University of Texas, 2006.

- [29] Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. of the National Academy of Sciences*, 79, pp. 2554-2558.
- [30] Jang, J.R., 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23, nº 3.
- [31] Jang, J.R., 1995. Neuro-fuzzy modelling and control. *Proc. of the IEEE*.
- [32] Jang, J.R., 1996. Input selection for ANFIS learning. In *Proceedings Of The International Conference On Fuzzy Systems*, 1493-1499.
- [33] Jang, J. -S. R, Sun, C. T. and Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing*. Prentice Hall.
- [34] Jirka G., Bleninger T., 2009. Far / near-field transition. Case studies. *Jornada Avances tecnológicos en el diseño y control del vertido de salmueras en el medio marino*. Valencia.
- [35] Jones, C.G., Lawton, J.H., Shachak, M., 1997. Positive and negative effects of organisms as physical ecosystem engineers. *Ecology* 78: 1946-1957
- [36] Kämpf, J., 2009. Impacts of blending on dilution of negatively buoyant brine discharge in a shallow tidal sea. *Marine Pollution Bulletin* 58, 1032–1038
- [37] Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, pp. 574-69.
- [38] Lattemann, S., Höpner, T., 2003. Seawater desalination: Impacts of brine and chemical discharge on the marine environment, Institute for Chemistry and Biology of the Marine Environment (ICBM), University of Oldenburg.
- [39] Lattemann, S., Höpner, T., 2008. Environmental impact and impact assessment of seawater desalination. *Desalination* 220 (2008) 1–15

- [40] Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulagnier, S., 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling* 90, 39–52.
- [41] Leendeertse J.J., 1967. Aspects of a computational model for long water wave Propagation. Rand Corporation, Memorandum RH-5299-RR, Santa Monica.
- [42] Lu, W.Z, Lo, S.M., Fan, H.Y., 2003. Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. *Neurocomputing, Volume 51, Pages 387-400*
- [43] Luk, K.C., Ball, J.E., Sharma, A., 2001. An application of artificial neural networks for rainfall forecasting. *Mathematical and Computer Modelling, Volume 33, Issues 6-7, Pages 683-693*
- [44] McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immament in nervous activity. *Bulletin of Mathematical Biophysics*, 5, pp. 115-133.
- [45] Martín, B., 2006. *Redes neuronales y sistemas borrosos*. 3ª edición. RA-MA.
- [46] Maier, H.R., 1995. Use of artificial neural networks for modelling multivariate water quality time series. PhD Thesis, The University of Adelaide.
- [47] Maier, H.R., Dandy, G.C., 1997a. Determining inputs for neural network models of multivariate time series. *Microcomputers in Civil Engineering* 12 (5), 353–368.
- [48] Maier, H.R., Dandy, G.C., 1997b. Modelling cyanobacteria (blue-green algae) in the River Murray using artificial neural networks. *Mathematics and Computers in Simulation* 43 (3–6), 377–386.
- [49] Maier, H.R., Dandy, G.C., Burch, M.D., 1998. Use of artificial neural networks for modelling cyanobacteria *Anabaena* spp. in the River Murray, South Australia. *Ecological Modelling* 105 (2/3), 257–272.

- [50] Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software* 15, 101–124
- [51] Martínez, B., 2008. Evaluación del estado de las aguas costeras: herramientas de vigilancia basadas en los ecosistemas de *Posidonia oceanica*. Tesis Doctoral. Universidad de Barcelona. Facultad de Biología, Departamento de Ecología.
- [52] Masters, T., 1993. *Practical Neural Network Recipes in C++*. Academic Press, San Diego, CA.
- [53] Mateo, M.A., Romero, J., 1997. Detritus dynamics in seagrass *Posidonia oceanica*: elements for an ecosystem carbon and nutrient budget. *Marine Ecology Prog Ser* 151: 43-53
- [53] Minns, A.W., Hall, M.J., 1996. Artificial neural networks as rainfall–runoff models. *Hydrological Sciences Journal* 41 (3), 399–417.
- [54] Marcé R., 2007. Ter River influence on Sau Reservoir limnology Empirical and watershed-scale modelling. Tesis doctoral. Universitat de Barcelona.
- [55] Mazzella, L., Scipione, M.B., Buia, M.C., 1989. Spatio-temporal distribution of algal and animal communities in a *Posidonia oceanica* meadow. *PSZN: Marine Ecology* 10: 107-129
- [56] Nayak, P.C., Sudheer, K.P., Rangan, D.M., Ramasastri, K.S., 2004. A neuro-fuzzy computing technique for modelling hydrological time series. *Journal of Hydrology, Volume 291, Issues 1-2, 31. Pages 52-66*
- [57] Niepelt, A., Bleninger, T., Jirka, G., 2008. Desalination brine discharge modelling. Coupling of hydrodynamic models for brine discharge analysis. Proc. Intl. Conf. Marine Waste Water Discharges - MWWD 2008, Cavtat (Croatia).
- [58] Palomar P., Losada, I., 2008. Desalinización de agua marina en España: Aspectos a considerar en el diseño del sistema de vertido para protección del medio marino. *Revista de Obras Públicas* 3486, 37-52.

- [59] Payo, A., Cortés, J.M, Molina, R., 2009. Effect of wind and waves on a nearshore brine discharge dilution in the east coast of Spain. *Desalination for the Environment: Clean Water and Energy*. Baden-Baden (Alemania).
- [60] Pergent G, Romero J, Pergent-Martini C, Mateo MA, Boudouresque CF (1994) Primary production, stocks and fluxes in the Mediterranean seagrass *Posidonia oceanica*. *Mar Ecol Progress Series*, 106: 139-146.
- [61] Programas de Vigilancia Ambiental de la desaladora de Alicante. Años 2004, 2005, 2006 y 2007. Departamento de Ciencias del Mar y Biología aplicada. Universidad de Alicante.
- [62] Puertos del Estado, 2001. ROM 0.0, Recomendaciones para obras marítimas, 3, 69-78.
- [63] Refenes, A. -P. N. and Zaprani, A. D. (1999). Neural model identification, variable selection and model adequacy. *Journal of Forecasting*, Vol. 18 (5): 299- 332.
- [64] Rosenblatt, F. *Principles of Neurodynamics*. Spartan Books, Nueva York, 1962.
- [65] Rumelhart DE, Hinton GE, Williams RJ, 1986. Learning representations by backpropagation errors. *Nature*, 323, pp. 533-6.
- [66] Ruiz Fernández, J.M., 2005. Impacto ambiental de las desaladoras sobre las comunidades bentónicas marinas. *Revista ingeniería y Territorio*, 72: Desalación, 40-47.
- [67] Ruiz Mateo, A., 2007. Los vertidos al mar de las plantas desaladoras. *Ambienta*, 51-57.
- [68] Ruiz Mateo, A., Antequera, M., González, J., 2008. Comportamiento y efectos de los vertidos hipersalinos al mar. *Planes de Vigilancia y Control*. V Seminario Internacional de Ingeniería y Operaciones Portuarias. Concepción, Chile.
- [69] Sadhwani, J.J., 2004. Impacto ambiental en la desalación de aguas. VII Congreso Nacional del Medio Ambiente.

- [70] Sahoo G.B., Ray C., Wade H.F., 2005. Pesticide prediction in ground water in North California domestic wells using artificial neural works. *Ecological Modelling*, 183, 29-46.
- [71] Sánchez-Lizaso, J.L., Romero, J., Ruiz, J., Gacia, E., Buceta, J.L., Invers, O., Fernández-Torquemada, Y., Mas, J., Ruiz-Mateo, A., Manzanera, M., 2008. Salinity tolerance of the Mediterranean seagrass *Posidonia oceanica*: recomendations to minimize the impact of brine dischargues from desalination plants. *Desalination* 221, 602-607.
- [72] Shanin, M.A., Maier, H.R., and Jaksa, M.B., 2004b. Data division for developing neural networks applied to geotechnical engineering. *Journal of Computing in Civil Engineering*, ASCE, 18(2), 105-114.
- [73] Shiklomanov I.A., 1999. World water resources and their use. SHI/UNESCO.
- [74] Smith, M. (1993). *Neural Networks for Statistical Modeling*. Van Nostrand Reinhold, New York.
- [75] Torres Corral, M., 2005. Desalación y planificación hidrológica hoy. *Revista ingeniería y Territorio*, 72: Desalación, 8-16.
- [76] UNEP, 2008. *Desalination. Resource and Guidance Manual for Environmental Assessments*.
- [77] Universidad de Alicante. Programas de Vigilancia Ambiental de la desaladora de Alicante. Años 2004, 2005, 2006 y 2007. Departamento de Ciencias del Mar y Biología aplicada.
- [78] Vairappan, C., Tamura, H., Gao, S., Tang, Z., 2009. Batch type local search-based adaptative neuro-fuzzy inference system (ANFIS) with self-feedbacks for time-series prediction. *Neurocomputing* 72, 1870-1877.
- [79] Vicente, C., 2005. Descripción del funcionamiento de la desaladora de agua de mar del Canal de Alicante. *Dyna*, nº 1, 20-24.

[80] Zadeh, L.A., 1965. Fuzzy Sets. *Information & Control*, 8, pp. 338-353.

[81] www.cth.gva.es/areas/SIG/sigshp/cartografia_basica/lics/fichas_lics/lic_tabarca.pdf