



Título del Trabajo Fin de Máster:

***MAPEO DE PROCESOS
AUTORREGRESIVOS MEDIANTE RED
NEURONAL PERCEPTRON
MULTICAPA CON COMPONENTE
ESTOCÁSTICA. APLICACIÓN A LA
CUENCA DEL JÚCAR.***

Intensificación:

ORDENACIÓN, RESTAURACIÓN Y GESTIÓN DE CUENCAS

Autor:

BAU IZQUIERDO, RAFAEL

Director/es:

DR. MANUEL PULIDO VELÁZQUEZ

DR. JUAN CAMILO OCHOA RIVERA

Fecha: SEPTIEMBRE, 2012

Título del Trabajo Fin de Máster:

Mapeo de procesos autorregresivos mediante red neuronal perceptron multicapa con componente estocástica. Aplicación a la cuenca del Júcar.

Autor: ***Bau Izquierdo, Rafael***

Tipo	A <input type="checkbox"/> B <input checked="" type="checkbox"/>	Lugar de Realización	VALENCIA
Director	<i>DR. Manuel Pulido Velázquez</i>	Fecha de Lectura	<i>Sept., 2012</i>
Codirector1	<i>DR. Juan Camilo Ochoa Rivera</i>		
Codirector2			
Tutor			

Resumen:

En la actualidad se está observando una disminución a escala global de la disponibilidad de recursos hídricos. Este problema se ha intensificado en las últimas décadas debido al cambio climático y a la contaminación de las aguas (ya sean masas de agua superficiales como subterráneas), hecho que repercute directamente en la calidad de vida de la población y en la producción de alimentos a escala mundial.

De este modo, ante los escenarios actuales de preocupación y de incertidumbre en el que se encuentra la sociedad, se hace necesario desarrollar nuevos modelos que sean capaces de prever posibles sequías y poder así, adoptar medidas tendentes a la minimización de sus efectos.

Si se atiende a los modelos tradicionales, estos se basan en el enfoque probabilístico, ya que permite efectuar la simulación de la gestión del sistema a gracias a las series equiprobables generadas a partir de los datos históricos disponibles para estimar los probables estados del sistema en el futuro. Dentro de esos modelos de generación de series, los de mayor uso tienden a no reproducir de forma satisfactoria las características de las sequías históricas, entre otras propiedades.

Por este motivo, en el presente trabajo final de máster se ha tratado de estudiar las redes neuronales artificiales para poder incorporarse de forma satisfactoria en el campo de la Hidrología Operacional y aportar una mayor calidad a la hora de generar series sintéticas.

Las aplicaciones del estudio consistieron en generar series sintéticas de caudales mensuales empleando los modelos tradicionales y un modelo de red neuronal perceptron multicapa con componente estocástica. En un primer estudio se analizaron los resultados para un modelo univariado. Estos resultados aparecen en el estudio Mapping autoregressive processes with a stochastic three-layer perceptron. Application to hydrologic time series modeling (Ochoa-Rivera et al., en revisión) y después, se aplicó el mismo procedimiento para un modelo bivariado en otra cuenca de estudio diferente. De los resultados se extrae que el modelo perceptron multicapa es más fiable y robusto que los métodos tradicionales, presentando un mejor rendimiento y fiabilidad.

Abstract:

Nowadays can be observed a severe global reduction of water resources, this problem has been increased due to the global warming and water pollution (surface water and groundwater) as a result affect global food production and quality of life.

Because of this, it is necessary to devevelop new models with ability to predict drought time series and offer measures that reduce the risk in future scenarios that may occur.

Classical models study this problem from the probabilistic viewpoint, meaning those kinds of models which use historical time series to predict future time series. For this reason, it is necessary to undertake further studies in Hydrology to develop new models for generating synthetic time series. Classical models use to exhibit some problems when they are applied to reproduce historical droughts.

In this study have tried to apply artificial neural networks to improve the classical models and obtain better results in time series.

The applications of the study consisted of generating synthetic monthly time series using classical models and a model of multilayer perceptron neural network with stochastic component. In a first study, we observed better results for univariate model, these results appear in the study "Study Autoregressive Processes Mapping with a stochastic three-layer perceptron. Application to hydrologic time series modeling" (Ochoa-Rivera et al., under review) and then the same procedure was applied to bivariate model in another separate study basin.

The results showed that the proposed neural network is more reliable and robust than classical methods, thus presenting a better performance and reliability.

Resum:

En l'actualitat s'observa una disminució a escala global de la disponibilitat de recursos hídrics. Aquest problema s'ha intensificat en les últimes dècades per l'acció del canvi climàtic i la contaminació de les aigües (tant superficials com subterrànies). Aquest fet repercuteix directament en la qualitat de vida de la població i en la producció d'aliments a escala mundial.

Per aquesta raó, davant dels escenaris de preocupació existents i d'incertesa en què es troba la societat moderna, cal desenvolupar nous models pugen prevenir episodis de sequera i poder adoptar les mesures més adequades per a minimitzar els seus efectes desfavorables.

Si s'estudien els models tradicionals, es veu que aquests es basen en el mètode probabilístic, permetent efectuar simulacions de la gestió del sistema gràcies a les sèries equiprobables generades a partir de l'informació històrica per a poder estimar situacions possibles del sistema en el futur. Dins dels models clàssics de generació, els més utilitzats no acaben de reproduir satisfactòriament les característiques històriques de les sequeres entre d'altres.

En aquest treball de fi de màster, s'ha tractat d'estudiar les xarxes neuronals artificials per a poder implementar-les de forma satisfactòria a l'àmbit de l'Hidrologia Operacional i aportar una major precisió a l'hora de generar sèries sintètiques.

Les aplicacions de l'estudi han consistit a generar sèries històriques de cabals mensuals a partir dels models tradicionals i amb una xarxa neuronal perceptró multicapa amb component estocàstica. En un primer estudi es treballà amb un model univariat, els resultats s'han inclòs a l'article Mapping autoregressive processes with a stochastic three-layer perceptron. Application to hydrologic time series modeling (Ochoa-Rivera et al., en revisió) I, després, es va aplicar el mateix procediment però amb un escenari bivariat i en un altra conca d'estudi. Dels resultats obtinguts en podem extraure que la xarxa neuronal perceptró multicapa es més robusta que els models tradicionals, presentant un major rendiment i fiabilitat.

Palabras clave:

Modelación estocástica, redes neuronales estocásticas, índices de bondad de ajuste, sequías, procesos autorregresivos

Índice general

Capítulo 1. Introducción	1
1.1 Metodología y aportaciones	1
1.2 Gestión de los recursos hídricos.	3
1.2.1 Enfoques clásicos.	3
1.2.2 La gestión Integrada del agua.	4
1.2.2.1 Sistemas de ayuda de decisión (SAD).	4
1.2.2.2 La gestión sostenible de los recursos hídricos.	6
Capítulo 2. Estado del arte	9
2.1 Generación sintética de escenarios hidrológicos.	9
2.2 Sequías en los sistemas de recursos hídricos.	13
2.2.1 Variables necesarias para la predicción de sequías.	14
2.2.1.1 Modelos de regresión.	15
2.2.1.2 Análisis de series temporales.	16
2.2.1.3 Modelos probabilísticos.	16
2.2.1.4 Redes neuronales.	17
2.2.1.5 Modelos híbridos.	18
2.2.2 Predicción a largo plazo de las sequías.	18
2.2.3 Modelación de sequías bajo escenarios de cambio climático.	19
2.2.4 Gestión de los episodios de sequía mediante sistemas de ayuda de decisión (SAD).	20
2.3 Las redes neuronales artificiales.	22
2.3.1 Desarrollo de las redes neuronales.	22
2.3.2 Aplicación de las redes neuronales en la gestión de los recursos hídricos.	23
2.3.3 Modelación de series temporales.	25
Capítulo 3. Estadísticos de bondad de ajuste en la modelación hidrológica	31
3.1 Índices de bondad de ajuste de modelos. Aplicación a la Hidrología.	31
Capítulo 4. Redes neuronales artificiales estocásticas	43
4.1 Introducción.	43
4.1.1 Modelo matemático de las redes neuronales determinísticas.	45
4.1.2 Proceso de aprendizaje.	46
4.1.2.1 Aprendizaje supervisado.	46
4.1.2.2 Aprendizaje no-supervisado.	47
4.1.3 El perceptron multicapa.	48

4.1.4	Funciones de activación empleadas.	49
4.1.5	El algoritmo de retropropagación del error (EBLA).	52
4.2	Construcción del modelo de red neuronal con la componente estocástica.	54
4.2.1	Exploraciones iniciales.	54
4.2.2	Preprocesamiento de caudales.	59
4.2.3	Componente determinística multivariada.	63
4.2.4	Entrenamiento de la red neuronal.	65
4.2.5	Red neuronal con componente estocástica.	66
4.3	Modelación con la red neuronal.	68
4.3.1	Ajuste del modelo.	68
4.3.2	Validación del modelo.	69
 Capítulo 5. Casos de estudio		71
5.1	Consideraciones previas.	72
5.1.1	Estadísticos de sequía y almacenamiento.	73
5.2	Caso Univariado. Cabecera del Tajo.	77
5.2.1	Descripción del sistema.	77
5.2.2	Ajuste del modelo estocástico de red neuronal.	78
5.2.3	Rendimiento de la red neuronal propuesta frente a los modelos autorregresivos.	79
5.3	Caso bivariado. Cuenca del Júcar.	84
5.3.1	Descripción del sistema.	84
5.3.2	Ajuste del modelo estocástico de red neuronal.	85
5.3.3	Rendimiento de la red neuronal propuesta frente a los modelos autorregresivos.	92
5.4	Parsimonia de los modelos aplicados.	99
 Capítulo 6. Conclusiones y líneas futuras de investigación		103
6.1	Conclusiones generales.	104
6.2	Comparación de los modelos empleados.	105
6.3	Líneas futuras de investigación.	106
 Referencias		107

Índice de figuras

Figura 2.1: Caracterización probabilística de los parámetros de sequía.	17
Figura 2.2: Esquema simplificado de funcionamiento de un sistema de ayuda de decisión (SAD).	21
Figura 3.1: Relación de CE con α y γ (asumiendo $\beta_0 = 0$).	34
Figura 4.1: Configuración tipo de una red neuronal artificial.	43
Figura 4.2: Esquema de una red neuronal con todos sus elementos.	45
Figura 4.3: Configuración tipo de una red perceptron multicapa.	48
Figura 4.4: Funciones escalón a) rango [0; 1] b) rango [-1; 1]	49
Figura 4.5: Funciones mixtas a) rango [0; 1] b) rango [-1; 1]	50
Figura 4.6: Funciones sigmoidales (a) logística; (b) tangente hiperbólica.	51
Figura 4.7: Imagen de una superficie de error de la red MLP.	52
Figura 4.8: red MLP 3-2-1 para el mapeo de la función $y = x^2$	55
Figura 4.9: Mapeo de la función $y = x^2$ con una red MLP 3-2-1 (a) entrenamiento (b) validación.	55
Figura 4.10: red MLP 1-2-1 para el mapeo de una serie de Henon.	56
Figura 4.11: Mapeo de la función $yt+1 = yt + 3[yt(1 - yt)]$ con una red MLP 1-2-1 (a) entrenamiento (b) validación.	56
Figura 4.12: Mapeo de la función AR(1), $z_t = 0.796z_{t-1} + 0.481t$ con una red MLP 1-2-1 (a) entrenamiento (b) validación.	57
Figura 4.13: Mapeo de modelos AR(2) (línea discontinua) con redes MLP 2 - 1 - 1 (línea continua) en fase de entrenamiento (a) AR(2)1 (b) AR(2)2 (c) AR(2)3 (d) AR(2)4.	58
Figura 4.14: Mapeo de modelos AR(2) (línea discontinua) con redes MLP 2 - 1 - 1 (línea continua) en fase de validación (a) AR(2)1 (b) AR(2)2 (c) AR(2)3 (d) AR(2)4.	58
Figura 4.15: Error de mapeo de los procesos AR(2).	59
Figura 5.1: Serie tipo para la caracterización de las propiedades de las sequias. Intensidad, magnitud y duración.	74
Figura 5.2: Localización del embalse de Bolarque y del sistema de Cabecera en la Cuenca hidrográfica del Tajo.	77
Figura 5.3: Evolución del error cuadrático medio para la RNA 2-1-1.	79
Figura 5.4: Estadísticos básicos del modelo univariado.	80
Figura 5.5: Intervalos de confianza del 95% para el valor del caudal medio mensual del embalses de Bolarque.	80
Figura 5.6: Porcentaje de error de predicción de los estadísticos de sequía del modelo univariado.	82
Figura 5.7 Coeficiente de Hurst para el escenario univariado.	82

Figura 5.8: Capacidades mínimas de embalse para el escenario univariado.....	83
Figura 5.9: Localización de los embalses de Alarcón y Contreras en la Cuenca hidrográfica del Júcar.	84
Figura 5.10: Porcentaje de la varianza explicada para la media y desviación típica de las series históricas de Alarcón y Contreras.	85
Figura 5.11: Influencia del número de armónicos y su relación con el error asociado. Modelos AR(1) y AR(2) para la media y desviación típica de las series históricas de Alarcón y Contreras.....	86
Figura 5.12: Configuración red neuronal propuesta.....	89
Figura 5.13: Evolución del error cuadrático medio para la RNA 4-4-2.	90
Figura 5.14: Estadísticos básicos para el embalse de Alarcón en el modelo bivariado.	92
Figura 5.15: Estadísticos básicos para el embalse de Contreras en el modelo bivariado.	93
Figura 5.16: Intervalos de confianza del 95% para el valor de los caudales medios mensuales de los embalses de Alarcón y contreras.....	93
Figura 5.17: Valores de sesgo para los Alarcón y Contreras en el modelo bivariado iniciando la generación de series sintéticas en Diciembre.	95
Figura 5.18: Porcentaje de error de predicción de los estadísticos de sequía del modelo bivariado para el embalse de Alarcón.	96
Figura 5.19: Porcentaje de error de predicción de los estadísticos de sequía del modelo bivariado para el embalse de Contreras.	97
Figura 5.20: Coeficiente de Hurst para el escenario bivariado.....	98
Figura 5.21: Capacidades mínimas de embalse para el escenario bivariado.....	98

Índice de tablas

Tabla 1.1: Principales fases y acciones a llevar a cabo del proceso de planificación hidrológica de la Directiva Marco del Agua.	7
Tabla 2.1: Equivalencia de los modelos estadísticos y funciones de activación empleadas en las redes neuronales artificiales.....	24
Tabla 2.2: Equivalencia de los modelos estadísticos y número de nodos ocultos empleadas en las redes neuronales artificiales.....	24
Tabla 3.1: Estadísticos propuestos para el análisis de los diferentes modelos.	41
Tabla 5.1: Valores resultantes de la reducción porcentual del caudal medio histórico Bolarque (escenario univariado).	72
Tabla 5.2: Valores resultantes de la reducción porcentual del caudal medio histórico de los embalses de Alarcón y Contreras (escenario bivariado).....	72
Tabla 5.3: Resultados de los índices empleados para la comparación de los diferentes modelos.....	81
Tabla 5.4: Valores de la media y la desviación típica de Alarcón en función del número de armónicos empleados.	87
Tabla 5.5: Valores de la media y la desviación típica de Contreras en función del número de armónicos empleados.	88
Tabla 5.6: Pesos de la conexión de la red neuronal MLP 4-4-2.....	91
Tabla 5.7: Resultados de los índices empleados para la comparación de los diferentes modelos Alarcón.....	94
Tabla 5.8: Resultados de los índices empleados para la comparación de los diferentes modelos Contreras.....	94
Tabla 5.9: Verificación de las reglas de parsimonia de Weigend et al. (1); Lachtermacher y Fuller (2) y Ochoa-Rivera (3).	100
Tabla 5.10: Índice de parsimonia globales de los modelos aplicados en los casos de estudio del modelo univariado.....	100
Tabla 5.11: Índice de parsimonia globales de los modelos aplicados en los casos de estudio del modelo bivariado.....	101
Tabla 5.12: Resultados del AIC de los modelos aplicados en los casos de estudio del modelo bivariado.	101
Tabla 5.13: Resultados del FPE de los modelos aplicados en los casos de estudio del modelo bivariado.	102

Capítulo 1

Introducción

Este primer capítulo está dedicado a la descripción de los aspectos metodológicos y de procedimiento empleados en este Trabajo Fin de Máster; también se expone su finalidad y el aporte que dicho aporte puede representar para estado del conocimiento.

1.1 Metodología y aportaciones.

El objetivo general de este trabajo es estudiar las ventajas de las redes neuronales artificiales (RNA) perceptron multicapa con componente estocástica sobre los modelos autorregresivos clásicos en la generación de series sintéticas de caudales mensuales para su aplicación en estudios de planificación y gestión de sistemas de recursos hídricos.

Los objetivos específicos son:

- a) Comparar mediante un caso de estudio real la bondad de ajuste del modelo RNA frente a los modelos autorregresivos en cuanto a los estadísticos más relevantes en el ámbito de la planificación de los sistemas de recursos hídricos.
- b) Comprobar que el modelo de red neuronal empleado reproduce bien los estadísticos más relevantes de sequía para la gestión de los recursos hídricos.
- c) Estudiar la capacidad de generalización de las redes neuronales perceptron multicapa con componente estocástica para la generación sintética de caudales en diferentes escenarios.

En este proyecto se investiga la generación de series sintéticas de caudales a escala mensual con métodos no paramétricos, no lineales, frente a los modelos paramétricos clásicos (ej. modelos ARMA, Salas et al., 1980). El principal problema observado con los métodos paramétricos lineales, es que a menudo no son capaces de reproducir de forma satisfactoria la estructura de dependencia temporal de las series hidrológicas. Los estadísticos de dependencia espacial, persistencia temporal y estadísticos de sequías no suelen ser bien preservados por dichos métodos en las series sintéticas. Aquí

se analizan las mejoras conseguidas con la aplicación de un modelo basado en redes neuronales, propuesto por Ochoa-Rivera (2002).

La configuración de la red neuronal artificial se ha hecho en cada caso de estudio aplicando el principio de parsimonia, es decir, se ha tratado de buscar la arquitectura más sencilla posible que proporcione unos resultados satisfactorios.

La experimentación se ha llevado a cabo en distintos sistemas de recursos hídricos españoles. Para el análisis univariado, se ha empleado la serie histórica de caudales mensuales (octubre de 1940 - septiembre de 1993) del embalse de Bolarque en la cabecera del río Tajo. Los resultados aparecen descritos en el artículo Mapping autoregressive processes with a stochastic three-layer perceptron (Ochoa-Rivera et al., en revisión). Por otro lado, el estudio bivariado se ha realizado a partir de la información disponible de las series mensuales históricas de los embalses de Alarcón y Contreras en el periodo comprendido entre octubre de 1940 y septiembre de 2006, que forman parte del sistema de cabecera de la cuenca del río Júcar. Para la obtención de los caudales sintéticos se han empleado los modelos MODESTE, GENESIS y SERENA (UPV, 2000a; 2000b; 2001).

La comparación de los resultados obtenidos a partir de los modelos clásicos y de la red neuronal se ha llevado a cabo en dos niveles. En primer lugar se ha comparado las diferencias entre los estadísticos básicos sintéticos (media, desviación típica, sesgo y coeficiente de autocorrelación) y sus homólogos históricos, proponiendo para esta comparación una serie de estadísticos básicos que permitan determinar cual proporciona mejores resultados de forma objetiva. En segundo lugar, empleando los estadísticos de sequía a escala mensual y anual, se comparan las magnitudes de las probabilidades de fallo de las demandas del sistema de recursos hídricos, ya que estos son los índices de gestión en los que más influyen las sequías de las series de caudales.

Por último se presentan las conclusiones obtenidas y las futuras líneas de investigación.

1.2 Gestión de los recursos hídricos.

La gestión de los sistemas de los recursos hídricos es un tema especialmente delicado y uno de los grandes retos a los que la humanidad se enfrenta en el siglo XXI. Son numerosas las metodologías clásicas empleadas para su estudio durante el siglo pasado, pero a su vez, se hace necesaria la inclusión de las nuevas herramientas informáticas para abrir nuevos campos de investigación.

Así pues, se considera importante realizar una breve descripción acerca de la gestión de los recursos hídricos, la generación de escenarios hidrológicos y la aplicación de las redes neuronales en este ámbito.

1.2.1 Enfoques clásicos.

Para poder comprender los sistemas de recursos hídricos y sus interacciones, tanto internas como externas, hacen falta desarrollar teorías, técnicas, modelos y algoritmos, mediante los cuales se pueda tener una aproximación a su realidad física. Si bien ese conjunto de teorías y métodos se vienen desarrollando desde hace bastante tiempo, es a partir de la década de los años cincuenta cuando empiezan a desarrollarse de forma más notoria y acelerada, gracias en parte, a la aparición de los Análisis de sistemas utilizando los nuevos sistemas informáticos.

El Análisis de sistemas puede decirse que es un enfoque de tipo analítico que, a partir de hipótesis, objetivos y criterios preestablecidos, permite seleccionar una entre varias alternativas posibles, mediante la aplicación de técnicas matemáticas de diseño y planificación (Andreu, 1993; Simonovic, 2000).

La principal herramienta del Análisis de sistemas es la simulación matemática. La simulación permite predecir condiciones específicas, siendo una herramienta más a tener en cuenta a la hora de tomar decisiones.

Para llevar a cabo la simulación de un sistema de recursos hídricos se puede utilizar el siguiente esquema (Andreu, 1993):

- Construcción de un modelo matemático.
- Calibración y validación del modelo.
- Estudio del caso a analizar.
- Análisis e interpretación de resultados.

Es posible repetir con distintas iteraciones el estudio en el caso que los resultados no sean satisfactorios.

Según Wurbs (1993), los modelos de simulación pueden clasificarse en dos grandes grupos, dependiendo de la posibilidad de ser aplicados a un sistema de recursos hídricos en particular o a cualquier sistema en general: modelos específicos y modelos de propósito general.

1.2.2 La gestión Integrada del agua.

La gestión del agua debe incluir “la integración, coordinación y administración de actividades humanas entre los límites naturales de una cuenca hidrográfica o demarcación” (Reimold, 1998). Este enfoque siempre ha estado condicionado por dos cuestiones fundamentales, por un lado, la cantidad de agua de la que se dispone (oferta) y, por otro lado, la cantidad de agua que se necesita (demanda). A pesar de que estas cuestiones siempre han estado presentes a la hora de gestionar los recursos hídricos, en las últimas décadas los impactos ambientales y socio-económicos se han hecho más persistentes, siendo necesaria una mejor gestión de los recursos hídricos. Además, las capacidades de simulación, modelación y predicción se han desarrollado de forma importante en las últimas décadas junto a una mayor disponibilidad de datos reales (Loucks et al.1981).

Por tanto, los nuevos retos a los que nos enfrentamos en el siglo XXI pueden resumirse en:

- Sostenibilidad de los recursos hídricos a largo plazo.
- El agua es un elemento vital para los seres vivos y para la economía.
- Integrar los aspectos económicos y ecológicos del agua en las cuencas hidrográficas (enfoque holístico).
- Incluir a todos los agentes implicados en la toma de decisiones.

1.2.2.1 Sistemas de ayuda de decisión (SAD).

Los sistemas de ayuda de decisión surgen tímidamente en la década de los setenta y a partir de los ochenta comienzan a publicarse artículos acerca de su utilización. Ya en los noventa es cuando comienzan a generalizarse y surgen programas específicos (por ejemplo el NELUP, Dunn et al. 1996) donde se integran modelos de economía, ecología e hidrología.

El desarrollo de los SAD se fundamenta en gran parte en la evolución de numerosos programas que simulaban los procesos hidrológicos de cuenca (lluvia-escorrentía). Se puede decir que estos primeros sistemas fueron la base de los actuales SAD, ya que acumulan más tiempo de desarrollo y en la actualidad se pueden encontrar implícitamente dentro de los mismos.

En la actualidad, las técnicas tradicionalmente empleadas dentro de algunos modelos de simulación, como lo son la programación lineal, la programación dinámica, y la programación no-lineal, han ido abriendo espacio a nuevos enfoques matemáticos como algoritmos de evolución o las teorías de conocimiento basadas en los sistemas biológicos, de los cuales hacen parte de la programación evolutiva, los algoritmos genéticos, la lógica difusa, las redes neuronales y la teoría del aprendizaje, entre otras.

Ahora, tienden a implementarse dentro de los sistemas de soporte de decisión los sistemas de información geográfica (SIG), una herramienta nueva que permite ubicar en el espacio todo el sistema y facilitar de este modo su comprensión y correlación espacial ayudando aún más a la toma de decisiones.

A la vista de lo citado anteriormente, el SAD es una herramienta que emplea la informática y a partir de la información disponible y los criterios de simulación deseados, arroja unos resultados útiles para los gestores que deben de tomar las decisiones en la vida real. Además, hay que resaltar que este sistema puede responder de forma interactiva con los gestores y así poder predecir diferentes resultados en función de las diferentes prioridades seleccionadas. También puede definirse un SAD como un ambiente en el que se integran modelos, sistemas de información y bases de datos, concebido para responder de manera interactiva al gestor los planteamientos que éste se formula acerca del comportamiento del sistema de recursos hídricos (Andreu et al., 1996).

Los componentes principales según Reitsma et al. (1996) de un SAD son:

- Información de estado.
- Información de proceso.
- Herramientas de evaluación.

La *información de estado* hace referencia a los datos que representan la situación del sistema de recursos hídricos en un momento determinado, suelen emplearse bases de datos (elementos que constituyen el núcleo de los actuales SAD). La *información de proceso* representa el comportamiento del sistema a lo largo del tiempo, dentro de este componente se encuentran los modelos que simulan la dinámica propia del sistema de recursos hídricos. Las *herramientas de evaluación* son las utilidades de software que permiten transformar los datos brutos del sistema en información relevante para la toma de decisiones; es decir, por medio de técnicas de filtrado, modificación y presentación de esos datos en forma idónea, se hace la evaluación de alternativas.

Como ejemplo de estos sistemas se puede citar el sistema de ayuda de decisión AQUATOOL (Andreu, 1996) desarrollado e implementado por el departamento de Ingeniería Hidráulica y Medio Ambiente de la Universitat

Politécnica de València, el cual consiste en tres unidades de control primario: un módulo para la simulación de la gestión de cuencas (SIMGES), un módulo para la optimización de la gestión de cuencas (OPTIGES), un módulo para la modelación del flujo en acuíferos (AQUIVAL). Como complementos, se han ido desarrollando otras aplicaciones como son: dos módulos para evaluación de riesgo (SIMRISK), otro módulo desarrollado para el análisis de la calidad de aguas en las cuencas (GESCAL), un módulo para calcular la incidencia económica de la gestión de las aguas (EVALGES), un módulo en desarrollo para los indicadores de calidad biológicos de las cuencas (CAUDECO) y el módulo para la generación de series sintéticas de caudal a partir de series históricas (MASHWIN).

Una de las ventajas de este sistema es su facilidad para implementar datos y la comodidad de observar y modificar los resultados gracias a su interfaz gráfica. Entre sus principales funciones destaca la simulación y optimización en la gestión de cuencas hidrográficas debido al uso conjunto de aguas superficiales y subterráneas. El manejo gráfico de las variables de decisión y de los parámetros de gestión permite dar respuesta a las diferentes incógnitas que puedan plantearse.

El SAD AQUATOOL ha sido desarrollado y aplicado en casos reales de estudio en las cuencas de los ríos Tajo y Segura, entre otros. Además se ha empleado para la planificación y gestión de varios sistemas de recursos hídricos en España (Confederación Hidrográfica del Júcar, 2001) y en otros países europeos, así como dentro de los trabajos del Libro Blanco del Agua en España (Ministerio de Medio Ambiente, 1998).

1.2.2.2 La gestión sostenible de los recursos hídricos.

Algunos autores (Chaturvedi, 2000; Loucks, 2000) han señalado la necesidad de gestionar los sistemas de recursos hídricos de forma sostenible, entendidos como aquéllos que se diseñan y gestionan para contribuir plenamente, ahora y en el futuro, a los objetivos de la sociedad, preservando su integridad hidrológica y medioambiental (UNESCO, 1999).

En Europa surge este compromiso firme de gestión del agua y protección del medio ambiente a partir de la Directiva Marco del Agua (2001), las principales fases y acciones se recogen en la tabla 1.1 y, en previsión de las necesidades que van a surgir para el desarrollo de los nuevos Planes Hidrológicos de Cuenca, se empiezan a utilizar los SAD para poder racionalizar mejor los recursos hídricos disponibles.

Tabla 1.1: Principales fases y acciones a llevar a cabo del proceso de planificación hidrológica de la Directiva Marco del Agua.

Fuente: Directiva 2000/60/CE del Parlamento Europeo y del Consejo, de 23 de octubre de 2000.

Fases principales	Acciones
Evaluación de escenarios actuales y potenciales	<ul style="list-style-type: none"> - Caracterizar la escena de actuación - Evaluación del estado actual y análisis de lagunas actuales.
Establecimiento de objetivos	<ul style="list-style-type: none"> - Análisis de lagunas de información. - Establecimiento de objetivos medioambientales.
Desarrollo de programas de medidas con toma de decisiones	<ul style="list-style-type: none"> - Establecimiento de programas de medidas. - Realización de planes hidrológicos de cuenca. - Implementación de programas de medidas y preparación de informes.
Fases transversales	<ul style="list-style-type: none"> - Establecimiento de programas de control. - Evaluación de la primera y segunda fase. - Información, consulta y participación activa de los grupos de interés (stakeholders).

A día de hoy es un hecho que a la hora de planificar y gestionar una cuenca hidrográfica se deben tener en cuenta los requisitos ambientales. Para responder a este reto se han venido desarrollando en las dos últimas décadas varios tipos de nuevas tecnologías y enfoques, que intentan ser, dentro del proceso de toma de decisiones en problemas de gestión, un medio de apoyo más eficiente que la mera aplicación de modelos de forma independiente. Dentro de esta corriente se encuentran los sistemas soporte a la decisión y de este modo, el presente trabajo pretende trabajar en esa dirección.

Por tanto, los objetivos futuros en el ámbito de los recursos hídricos deberán encaminarse a integrar tanto conocimientos como herramientas que permitan estimar los impactos de las decisiones actuales en el futuro a medio-largo plazo.

Para poder aplicar políticas *sostenibles* se debe de evaluar en muchos casos futuros inciertos y además en el ámbito de los recursos hídricos se ha de tener presente la aleatoriedad de los mismos. De hecho, la escasez del agua es una de las principales preocupaciones dentro de la gestión de los sistemas de recursos hídricos, y la necesidad de perfeccionar los métodos de su previsión se ha convertido en el motor de nuevos desarrollos teóricos y de enfoques dirigidos a comprender mejor un escenario futuro incierto.

Capítulo 2

Estado del arte

En el presente trabajo final de máster se ha aplicado la metodología propuesta por Ochoa-Rivera (2002) para el estudio de las redes neuronales con componente estocástica comparándolas con los modelos clásicos utilizados en el ámbito de la Hidrología Operacional. Por esta razón, previamente se ha considerado necesario realizar una breve revisión bibliográfica de la evolución histórica en los modelos y temas empleados en este trabajo.

2.1 Generación sintética de escenarios hidrológicos.

La Hidrología Operacional podría definirse como la parte de la hidrología que trata la derivación de series temporales sintéticas a partir de los registros históricos (OMM, 1971). Dentro del estudio de los sistemas de recursos hídricos, la modelación de series temporales hidrológicas ha ido ganando cada vez más terreno desde el inicio de su desarrollo. En este capítulo se hace una breve reseña histórica acerca de la generación de caudales sintéticos.

En Salas et al. (1980) se presenta un resumen acerca del inicio y desarrollo de la modelación de series temporales hidrológicas, en el cual se anota que ya desde principios del siglo pasado algunos autores habían demostrado la idoneidad de emplear la estadística para la generación de series sintéticas. Uno de los avances más significativos para la generación de caudales sintéticos es la aplicación de los modelos autorregresivos, destacando los trabajos de Thomas y Fiering (1962) y Yevjevich (1963), dan inicio de manera formal al desarrollo de la modelación estocástica. No obstante, una de las principales y primeras críticas que se les hace a los modelos autorregresivos es su baja memoria, por estar basados en procesos estocásticos gaussianos, lo cual les impide reproducir las características de las series a largo plazo (Mandelbrot y Wallis, 1968). Por esta razón, dichos autores proponen la utilización de ruidos fraccionales basados en el fenómeno de Hurst.

Al poco tiempo, Mandelbrot (1971) publica un nuevo trabajo donde reconoce los inconvenientes de la técnica inicialmente propuesta, en lo que a su aplicación se refiere, e introduce como simplificación de la misma el generador

veloz de ruidos gaussianos fraccionales (fast fractional gaussian noise generator), con el que intenta reducir el número de operaciones necesarias para generar los ruidos sin deteriorar la estructura de dependencia temporal. Pero como otros autores indican (Jackson, 1975; Bras y Rodríguez-Iturbe, 1985), este método no consigue instalarse en la Hidrología Estocástica por la gran cantidad de operaciones a realizar y la dificultad de estimar los parámetros de Hurst.

En medio de la controversia entre los modelos autorregresivos y los modelos de ruido fraccional, Box y Jenkins (1970) presentan nuevas herramientas para el análisis y la modelación de series temporales, relacionadas específicamente con mejoras en la estimación de parámetros, verificación de las condiciones que estos cumplen y de la revisión de las hipótesis del modelo. Además, se presenta otro de los nuevos desarrollos, la adición de una componente de media móvil (MA) a los modelos autorregresivos, estableciendo así una nueva categoría de modelos, denominados autorregresivos de media móvil (ARMA). La componente de media móvil va más asociada a la descripción de la porción del caudal correspondiente a la escorrentía superficial, mientras que la componente autorregresiva se vincula más a la fracción subterránea del caudal, razón por la cual los modelos ARMA pueden resultar más parsimoniosos que los modelos puramente autorregresivos cuando la escorrentía superficial tiene un peso importante en la serie de caudales (Salas et al., 1980).

Con el fin de hacer nuevos aportes a la solución del problema de la dependencia temporal a largo plazo, Valencia y Schaake (1973) presentan un estudio en el que señalan que los modelos autorregresivos pueden ser inapropiados en casos en los que se requiera reproducir ciertos estadísticos relevantes en una escala temporal superior a la modelada; es decir, un modelo autorregresivo de escala mensual probablemente no reproduzca ciertos estadísticos a nivel anual. Todos los trabajos posteriores sobre modelos de desagregación parten de este estudio, ya que es éste el que sienta las bases de dicha materia dentro de la modelación estocástica.

A partir de lo anterior, Mejía y Rousselle (1976) publican un breve artículo, donde se propone una modificación de los modelos autorregresivos, porque a pesar de que el modelo básico reproduce las correlaciones entre los valores de los subperiodos de un mismo año, no mantiene las correlaciones entre el último subperiodo de un año y el primero del siguiente. Por ello, proponen la adición de un nuevo término al modelo básico, de tal forma que las correlaciones en cuestión puedan preservarse directamente. A la vez, Tao y Delleur (1976) presentan un modelo multivariado que permite desagregar series hidrológicas de alto nivel (anuales) en series de bajo nivel (mensuales o semanales), sin necesidad de transformar las series originales para remover sus periodicidades

intra-anales. Esto hace que el modelo esté libre de cualquier error o sesgo en la preservación de dichas periodicidades.

Debido al gran número de parámetros que puede llegar a tener en ciertas aplicaciones el modelo extendido de Mejía y Rousselle, Lane (1979) presenta un modelo de desagregación de series anuales en series estacionales, que comparativamente con el modelo extendido, incluso con el modelo básico, tiene una reducida cantidad de parámetros. Este modelo, conocido también como modelo condensado de desagregación espacial, permite que las subestaciones dependan de una estación clave.

Salas et al. (1980) publican una de las obras más completas en este aspecto, en la que se incluyen los modelos autorregresivos (AR), autorregresivos de media móvil (MA) y modelos de desagregación, al tiempo que presenta metodologías para enfrentar problemas prácticos.

Para poder comprobar si el modelo propuesto para un determinado estudio es adecuado o no, deben seguirse los pasos propuestos por Stedinger y Taylor (1982) quienes introducen los conceptos de *verificación* y *validación*.

Otros autores que cabe destacar en esta revisión son Bras y Rodríguez-Iturbe (1985) quienes publican un libro en el que dedican varios capítulos a la generación y predicción de series temporales con modelos autorregresivos, modelos de media móvil, modelos autorregresivos de media móvil y variantes o derivaciones de éstos, y modelos de desagregación. Al poco tiempo Fernández y Salas (1986) presentan un nuevo tipo de modelos basados en procesos autorregresivos para la generación de series sintéticas. Se trata de modelos capaces de reproducir la estructura de covarianza de series periódicas con distribuciones marginales no-gaussianas.

También se hace necesario citar a Ula (1992), autor que estudia los efectos que durante la fase de generación sintética tienen los valores iniciales sobre los momentos de primer y segundo orden en los modelos autorregresivos periódicos. Anota que aunque esos efectos son transitorios en los modelos estacionarios, su naturaleza y extensión no es clara en los modelos periódicos, y añade que por esta razón, el procedimiento común que consiste en eliminar una primera parte de las series sintéticas para remover el efecto de esos valores iniciales, puede ser inadecuado cuando se trata de procesos no-estacionarios.

Lettenmaier (1993) presenta una interesante serie de ideas relacionadas con el estado del arte de los modelos de generación sintética hasta ese momento. En sus análisis, el autor indica los problemas que a su juicio pueden resolverse satisfactoriamente con los modelos estocásticos, y señala los inconvenientes que existen en la aplicación de los mismos a cierto tipo de problemas; también da algunas luces sobre lo que pueden ser las futuras líneas de avance en

modelación sintética. Destaca, por ejemplo, que para la generación de caudales estacionales, el enfoque más práctico parece ser el de los modelos de desagregación, y cita específicamente el modelo condensado de Lane, aunque aclara que su uso depende de las correlaciones que se desee mantener; igualmente, advierte sobre la poca o nula utilidad de los modelos de generación sintética en aquellos casos en que los caudales no se pueden normalizar, como sucede en algunas regiones de clima árido o allí donde existen corrientes intermitentes o efímeras. Destaca también este autor que gran parte del valor de los modelos de generación sintética depende de la longitud de la muestra histórica, y finalmente presenta algunas reflexiones sobre la predicción de series temporales, en las que cuestiona la eficiencia de estos modelos para escalas temporales grandes.

Firor et al. (1996) utilizan modelos estocásticos para generar series sintéticas y con éstas alimentan un modelo simple de planificación hidrológica, con el fin de ver los efectos potenciales de los cambios de clima sobre los sistemas de recursos hídricos.

Koutsoyiannis (2000) presenta un procedimiento general que permite obtener series hidrológicas sintéticas, conservando su estructura de dependencia temporal a corto y largo plazo, y que puede adaptarse para propósitos de predicción. El procedimiento combina modelos autorregresivos de media móvil para reproducir la dependencia temporal a corto plazo, y modelos de ruido gaussiano fraccional para preservar la dependencia temporal a largo plazo. Esto se hace mediante el ensamble de lo que el autor denomina una función de autocovarianza generalizada, la cual está dada en términos de tres parámetros, de los que a su vez dependen los coeficientes del esquema de generación; de éstos se requieren tantos como términos de autocovarianza se desea reproducir.

Sánchez et al. (2001) integra el uso de modelos estocásticos en el marco de una nueva metodología desarrollada por los autores para la gestión probabilística de sistemas de recursos hídricos. Inicialmente hacen un análisis estadístico de las series de caudales históricos, para ajustarles un modelo estocástico y generar series sintéticas condicionadas al valor presente y con una longitud igual al período de anticipación, con el fin de alimentar con ello un modelo que simula la gestión del sistema de recursos hídricos en ese período y determina las probabilidades de estado del sistema.

Anderson et al. (2007) emplean los modelos autorregresivos periódicos de media móvil (PARMA) para la generación de caudales sintéticos semanales y diarios. Como los modelos estocásticos de bajo nivel presentan un gran número de datos, en este estudio emplean el análisis de Fourier para obtener una adecuada modelación del sistema. Al tratar series semanales, se obtienen 26 armónicos, hecho que dificulta a priori la selección de los armónicos más

importante para la posterior modelación de las series sintéticas. Para ello, emplean el estadístico α -Bonferroni para detectar los armónicos más influyentes.

Actualmente, las tendencias en modelación buscan mayormente perfeccionar las técnicas de calibración y los modelos existentes: se comparan unos modelos con otros, se intenta mejorar las técnicas de estimación de sus parámetros, se pone a prueba su eficiencia en estudios de sistemas de recursos hídricos, se combinan entre sí para crear nuevos esquemas de modelación sintética y, si se añade la informática, se puede implementar nuevos algoritmos y nuevas herramientas para la modelación de series sintéticas, tal y como se pretende mostrar en este trabajo. Concluyen indicando que la combinación Fourier-PARMA presenta mejor rendimiento que los modelos PARMA.

2.2 Sequías en los sistemas de recursos hídricos.

Las sequías son uno de los riesgos más complejos en los recursos hídricos (Wilhite, 1993). Los impactos suelen presentar mucha incertidumbre en cuanto a duración, extensión y magnitud. De ahí la importancia de poder analizar y prever de forma anticipada estos episodios de forma que podamos tratar de actuar de forma correcta y tratar de mitigar los efectos negativos que ocasiona.

Si bien es cierto que la sequía se debe a una reducción considerable de los recursos hídricos en un periodo dado, algunos autores señalan la importancia de definir el termino sequía debido a la diversidad de disciplinas afectadas por este fenómenos y a la variabilidad geográfica de los factores que los determinan tales como clima y economía. Además, al aumentar las demandas de agua durante el último siglo también han aparecido otros factores negativos como la contaminación de las aguas y el cambio climático (Bates et al. 2008).

Si se atiende al concepto de sequía aportado por Linsley et al. (1975), se puede decir que este término hace referencia a un periodo conocido en el que los caudales del sistema hídrico son insuficientes para satisfacer todas las demandas existentes. Con el paso del tiempo, los gestores de recursos hídricos han ido encaminados a reducir los daños y las pérdidas económicas; estas medidas podrían clasificarse como medidas preventivas y medidas reactivas. Las primeras tienen como principal objetivo tratar de anticiparse a las sequías y gestionar el riesgo con el menor impacto posible, mientras que las segundas actúan cuando ya se ha manifestado el fenómeno y se realizan actuaciones tendentes a superar el periodo desfavorable.

En este sentido, todas las acciones tendentes a la anticipación de los episodios de sequía van a proporcionar mejores resultados, ya que están orientadas a la planificación, gestión y asignación de los recursos disponibles y de las demandas existentes, además de mejorar la eficiencia al reducir las pérdidas. La mayoría de la bibliografía existente encamina muchas de estas medidas a la programación del suministro, aportaciones desde otros sistemas, reutilización y reciclaje de aguas, desalinización de aguas salobres, explotación de recursos subterráneos, entre otros (Bruins, 1993; Menéndez y Aguilera, 1996)

Al ser la planificación y la gestión una herramienta fundamental a la hora de abordar los fenómenos de la sequía, deben actuar proporcionando información y previsión de los eventos de sequía que pueden ocurrir. Es por ello, que la generación de series sintéticas es útil en este sentido, ya que los modelos de predicción suelen discrepar en función de los variables empleadas (Werick y Whipple, 1994) más aún en el ámbito mediterráneo, donde la variabilidad estacional de precipitación es muy acusada. Este planteamiento se recoge en el estudio de Menéndez y Aguilera (1996) donde se destaca la dificultad de aplicar modelos climatológicos de circulación general para la predicción de sequías ya que existe gran diferencia en la escala de trabajo de estos modelos con los modelos hidrológicos y la extrema variabilidad climatológica existente en España.

2.2.1 Variables necesarias para la predicción de sequías.

Las variables empleadas para la detección de las sequías dependen del ámbito a estudiar, es decir, en función del estudio que se quiera analizar se utilizarán unos parámetros u otros. (Mishra y Singh, 2010).

Por lo general, las variables hidroclimáticas comunes en cualquier estudio de sequías incluyen:

- Precipitaciones.
- Caudales aforados.
- Temperatura.
- Evapotranspiración
- Niveles de aguas subterráneas.
- Humedad del suelo.
- Niveles de lagos/embalses.

En relación con lo anterior, los parámetros empleados para determinar la gravedad de las sequías son los siguientes:

- Intensidad.
- Duración.
- Magnitud.
- Frecuencia.
- Extensión.

Una vez definidas las principales variables empleadas para el estudio de las sequías, conviene hacer un breve repaso de algunos métodos empleados en la predicción de sequías, analizando tanto los métodos tradicionales como los más avanzados.

2.2.1.1 Modelos de regresión.

Las relaciones existentes entre dos o más variables cuantitativas, en la cual una de ellas, la variable dependiente, que se pretende predecir y, la variable independiente, de la que se dispone información pueden ser relacionadas mediante este tipo de análisis.

La variable dependiente es el parámetro que cuantifica las sequía (por ejemplo: índice de sequía), mientras que la variable independiente es aquella que explica el porqué de las sequías (por ejemplo: precipitación, caudal, humedad del suelo).

En este sentido, existen algunos estudios centrados en el ámbito de la producción agrícola, puesto que, la situación de sequía es un periodo de estrés hídrico muy importante para las plantas, y por ello es posible encontrar estudios en los cuales se trata la predicción de las sequía en la fachada Este de Brasil utilizando el índice normalizado de variación de la vegetación en función como consecuencia los fenómenos de El niño (ENSO en Inglés) por los investigadores Liu y Negron-Juarez (2001).

A pesar de que la regresión lineal es un método muy utilizado, presenta algunas limitaciones que cabe destacar:

- La aceptación como relación lineal entre la variable dependiente e independiente, que por el contrario, dificulta la predicción a largo plazo, puesto que el comportamiento no es completamente lineal.
- Otro aspecto es la dificultad en la comprensión de los mecanismos subyacentes causales.

2.2.1.2 Análisis de series temporales.

Este tipo de análisis se encuentra muy extendido en muchos ámbitos científicos, sin embargo, para la predicción de sequías, su uso ha sido más bien escaso. Los principales aportes a este campo son la capacidad de identificación, estimación y comprobación de los modelos desarrollados (Mishra y Desai, 2005).

Los modelos autorregresivos integrados de media móvil (ARIMA) contribuyen con la aportación de considerar la correlación lineal entre observaciones, mientras que, los modelos autorregresivos integrados de media móvil estacionales (SARIMA) describen el comportamiento no estacionario dentro de la misma estación, así como entre las diferentes estaciones consideradas. (Box et al., 1994).

La inclusión de las series temporales en el estudio de las sequías ayuda a pronosticar las sequías utilizando las series históricas disponibles, si bien es cierto, que en la mayoría de los casos se dispone de series relativamente cortas para poder predecir con éxito las sequías.

Fernández et al. (2009) emplean un modelo SARIMA para predecir a 12 meses vista y obtener resultados para los índices de sequía empleando caudales mensuales y semanales. Sin embargo, han necesitado realizar un estudio específico para cada uno de los caudales utilizados, hecho que dificulta enormemente el modelo de predicción.

2.2.1.3 Modelos probabilísticos.

Los modelos probabilísticos son herramientas muy útiles para la predicción de sequías ya que son apropiados para cuantificar la incertidumbre asociada a este fenómeno. En la figura 2.1 se muestra la caracterización probabilística de los parámetros de sequía.

Las cadenas de Markov son un proceso estocástico que tiene la propiedad de que el valor procesado en un tiempo t (X_t), depende únicamente de los valores existentes en $t-1$ (X_{t-1}); y no de las secuencias de los valores X_{t-2} ; X_{t-3} ;...; X_{t-n} . La probabilidad condicional que aporta este modelo es que para un instante de tiempo el estado del sistema es diferente, fenómeno que es conocido como probabilidad de transición paso a paso (one-step transition probability) y para la cadena de Markov, la probabilidad de transición es un elemento básico.

Gabriel y Neumann (1962) fueron los primeros en emplear este método de la cadena para periodos secos. Lohani y Lognathan (1997) utilizando el Índice de severidad de sequía de Palmer (PDSI) caracterizan el comportamiento

estocástico de las sequías para los sistemas de alerta temprana. Paulo et al. (2005) han desarrollado un esquema condicionado empleando el método de Markov para predicción a corto plazo. Morerira et al. (2006) han aplicado un modelo lineal logarítmico para conocer las probabilidades de transición entre las clases de sequía para tratar de establecer un posible comportamiento de las mismas y ver si estas pueden tener relación con el cambio climático. Aún no se han encontrado resultados concluyentes.

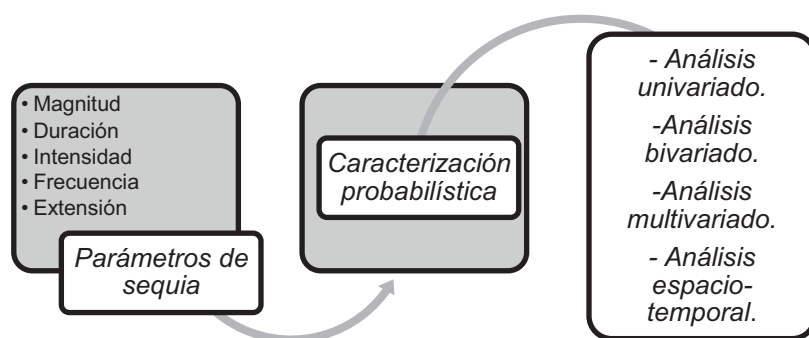


Figura 2.1: Caracterización probabilística de los parámetros de sequía.

2.2.1.4 Redes neuronales.

Las redes neuronales son un tipo de modelos no-lineales que aportan gran flexibilidad a la hora de afrontar distintos problemas. En este apartado se comenta la aplicación de las redes neuronales para la predicción de sequías, si bien, este es un campo en el que existen pocos estudios aplicando este tipo de modelo.

Morid et al. (2007) predice valores de índices de sequía empleando diferentes episodios de precipitación. Mishra y Desai (2006) en otro estudio, comparan los modelos estocásticos lineales y diferentes configuraciones de redes neuronales, mostrando que las redes neuronales presentan mejores resultados, si bien es cierto, que la naturaleza de estas redes se asemeja a una caja negra hecho que dificulta establecer un procedimiento bien definido.

2.2.1.5 Modelos híbridos.

Los modelos híbridos están comenzando a implementarse recientemente debido a la combinación de dos o más modelos permitiendo tener una mejor predicción que si se emplean los modelos por separado. Por ejemplo, Kim y Valdes (2003) muestran en su estudio la combinación de los modelos de wavelet y las redes neuronales artificiales; los primeros permiten capturar la información necesaria en todos los niveles, mientras que las redes neuronales captan la información a diferentes niveles gracias a las diferentes capas que forman su estructura. Otro ejemplo es el de Bacanli et al. (2009) para la región de Anatolia (Turquía) donde comparan el modelo de interferencia neurodifuso (ANFIS) frente a las redes neuronales y en el mismo se observan ventajas a la hora de emplear los modelos ANFIS para predecir las sequías utilizando el índice de precipitación estandarizada (SPI).

2.2.2 Predicción a largo plazo de las sequías.

Para poder predecir las posibles sequías que puedan ocurrir en el futuro se debe atender a diferentes variables:

- El papel de la información disponible: Durante mucho tiempo existen registros de precipitación, pero estos registros no se encuentran homogéneamente distribuidos por toda la superficie terrestre. Sin embargo, se conoce que existe una relación inequívoca entre precipitación y aparición de sequías. A escala global, se conoce muy bien el fenómeno atmosférico, pero en cada región existen particularidades orográficas que hacen variar los resultados.
- Los modelos de circulación global (GCMs) pueden servir para establecer patrones o comportamientos de sequías con bastante fiabilidad. Sin embargo, la particularidad de cada región hace que todavía se tenga gran incertidumbre en la predicción de las sequías.
- La importancia de la periodicidad en la predicción de sequías: A menudo, las variables hidroclimáticas presentan periodicidades relacionadas con las sequías, si bien es cierto, que estas pueden ocurrir cada 2 o 3 años o incluso, que no aparezcan en décadas, hecho que demuestra que aún no se comprende bien este fenómeno. Por poner un ejemplo, Siscoe (1978) pudo establecer una generalización con un periodo de 20 años donde podía apreciarse la expansión y contracción de las áreas afectadas por las sequías en el Oeste americano, a partir del ciclo solar de Hale (presenta una periodicidad de 22 años).

- La importancia de la extracción de datos adecuados para predecir sequías: Por lo general, dependiendo de la escala espacial a estudiar, se emplea un tipo u otro de información. Pero, la verdadera labor en la predicción de sequías, está en determinar qué datos son los útiles, hecho que dificulta aún más la capacidad de predecir las sequias. En este aspecto, cabe destacar el estudio de Tadesse et al. (2004) que ha conseguido identificar la relación entre los parámetros oceánicos y los índices de sequía en la región de Nebraska, EE.UU. En este estudio ha sido capaz de diferenciar las sequias separadas por episodios secos y húmedos.

2.2.3 Modelación de sequías bajo escenarios de cambio climático.

La temperatura global del planeta se ha visto incrementada durante el último siglo y todo apunta que lo seguirá haciendo durante el presente siglo. Los efectos del cambio climático son muy diversos y varían en función de la zona de estudio. Los modelos de circulación global (GCMs) pueden proporcionar información para la predicción de sequías en función de distintas hipótesis de trabajo.

El ejemplo de la aplicación de los GCMs para la predicción de sequías puede verse en el estudio de Blenkinsop y Fowler (2007), en el cual se obtuvo una modificación del Índice de sequía basado en las anomalías de las precipitaciones mensuales para seis regiones de Europa. Se extrajeron algunas observaciones interesantes como fueron, por un lado, el incremento de la duración de los episodios de sequía en la zona sur de Europa y por otro lado, menos sequías en la zona norte de Inglaterra.

En general, todos los resultados de los GCMs muestran una tendencia a tener veranos más cálidos y secos e inviernos más fríos y húmedos que los que se han registrado en el siglo pasado, además de apuntar que el hemisferio norte tendrá un balance global de precipitación menos favorable que en el siglo XX.

En este sentido, la predicción de sequías en el futuro a largo plazo se va a convertir en un área de estudio muy importante para poder anticiparse a estos fenómenos y gestionar los recursos hídricos de forma óptima.

Por tanto, se puede afirmar que las sequías constituyen un tema esencial en contexto de los sistemas de recursos hídricos y repercuten de forma considerable en la planificación y gestión de los recursos hídricos, por lo que se hace necesario un enfoque probabilístico para poder abordar el problema con suficiente antelación y tener definido un protocolo de actuación.

2.2.4 Gestión de los episodios de sequía mediante sistemas de ayuda de decisión (SAD).

Para una gestión efectiva de los periodos de sequía, se deben tener en consideración los siguientes factores:

- Impacto natural debido a la variabilidad hidro-meteorológica del sistema.
- La información obtenida a partir de modelar diferentes componentes cualitativos de la sequía.

Sin embargo, comprender la variabilidad natural de los procesos es algo complejo tal y como se ha visto en los apartados anteriores. Prueba de ello es que una alta variabilidad de los factores hidro-meteorológicos incrementa de forma considerable la incertidumbre para la gestión en las etapas iniciales de la alerta por sequía. En la figura 2.2 aparece descrito de forma simplificada el funcionamiento de un sistema de ayuda de decisión.

Tal y como se vio en el apartado 2.1.3, los SAD proporcionan una gran ayuda a los expertos a la hora de actuar, tanto en condiciones normales del sistema, como en condiciones de sequía. A la hora de trabajar en una situación de sequía, es importante atender a las salidas del sistema, tales como:

- La capacidad del embalse en función de los diferentes escenarios estudiados.
- La caracterización multivariada del sistema al modificar diferentes parámetros como, capacidad de reserva, déficits, duración de la sequía, etc.
- La respuesta del sistema en función de la duración, magnitud, resiliencia del fenómeno de la sequía.

Como ejemplos de los SAD aplicados a la gestión de sequías se pueden indicar los estudios de Chang et al. (1996) donde se desarrolla una monitorización diaria de la sequía empleando la memoria histórica de la que se dispone. Por otro lado, Pallotino et al. (2005) ha desarrollado un sistema SAD que permite generar un escenario en el cual pueden estudiarse un conjunto de estadísticos de sequía basados en la evolución temporal y espacial de la sequía para tratar de identificar su posible tendencia y/o evolución, de modo que se pueda disponer de una base más sólida para tomar decisiones.

Por último, habría que destacar los mecanismos de toma de decisiones multicriterio (MCDAs), ya que las sequías al ser un fenómeno multidimensional, en el que se incluyen aspectos meteorológicos, ecológicos, hidrológicos, medio ambientales y socio-económicos, hace que cualquier decisión que se tome a la hora de gestionar un episodio de sequía, pueda desencadenar efectos negativos en otros aspectos, que a priori, parecían que no fuesen a estar

afectados. Sin embargo, el mayor inconveniente de estos sistemas de decisión está a la hora de entender cómo se va a comportar la situación de riesgo si se hace persistente en el tiempo, hecho que en la mayoría de ocasiones, termina por hacer que los expertos no tomen las decisiones más adecuadas. (Whestphal et al., 2007).

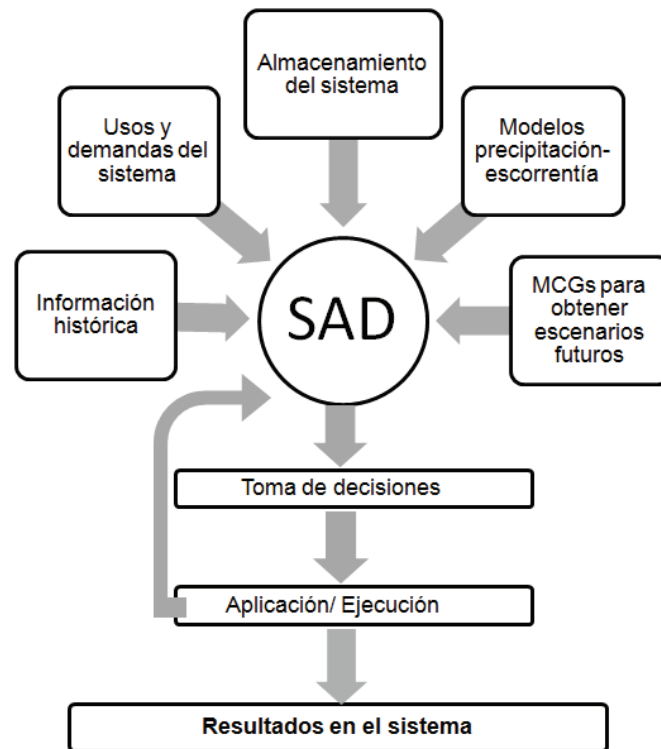


Figura 2.2: Esquema simplificado de funcionamiento de un sistema de ayuda de decisión (SAD).

2.3 Las redes neuronales artificiales.

Tal y como se ha remarcado anteriormente, gracias al desarrollo de la informática se han podido incluir las redes neuronales en la gestión de los recursos hídricos. Las redes neuronales, gracias a su estructura, algoritmos utilizados y robustez de sus resultados, han permitido asentarse dentro de este ámbito.

La explicación reside en el uso de la computación cognitiva, la cual comprende varias teorías que tienden a imitar el comportamiento de los sistemas biológicos; una de esas teorías es justamente la de las redes neuronales.

Las redes neuronales ofrecen numerosas ventajas. La primera es su capacidad de aprendizaje ya que permiten generalizar sucesos pasados en situaciones futuras, además tolera ciertos errores o ausencias de datos ya que su estructura en paralelo permite redirigir la información y evitar la pérdida de funcionalidad (aunque sea a costa de anular parte de la red). Otra de las virtudes que presenta es su capacidad de procesar grandes cantidades de datos de forma rápida sin saturarse. Por último, destacar que pueden implementarse fácilmente gracias a las nuevas herramientas informáticas.

2.3.1 Desarrollo de las redes neuronales.

Las redes neuronales surgen en la década de los cuarenta del siglo pasado, cuando Pitts y Mc Culloch (1943) proponen una conceptualización del funcionamiento del cerebro humano basada en una red de células interconectadas. El objetivo último consistía en encontrar una explicación de cómo funcionaba realmente el cerebro. El modelo propuesto fue un dispositivo lógico (matemático) y no fisiológico y/o biológico. Con estas premisas el mecanismo parece sencillo, la neurona recibe una señal binaria que produce respuesta o no. Por tanto, las redes neuronales quedan dispuestas por unidades de proceso elementales (nodos) interconectadas entre si y organizadas de forma jerárquica, con capacidad para adquirir conocimiento de tipo experimental. En concreto, las redes neuronales artificiales tienen la capacidad de simular y predecir procesos no lineales.

Años más tarde, Minsky y Papert (1969) demuestran matemáticamente las limitaciones del perceptron simple, demostrando que sólo puede resolver problemas sencillos de clasificación de patrones, en los que las regiones de decisión estén separadas por una línea, o un hiperplano en el caso más general.

Von der Malsburg (1973) puede considerarse como uno de los fundadores de las redes neuronales gracias al aporte realizado sobre los trabajos de simulación computacional en mapas auto-organizativos basados en el aprendizaje competitivo. Por otro lado, Kohonen (1982) presenta un modelo de red neuronal capaz de formar mapas topológicos de características, a partir de una cierta información de entrada, de manera similar a como ocurre en el cerebro humano. Ackley et al. (1985) utilizan las aportaciones presentadas por Kirpatrick et al. (1983, Procedimiento del temple simulado) para crear la máquina de Boltzman (la primera red multicapa que resulta exitosa).

Posteriormente, Rumelhart et al. (1986) reportan el desarrollo del algoritmo de retropropagación del error (error-backpropagation algorithm), que desde entonces se convierte en la técnica más popular para el entrenamiento de redes perceptron multicapa, dando así gran impulso al desarrollo y aplicación de nuevas investigaciones en el campo de las redes neuronales artificiales.

En la actualidad se avanza en la investigación de nuevos modelos de redes neuronales, en nuevas aplicaciones de modelos existentes, en el perfeccionamiento de los algoritmos de entrenamiento y en la búsqueda de metodologías que permitan determinar más directamente las tipologías de varios modelos de uso frecuente.

2.3.2 Aplicación de las redes neuronales en la gestión de los recursos hídricos.

Las redes neuronales llevan aproximadamente dos décadas dentro de la ciencia del agua, esto se debe, según algunos autores (Govindaraju y Rao, 2000) a la tendencia que existe entre los técnicos a aplicar sólo metodologías de uso consolidado. Otro de los motivos es el hecho de que las redes neuronales son, desde el punto de vista de la modelación matemática cajas negras, es decir, no tienen en cuenta los principios físicos del sistema en su modelación.

Las redes neuronales gracias a la potencialidad de sus aplicaciones han permitido desarrollar nuevas líneas de investigación en el campo de la Hidrología Operacional (Dowla y Rogers, 1995).

Algunos estudios indican que ciertos resultados mostrados por algunos modelos de redes neuronales son similares a los obtenidos por otros modelos tradicionales, por lo que en realidad según algunos autores, no son una nueva aplicación, sino otra forma de interpretar lo que ya se conocía anteriormente (Bienenstock y Geman, 1994). Sin embargo, los modelos estadísticos pueden incluirse dentro de las redes neuronales de forma que puedan realizar las

funciones como la regresión lineal simple, regresión polinómica, regresión logística, etc....

Sarle (1994) analiza cómo de forma individual los modelos estadísticos se pueden obtener simplemente con la variación de la función de transferencia y el número de nodos ocultos en una red neuronal. Estas relaciones se recogen en la tabla 2.1. También se ha sugerido que ciertos modelos de redes neuronales son equivalentes a modelos tipo ARMA (Hill et al., 1994), algunos de estos ejemplos aparecen en la tabla 2.2. Por otro lado, Connor et al. (1994) demostraron que las redes neuronales alimentadas hacia delante (feedforward) son un caso especial de las no lineales autorregresivas y que las redes neuronales recurrentes se asemejan a las no lineales autorregresivas con media móvil. Chon y Cohen (1997) han demostrado los modelos ARMA y no lineales autorregresivas con media móvil como redes equivalentes cuando se emplean funciones polinómicas en su configuración.

Tabla 2.1: Equivalencia de los modelos estadísticos y funciones de activación empleadas en las redes neuronales artificiales.

Fuente: Neural networks and statistical models. 1994.

Función de activación.	Modelo estadístico equivalente.
Umbral/escalón.	Función discriminación lineal.
Lineal.	Regresión lineal.
Logística.	Regresión logística.

Tabla 2.2: Equivalencia de los modelos estadísticos y número de nodos ocultos empleadas en las redes neuronales artificiales.

Fuente: Neural networks and statistical models. 1994.

Número nodos ocultos	Tipo modelo	Modelo estadístico equivalente.
Bajo	Paramétrico	Regresión polinómica.
Moderado	Quasi-paramétrico	Proyección pursuit.
Incrementado con el tamaño de la muestra	cribado no paramétrico	Regresión Kernel

Otros autores como Govindaraju y Rao (2000) presentaron una recopilación de artículos en los que se reportan aplicaciones de varios tipos de redes neuronales artificiales aplicadas a diferentes ámbitos de la hidrología, como es caso de predicción de caudales con redes multicapa alimentadas hacia delante; modelación de la escorrentía de una cuenca por medio de redes neuronales; caracterización de datos espaciales de la subsuperficie del suelo, empleando distintos tipos de redes neuronales, entre ellas las de Kohonen; modelación de

la recuperación de acuíferos mediante redes perceptron multicapa entrenadas con el algoritmo de retropropagación; regulación de ríos aplicando redes multicapa y redes de función de base radial; entre otros.

En la actualidad, las redes neuronales proporcionan una capacidad de generalización mejor que las metodologías clásicas. Además, han permitido comprender las relaciones existentes entre las diferentes subsistemas que conforman una cuenca hidrográfica. Por esta razón, siguen realizándose investigaciones y aplicaciones en casi todos los campos de la Hidrología, gracias a la robustez y bondad de los resultados conseguidos. La modelación de series temporales de caudales no ha sido ajena a esta realidad, aunque dicha área en particular no ha sido tan explorada como otras. A continuación se muestran las áreas más desarrolladas:

- Predicción de caudales y aportaciones. (Escala mensual, semanal, diaria, horaria).
- Generación de series sintéticas de aportaciones.
- Completado de datos en series hidrológicas.
- Predicción de lluvias.
- Estimación de precipitación a partir de imágenes de satélite.
- Aplicación de los Sistemas de Información Geográfica.
- Clasificación de patrones sinópticos y análisis de relaciones entre variables meteorológicas e hidrológicas.
- Aplicaciones en tiempo real: Niveles y caudales en situación de crecida.
- Modelación lluvia-escorrentía.
- Gestión de embalses.
 - o Contaminación y niveles de acuíferos.
 - o Precipitación.
 - o Aportaciones.
 - o Regulación de ríos.
 - o Transporte de sedimentos,
 - o Calidad de las aguas (contaminación y evolución).

2.3.3 Modelación de series temporales.

Las técnicas clásicas empleadas en la modelación de series de caudales asumen relaciones lineales entre las distintas variables, hecho que contrasta con la realidad, produciendo en algunas ocasiones resultados imprecisos. En el caso de emplear modelos no-lineales se hace necesario emplear un gran número de variables, las cuales exigen gran cantidad de datos externos que, en la mayoría de casos, es difícil de obtener (Deo y Thirumalaiah, 2000). Frente a esta situación, las redes neuronales artificiales gracias a su capacidad

para trabajar en procesos no-lineales (como es el caso de muchos problemas hidrológicos) proporcionan resultados más precisos que los modelos lineales. En virtud de lo anteriormente citado, puede destacarse el trabajo de Lapedes y Farber (1988) quienes trabajan con una serie temporal caótica para ilustrar una aplicación de redes neuronales alimentadas hacia adelante y entrenadas mediante el algoritmo de retropropagación del error. En este estudio advierten sobre la lentitud de convergencia de los métodos de optimización basados en el descenso del gradiente. Además, demuestran que las redes neuronales son una técnica precisa en tareas de predicción. Destacan la mejor calidad de los resultados que se obtienen al hacer predicción con pocos intervalos de tiempo hacia adelante, en lugar de emplear muchos. Un aspecto interesante de este trabajo es la interpretación geométrica que ofrecen de una red neuronal alimentada hacia adelante con dos capas ocultas.

El principal aporte de este estudio es la modificación que se hace al algoritmo clásico de retropropagación, con el fin de obtener la mejor arquitectura durante su proceso de aprendizaje. Dicha modificación consiste en sumarle a la función objetivo clásica (sumatoria de los errores cuadrados) un término denominado complejidad, con el cual se castiga el exceso de pesos, eliminando durante el entrenamiento los que no sean necesarios.

Comparan los resultados de la red con los de un modelo autorregresivo de umbral, siendo superiores a los de éste; así mismo, analizan la sensibilidad del error de la red neuronal ante el cambio del número de nodos de la capa de entrada: después de un cierto límite, el error permanece constante. Analizan también el preprocesamiento de los datos, y concluyen que este procedimiento permite que la red aprenda más rápido, pero advierten que de esta forma los errores en la escala original son mayores que cuando se prescinde del preprocesamiento.

Posteriormente, aparece el primer estudio de Raman y Sunilkumar (1995) para la generación de series de caudales mensuales en dos estaciones constituido por doce redes neuronales perceptron multicapa, las cuales entrenan con el algoritmo de retropropagación del error. Antes de presentar los datos a la red, realizan un preprocesamiento de los mismos, el cual consta de tres fases:

- 1) transformación normalizadora.
- 2) tipificación de los valores normalizados.
- 3) reducción de los sesgos mensuales de las series tipificadas.

La cantidad de neuronas de la capa oculta se establece mediante un proceso de ensayo y error. Los autores entrenan cada una de las doce redes del modelo de forma independiente, empleando para ello la serie de caudales del mes al que corresponde cada red; luego realizan su validación cruzada con dos años de datos, y después la prueban con otros dos años más. Generan diez series mensuales de 14 años de longitud y comparan sus estadísticos básicos

mensuales (media, desviación típica y coeficiente de sesgo) con los de las series generadas mediante un modelo AR(2) bivariado; también comparan las series generadas con las históricas, así como sus correlogramas. Los resultados obtenidos llevan a los autores a concluir que las redes neuronales artificiales ofrecen una alternativa viable para la modelación multivariada de series temporales en estudios de análisis de sistemas de recursos hídricos.

Por otro lado, Stern (1996) genera una red perceptron multicapa alimentada hacia delante y con el algoritmo de retropropagación del error para generar series temporales determinísticas y estocásticas. Ello, con el fin de analizar la capacidad de las redes neuronales para la predicción de series que tengan una componente aleatoria. Utiliza redes perceptron de tres capas, ya que estas son las más empleadas en problemas reales. Interpreta las neuronas de las capas ocultas como una forma de procesamiento intermedio de los datos, cuya función de activación suele ser la sigmoideal. Destacar de este estudio que la series empleadas parten de modelos previamente definidos (no son ni empíricas ni observadas) y sus resultados muestran que la red neuronal por si misma no es capaz de generar ruidos aleatorios.

Imrie et al. (2000) destacan la idoneidad de las redes neuronales en problemas de predicción hidrológica, aunque indican que en muchos casos no logran reproducir de manera satisfactoria los valores extremos de las series de datos con los cuales se las valida. Como solución a este problema, proponen la aplicación del algoritmo de correlación en cascada para realizar el entrenamiento de la red, ya que tiene la propiedad de configurar la tipología óptima de la red dentro del algoritmo de entrenamiento. Con dicho algoritmo, los autores entrenan una red neuronal para predecir caudales horarios, tratando especialmente de reproducir los valores extremos. Comparan la red así entrenada con otra a la que le aplican el algoritmo de retropropagación del error, y obtienen mejores resultados con la primera. En el estudio de Imrie et al., la modelación es univariada y las series empleadas no presentan periodicidades. La bondad de los resultados, especialmente la capacidad que presenta la red para reproducir los valores extremos, se deriva en buena parte del empleo de polinomios cúbicos como funciones de activación, aunque los mismos autores advierten que su aplicación a otros casos no garantiza necesariamente el buen ajuste de esos eventos extremos.

Salas et al. (2000) presentan un estudio en el que modelan la predicción de caudales mensuales mediante redes neuronales perceptron multicapa, entrenadas con el algoritmo de retropropagación del error. El modelo que desarrollan es univariado y consiste en predecir los caudales de julio a partir de los caudales registrados en junio sobre un horizonte de diez años. Tipifican los datos originales antes de presentárselos a la red, cuya arquitectura es bastante simple, ya que está formada por tres capas, con una neurona en la primera capa, dos en la segunda y una en la tercera. Emplean como funciones de

activación, sigmoides bipolares. Validan el modelo con un conjunto de datos de igual longitud a los empleados en el entrenamiento, y juzgan el desempeño de la red entrenada comparando varios estadísticos básicos de la serie generada con los de la serie histórica. Los valores predichos por la red se aproximan a los históricos.

Cannas et al. (2000) construyen una red neuronal perceptron de tres capas, cuyo entrenamiento lo realizan con el algoritmo de retropropagación del error, y también implementan una red localmente recurrente para predecir caudales mensuales a uno, dos y tres meses de anticipación. Obtienen la arquitectura de cada red mediante un proceso de ensayo y error: para la red perceptron logran los mejores resultados con doce unidades en la capa de entrada, e igual número de neuronas en la capa oculta, y en la capa de salida tantas neuronas como intervalos a predecir; para la red recurrente obtienen una capa oculta con cinco neuronas. El entrenamiento se lleva a cabo con diez años de datos mensuales y para la validación utilizan datos de otros diez años.

Para juzgar la validez de la predicción de caudales utilizan un parámetro denominado índice de eficiencia del modelo, que es función de los errores medios cuadrados. Los autores reportan como satisfactorios los resultados de la modelación, especialmente los que se consiguen con la red neuronal localmente recurrente, aunque advierten que la calidad de los mismos se degrada cuando aumenta el número de intervalos de predicción. Este modelo es un caso de predicción univariada, en el que a pesar de que los resultados son buenos en general, los caudales pico tienden a ser subestimados y los mínimos sobreestimados, comportamiento que también detectan en su estudio Imrie et al. (2000). Cabe destacar, que la tipología de la red perceptron empleada para la predicción a un intervalo de tiempo tiene 156 conexiones, es decir, se trata de un modelo con un mínimo de 156 parámetros para calibrar.

En el año 2002 se publica la tesis doctoral *Modelo estocástico de redes neuronales para la síntesis de caudales aplicados a la gestión probabilística de sequías* por Ochoa-Rivera, con el objetivo de buscar una metodología capaz de generar con redes neuronales series sintéticas de mayor memoria hidrológica que las que se obtienen con los modelos autorregresivos clásicos. El autor propone un modelo perceptron multicapa con una componente aleatoria estocástica embebida, el cual se aplica a cuatro casos de estudio en los que los resultados se comparan con los de los modelos ARMA y modelos de desagregación temporal. Las acciones del trabajo consistieron en generar series sintéticas de caudales mensuales mediante el nuevo modelo y las técnicas tradicionales, para comparar su capacidad de preservación estadística, después se realizó la simulación, en términos probabilísticos, de la gestión de varios sistemas de recursos hídricos de distinta complejidad en condiciones de sequía. Los resultados finales muestran que el modelo propuesto es más robusto y presenta mayor memoria hidrológica que los

métodos tradicionales. Este estudio ha servido de base de partida para el presente trabajo fin de máster.

Continuando con el estudio anterior, Ochoa-Rivera et al., (2007) utilizan las redes neuronales para la modelación de los datos de entrada en un sistema de soporte de decisión enfocado a la gestión de sistemas de recursos hídricos. La finalidad era tratar de determinar hasta qué punto la componente aleatoria presente en los recursos hídricos en la cuenca del Júcar influye en la gestión de los mismos y, gracias a este estudio se puede indicar que es necesario aplicar modelos no lineales para los sistemas de soporte de decisión. El modelo utilizado de perceptron multicapa era el que mejor resultados proporcionaba, en comparación con los modelos ARMA(1,1) y el modelo de desagregación temporal de Lane. Posteriormente se publica un nuevo trabajo de Ochoa-Rivera (2008) donde se destaca la aptitud de las redes neuronales estocásticas con la modelación de eventos de sequía a partir de series de caudales mensuales.

Capítulo 3

Estadísticos de bondad de ajuste en la modelación hidrológica

El éxito del proceso de generación sintética de caudales depende en gran medida entre otras, del tipo de modelo que se emplee; los modelos presentan ciertas limitaciones bien identificadas en la literatura de modelación de series de caudales. Éstos son de gran relevancia cuando se simula el comportamiento de los sistemas de recursos hídricos. Es por ello que resulta necesario pensar en nuevos tipos de modelos que mejoren la reproducción de dichos estadísticos y en consecuencia determinar la validez de los diferentes modelos propuestos para los objetivos planteados.

Uno de los primeros objetivos a la hora de determinar si un modelo es adecuado o no es comparando los valores simulados con los valores observados.

Es muy frecuente utilizar un número de estadísticos y/o técnicas para determinar cuál es el modelo más adecuado. Normalmente en este tipo de herramientas se incluyen las medidas de bondad de ajuste o medidas del error relativo.

Sin embargo, conforme se desarrollan los modelos estos a su vez se vuelven más complejos, al igual que ocurre en el ámbito de la Hidrología. En un principio, los procesos de calibración estocástica han sido muy bien empleados si bien es cierto, que por esa misma razón se hace necesaria una revisión para determinar cuáles de los siguientes estadísticos pueden aplicarse como medida comparativa entre diferentes modelos y sus limitaciones.

3.1 Índices de bondad de ajuste de modelos. Aplicación a la Hidrología.

Para poder comprender la situación actual y el debate científico existente es conveniente realizar un pequeño paseo en el tiempo para conocer los orígenes de los índices más empleados y las diferentes opiniones existentes a la hora de establecer un marco estándar de índices que permitan comparar modelos.

El Coeficiente de eficiencia (CE) es presentado y desarrollado por Nash y Sutcliffe (1970). Este es un índice cuyo intervalo oscila entre $-\infty$ y 1. Siendo el valor 1 el ajuste perfecto, entendiendo por perfecto el caso en el que los valores simulados son idénticos a los observados. Cuando el valor es igual a 0 indica que las predicciones del modelo son tan precisas como la media de los datos observados, mientras que una eficacia menor que 0 se produce cuando la media observada es mejor que el modelo empleado, en otras palabras, cuando la varianza residual (descrita por el numerador de la expresión anterior), es mayor que la varianza de los datos (que se describe por el denominador). Este índice físicamente recoge la relación entre el error cuadrático medio (MSE por sus siglas en inglés) y la varianza de los datos observados y restado a 1.

Wilcox et al. (1990) analizaron la aplicación del CE. En su estudio se podía observar que si el cuadrado de las diferencias entre el modelo simulado es igual de grande que la variabilidad de los valores observados, el resultado del índice sería 0 y en el caso de que fuese este valor negativo, vendría a demostrar que la media de los observados sería mejor para predecir que los valores simulados.

El coeficiente de eficiencia representa una mejora respecto el coeficiente de determinación a la hora de evaluar los diferentes modelos ya que el coeficiente de determinación es muy sensible a las diferencias de las varianzas y de las medias de los valores observados y simulados, por el contrario, es demasiado sensible a los valores extremos (al igual que el coeficiente de determinación).

En referencia a lo anterior, el coeficiente de determinación (R^2) describe la proporción del total de la varianza que puede ser explicada por el modelo a partir de los datos observados. Para el coeficiente de determinación los valores oscilan entre 0 a 1. Sin embargo, el coeficiente de determinación es un índice limitado ya que estandariza las diferencias entre los valores observados y simulados de forma lineal entre las variables demostrando la insensibilidad del modelo en el caso que se añadan diferencias proporcionales entre los valores simulados y observados (Willmott, 1984).

En el ámbito de la Hidrología y la Climatología aparecen los primeros estudios de Mc Cuen y Sinder (1975) y Willmott (1984) comparando este índice con otros índices clásicos. Legates y Davis (1997) ilustran la que este coeficiente es muy sensible a valores extremos de los observados cerca de la media, por tanto, un modelo que trabaje en eventos de crecida va a presentar valores muy altos de r^2 ocultando la verdadera relación del resto de los valores observados y simulados. Mc Cuen y Sinder (1975) propusieron incluir un factor de ajuste en el denominador, si bien es cierto que comentan que en caso de eventos muy extremos el factor de correlación podría sobrepasar el valor 1.

Pese a que el índice de Nash-Sutcliffe es uno de los índices más utilizados para evaluar un modelo hidrológico, como se ha podido observar, se puede lograr un alto valor de este índice con un modelo no adecuado. Por tanto, no es aconsejable concluir el rendimiento de un modelo utilizando únicamente el CE. Otras herramientas estadísticas deben ser empleadas para llegar a una conclusión definitiva acerca del funcionamiento del modelo. Incluso un mal modelo puede dar valores en torno a 0,6 y en estos casos se hace necesario observar los datos simulados antes de sacar cualquier conclusión sobre la conveniencia o no del modelo.

En razón de lo anteriormente citado, numerosos autores muestran esta disconformidad, desde Martinec y Rango (1989) hasta Schaepli y Gupta (2007), permitiendo que aparezcan diferentes estudios acerca de la necesidad de modificar el índice CE para ajustarse de mejor forma a ámbitos específicos de la hidrología, así pues, Mathever et al (2006) propone una modificación muy limitada de este índice y poco después, Criss y Winston (2008), a raíz de Schaepli y Gupta (2007), definen un índice volumétrico para ser utilizado en lugar del CE en los modelos que simulan eventos de crecida.

De entre las limitaciones del CE está el hecho de emplear los valores medios observados como valor de referencia a partir del cual se evalúa el modelo. Este hecho presenta gran importancia cuando se evalúan variables como la precipitación en cuencas con gran variabilidad estacional, ya que puede producir grandes sobreestimaciones en el resultado del modelo. Por tanto, se hace necesario interpretar correctamente este índice, más aún cuando se intenta comparar diferentes estudios realizados en diferentes climas.

Por lo general, está aceptado emplear el CE y error cuadrático medio (MSE por sus siglas en Inglés) para determinar la capacidad o no del modelo de representar la realidad del sistema de porque pueden evaluarse por separado para diferentes periodos de observación; diferentes lugares y diferentes subsistemas dentro de la misma observación.

Debido a la gran popularidad de la que goza el CE, otros autores muestran la necesidad de descomponer este índice para poder comprender mejor sus resultados. Por esta razón Gupta et al. (2009) descomponen el índice CE de la siguiente manera:

$$CE = 2 \alpha r - \alpha^2 - \beta_n^2 \quad (3.1)$$

donde :

r= Coeficiente de correlación lineal.

$$\alpha = \sigma_s / \sigma_0$$

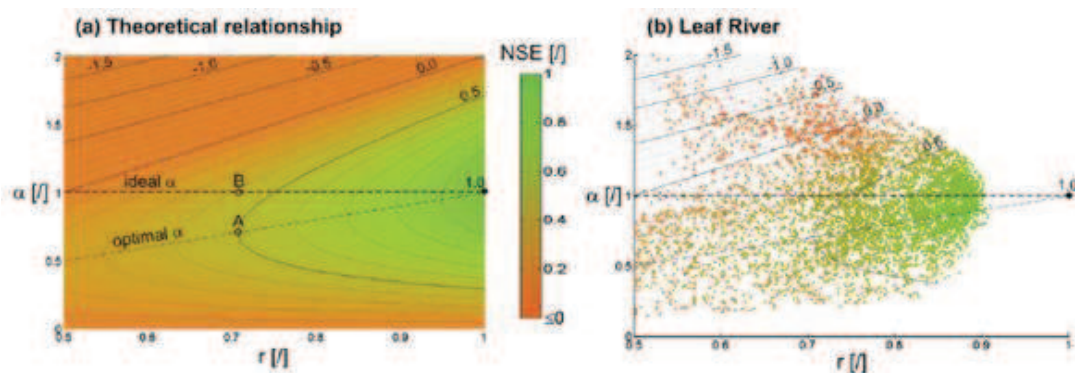
$$\beta_n = (\mu_s - \mu_0) / \sigma_0$$

Por tanto, el índice CE presenta dos componentes capaces de mostrar la relación existente entre la media y la desviación típica entre los valores simulados y observados, mientras que el tercer componente muestra la habilidad del modelo para reproducir en tiempo y forma los valores observados. En el caso del modelo perfecto, los valores serían $r=1$; $\alpha=1$; $\beta_n=0$.

Sin embargo, para utilizar el índice CE debe tenerse en consideración dos aspectos importantes:

- 1- El factor bias ($\mu_s - \mu_0$), componente que aparece en la forma normalizada de las series, estandarizado por la desviación típica de las series observadas. Esto debe tenerse muy en cuenta en áreas donde exista una gran variabilidad de precipitación-escorrentía ya que el componente bias tiende a reducirse y en consecuencia, el impacto en el índice CE es mínimo, hecho que puede ocasionar un buen resultado de CE cuando en realidad se están produciendo grandes errores en los volúmenes de agua simulados. De este modo, podría decirse que es equivalente a usar un modelo con diferentes pesos en el que el peso correspondiente al factor bias es despreciable.
- 2- Otro aspecto a destacar, es que el factor α aparece dos veces mostrando una interacción con el coeficiente de correlación lineal (r), que en el caso de obtener el valor de ajuste perfecto del índice CE, coincide cuando $\alpha=r$. Por tanto, r siempre va a ser menor que la unidad, lo que significa que para aumentar el valor de CE se va a infravalorar el valor del parámetro α en el caso de los caudales.

Si se considera a la vez los dos aspectos anteriormente citados se obtiene que cuando $\alpha=r$ y $\beta_n=0$ el índice CE es igual a r^2 (coeficiente de determinación) tal y como aparece en la figura 3.1.



- a) Relación teórica que muestra el valor ideal y óptimo de α . Para el punto A, el valor CE es mejor que para el punto B.
- b) Ejemplo de los resultados obtenido en un modelo hidrológico para Leaf River (Mississippi, EE.UU) 1924 Km² con una serie histórica de 11 años.

Figura 3.1: Relación de CE con α , γ (asumiendo $\beta_0 = 0$).

Fuente: Nash et al., 2009.

Por tanto, el índice propuesto por Gupta et al. (KGE) propone sustituir el valor β_n por β , de este modo, se puede obtener una mejor interpretación del bias utilizando la relación entre las medias de los caudales simulados y observados y de este modo, queda como valor óptimo $\beta=1$ (al igual que α y r).

$$KGE = 1 - ED \quad (3.2)$$

$$KGE_s = 1 - ED_s$$

donde:

$$ED = \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}$$

$$ED_s = \sqrt{[S_r(r - 1)^2] + [S_\alpha(\alpha - 1)^2] + [S_\beta(\beta - 1)^2]}$$

ED hace referencia a la distancia Euclídea desde el punto ideal, mientras que ED_s hace referencia a la distancia Euclídea al punto ideal en el espacio normalizado. S_r , S_α y S_β son factores para re-escalar el espacio desde el punto ideal.

De este modo, la corrección propuesta en el KGE obtiene una menor infraestimación de los resultados cuando se trabaja con caudales de eventos de precipitación, que cuando se trabaja con caudales de sequía, en el que la sobrestimación es menor que en el caso de emplear el índice CE.

La decisión de incluir el índice KGE para evaluar los modelos en el presente trabajo se debe a:

1. Éste índice permite determinar la distancia Euclídea de los tres componentes (α , β y r) respecto al valor ideal, que sería el equivalente al punto ideal de la regla de Pareto en un espacio tridimensional.
2. La infraestimación de los picos es menos severa que en el caso de CE.
3. Pone de manifiesto la importancia de los tres componentes en las perturbaciones que puedan aparecer en los valores parámetros.

Otros de los índices propuestos es el índice de concordancia (d) de Willmott (Willmott's y Wicks, 1980; Willmott, 1981). El índice de concordancia era inicialmente:

$$d = 1 - \frac{\sum_{i=1}^n [(P_i - O_i) - (O_i - \bar{O}_i)]^2}{\sum_{i=1}^n (|P_i - \bar{O}| + |P_i - \bar{O}|)^2} \quad (3.3)$$

Willmott razonó este índice para describir la covariabilidad entre los valores simulados y predichos con el valor medio de los observados. Inicialmente, los valores predichos respecto a la variabilidad de la media de los observados se comportaban de forma inversa a la diferencia entre los valores observados y los valores medios de los observados.

Tiempo después, Willmott y sus colaboradores se dieron cuenta que los errores grandes si se hacían al cuadrado influían de forma excesiva, proponiendo una modificación del índice de concordancia por el índice de concordancia modificado (d_j) (Willmott et al. 1985).

$$d_j = 1 - \frac{\sum_{i=1}^n |O_i - P_i|^j}{\sum_{i=1}^n (|P_i - \bar{O}| + |O_i - \bar{O}|)^j} \quad (3.4)$$

La principal ventaja del índice modificado respecto de su antecesor es que se aproxima este índice al valor 1 (ajuste perfecto) de forma más lenta que la aproximación que se produce entre los valores simulados respecto de los observados. Por tanto, proporcionan una importante diferencia cuando se comparan modelos relativamente buenos. También es menos sensible que la forma de distribución de error-frecuencia y a los valores extremos.

Legates y McCabe (1999) proponen una nueva expresión para el índice de concordancia modificado (Willmott et al., 1985). Con esta modificación se permite emplear un valor base de referencia (para el caso de los recursos hídricos pueden ser los valores medios mensuales para cada mes) y obtener así un resultado adimensional que oscila entre 0 y 1 evitando la utilización de los errores al cuadrado y, por tanto, utilizando un peso adecuado a la diferencia de los errores.

$$1 - \frac{\sum_{i=1}^N |O_i - P_i|}{\sum_{i=1}^N (|P_i - \bar{O}| + |O_i - \bar{O}|)} \quad (3.5)$$

Recientemente, Willmott (2011) ha modificado su índice d_1 por otro que presenta variación en sus límites (d_r), ya que los anteriores variaban desde 0 a 1, este permite variar desde -1 a 1. Este hecho es debido a la relación que mantiene con error absoluto medio (MAE por sus siglas en inglés). El objetivo ha sido tratar de solucionar el hecho de que en caso de mostrar el resultado 0 solo indicaba que los valores medios eran mejores que los simulados por el modelo, pero no proporcionaba más información. Con este nuevo intervalo por ejemplo puede mostrarse que un valor $d_r = 0$ nos indica, que la suma de la

magnitud de los errores y la suma de los errores entre el modelo y los valores observados son iguales. Para un valor $d_r=-1$ nos está diciendo que el modelo es peor estimador que el valor medio de los observados.

Respecto al error cuadrático medio (MSE por sus siglas en inglés), éste representa el promedio de la diferencia al cuadrado entre los valores simulados y los históricos. Si a este error le aplicamos su raíz cuadrada, entonces se obtiene la raíz del error cuadrático medio (RMSD):

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (P - O_i)^2} \quad (3.6)$$

Otra forma de establecer las diferencias de los estadísticos sintéticos con respecto a los históricos, es la raíz de la desviación media cuadrada relativa, RRMSD, índice de juicio que resulta apropiado cuando se tiene un número importante de resultados para contrastar (Fernández y Salas, 1986).

De dos modelos ajustados será mejor aquél que tenga menor RRMSD. Éste se define como la raíz cuadrada del promedio de las diferencias cuadráticas entre el valor medio de un estadístico sintético y su homólogo de la serie histórica, referida a este último.

$$\text{RRMSD} = \frac{\sqrt{\sum_{i=1}^N (\bar{P} - O_i)^2 / N}}{\bar{O}} \quad (3.7)$$

donde:

$$\bar{P} = \frac{1}{N_{sr}} \sum_{i=1}^{N_{sr}} P_i$$

N_{sr} = Numero de series sintéticas generadas.

O_i =valores observados.

N = Número de observaciones.

\bar{O} = Media de los valores observados.

Además, existen un par de razones que justifican el uso de la RRMSD para juzgar el rendimiento de un modelo frente a otro, según su capacidad de preservación de estadísticos. La primera de ellas es porque las diferencias correspondientes a los umbrales más altos en los estadísticos de sequía y

almacenamiento están más penalizadas que las que corresponden a los umbrales más bajos. De esta forma se cuida más la preservación de los estadísticos para los umbrales más cercanos al caudal medio de la serie. De este modo, se favorece la simulación de la gestión de sistemas de recursos hídricos, si se tiene en cuenta que las demandas suelen ser valores cercanos a los caudales medios históricos, tal como se reporta en algunos estudios sobre esta materia presentados por Vogel y Stedinger (1988) y; Sánchez et al., (2001).

La segunda razón es, que debido a que RRMSD es una medida referida a los valores medios de los estadísticos históricos, permite tener una idea de la magnitud de la desviación de los estadísticos sintéticos con respecto a los primeros.

Akaike (1974), propone estudiar el problema de la identificación del mejor modelo, desde el punto de vista de la decisión estadística, es decir, traslada el criterio de elección del modelo más adecuado atendiendo a aquel que presente una función de pérdida menor, (valor AIC más pequeño). La solución propuesta por Akaike es el AIC, el criterio de especificación con el que pretende seleccionar el modelo más adecuado a partir de la diferencia entre dos distribuciones (para nuestro caso la diferencia entre los valores históricos y los generados por el modelo). De este modo, para el presente trabajo, se trata de seleccionar el modelo más adecuado para la generación de caudales sintéticos, a partir de su expresión general:

$$AIC = -2 \ln(\text{máxima verosimilitud}) + 2(\text{n}^{\circ} \text{ de parámetros independientemente ajustados})$$

En este caso, el criterio Akaike quedará definido de la siguiente forma:

$$AIC = -\frac{2}{N} \ln(RMSD) + \frac{2K}{N} \quad (3.8)$$

donde:

N: Número de series históricas disponibles.

RMSD: raíz del error cuadrático medio

K: número de parámetros del modelo empleado.

El criterio Akaike es uno de los índices más empleados para comparar diferentes modelos de una misma categoría. Por eso, no es adecuado para comparar los modelos empleados en este ejercicio puesto que pertenecen a distintas categorías: modelos autorregresivos y modelos de redes neuronales.

Esto plantea la necesidad de buscar otros criterios similares al AIC, pero que salven dicho inconveniente.

Como otros criterios se puede seleccionar, en primer lugar, el “*Bayesian information Criteria*” (BIC) (Akaike 1978b, 1979) que surge a partir del AIC al introducirle una modificación de tipo bayesiano, tal como se indica en la siguiente expresión general:

$$BIC = \{-2 \ln(\text{máxima verosimilitud}) + \ln(N) k\} \quad (3.9)$$

Si se compara el criterio AIC con el BIC se puede ver que la diferencia principal se encuentra en BIC ya que este último penaliza mucho más los modelos con un mayor número de parámetros estimados (debido a la sustitución del 2 por $\ln(N)$) y por tanto se intenta corregir la sobreestimación observada en los resultados del AIC.

Este criterio también se descarta para el presente caso de estudio porque tampoco permite comparar modelos con diferente base matemática.

Otro criterio es el “*Final prediction error*” (FPE) propuesto también por Akaike (1969). Este índice presenta un criterio elaborado inicialmente para los procesos autorregresivos, en el que concretamente se basa en la estimación del error cuadrático medio de predicción a un paso. Esta dado por:

$$FPE = \frac{1 + \left(\frac{k}{n}\right)}{1 - \left(\frac{k}{n}\right)} \hat{\sigma}^2 \quad (3.10)$$

Para los casos de estudio del presente ejercicio, puede reformularse como:

$$FPE = \frac{n+q}{n-q} RMSD^2 \quad (3.11)$$

donde:

N: la diferencia entre número total de observaciones (N) y parámetros del modelo (k).

q: el orden del proceso autorregresivo.

Gracias a este índice puede compararse los modelos autorregresivos y los modelos de redes neuronales mediante la estimación del error cuadrático. Otra

particularidad que presenta este índice es que en el caso de disponer de un número suficientemente elevados de observaciones, puede relacionarse con el AIC de la siguiente forma:

$$AIC \approx N \ln(FPE) \quad (3.12)$$

Por último, quedaría por hablar del coeficiente de persistencia, PI (Kitadinis y Bras, 1980; Corradini et al., 1986), que se utiliza para comparar el rendimiento del modelo al comparar el valor observado $n-1$ con el valor generado por el modelo para el valor n .

El coeficiente de persistencia permite comparar las predicciones del modelo con las predicciones obtenidas en el supuesto de que el proceso se tratase como uno de Wiener (Varianza creciente linealmente con el tiempo), en cuyo caso, la mejor estimación para el futuro está dada por la última medición.

La eficiencia de la persistencia del modelo es un cálculo estadístico normalizado que cuantifica la magnitud relativa de la varianza residual (ruido) respecto a la varianza de los errores obtenidos por el uso de un modelo de persistencia simple (Moriasi et al., 2007).

El índice PI oscila entre 0 a 1, siendo este último el valor óptimo y debe ser mayor que 0 para indicar un rendimiento del modelo como mínimo, aceptable. Valores por debajo de cero indican que el modelo no está captando bien la información original.

Pese a que en un primer momento, el índice de persistencia podría ser muy útil en el caso de trabajar con registros diarios, se ha decidido incluirlo en el presente trabajo debido a que indica si el modelo está reproduciendo los datos anteriores observados o si por el contrario está generando nuevos datos.

En la tabla 3.1 se muestran los índices empleados para comparar los diferentes modelos propuesto en el trabajo final de máster:

Tabla 3.1: Estadísticos propuestos para el análisis de los diferentes modelos.

Estadístico	Identificador	Fórmula	Definición
Coefficiente de determinación	R^2	$\left\{ \frac{\sum_{i=1}^N (O_i - \bar{O}) (P_i - \bar{P})}{\sqrt{[\sum_{i=1}^N (O_i - \bar{O})^2] [\sum_{i=1}^N (P_i - \bar{P})^2]}} \right\}^2$	<p>O_i=Valores observados.</p> <p>\bar{O}= Media de los valores observados.</p> <p>P_i=Valores generados.</p> <p>\bar{P}= Media de los valores generados.</p>
Raíz relativa del error cuadrático medio	RRMSD	$\frac{\sqrt{\sum_{i=1}^N (\bar{P} - O_i)^2 / N}}{\bar{P}}$ $\bar{P} = \frac{1}{N_{sr}} \sum_{i=1}^{N_{sr}} P_i$	<p>O_i= Valores observados.</p> <p>N= Número de observaciones.</p> <p>\bar{O}= Media de los valores observados.</p> <p>N_{sr}= Número de series sintéticas generadas.</p>
Índice de concordancia modificado	d_1'	$1,0 - \frac{\sum_{i=1}^N O_i - P_i }{\sum_{i=1}^N (P_i - \bar{O}' + O_i - \bar{O}')}$	<p>O_i=Valores observados.</p> <p>\bar{O}' = Valor referencia observados.</p> <p>P_i= Valores generados.</p>
Índice Nash-Sutcliffe	CE	$1,0 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}$	<p>O_i=Valores observados.</p> <p>P_i=Valores generados.</p> <p>\bar{O}= Media valores observados.</p>
Índice Kling-Gupta	KGE	$KGE = 1 - ED$ $ED = \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}$ $\alpha = \frac{\sigma_P}{\sigma_O} \quad \beta = \frac{\bar{P}}{\bar{O}}$	<p>σ_0=Desv. Estándar valores observados.</p> <p>σ_P=Desv. Estándar valores generados.</p> <p>\bar{O}= Media valores observados.</p> <p>\bar{P}= Media valores generados.</p> <p>r = Coeficiente correlación lineal.</p>
Error de Predicción Final	FPE	$\frac{n + q}{n - q} RMSD^2$	<p>N:= la diferencia entre número total de observaciones (N) y parámetros del modelo (k).</p> <p>q:=el orden del proceso autorregresivo.</p>
Índice persistencia	PI	$1,0 - \frac{\sum_{i=1}^N (O_i - P_i)^2}{\sum_{i=1}^N (O_i - O_{i-1})^2}$	<p>O_i=Valores observados.</p> <p>P_i=Valores generados.</p> <p>P_{i-1}= Valor generado $i-1$</p>

Capítulo 4

Redes neuronales artificiales estocásticas

4.1 Introducción.

La neurona es el elemento básico de una red neuronal al igual que ocurre en los sistemas biológicos. La neurona artificial, más conocida como nodo, consta de conexiones tanto de entrada como de salida, y en su núcleo ó unidad central se procesan los datos que le permiten aprender y generalizar la información que recibe mediante los mismos. En la figura 4.1 se puede observar la arquitectura básica de una red neuronal.

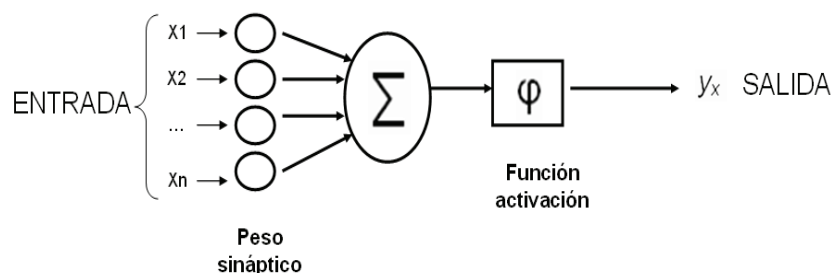


Figura 4.1: Arquitectura básica de una red neuronal artificial.

Cada neurona de la red procesa la información de entrada para generar una respuesta, la cual depende en primer lugar la información de entrada y, en segundo lugar de las operaciones que se realicen en el interior de la neurona. De esta forma, cuando se entrena a la neurona para obtener una salida adecuada a partir de las entradas será necesario modificar los valores de los parámetros libres para conseguir que la diferencia entre el valor de salida y el valor observado sea la mínima posible. Este procedimiento básico sienta las bases de entrenamiento para los sistemas basados en la neurona como unidad básica.

Para poder explicar el proceso de aprendizaje de una red neuronal es importante definir en primer lugar una serie de conceptos:

- Estado de activación.

Conocido también como *entrada neta*, *potencial postsináptico* o *campo local inducido*, es un valor que representa el nivel de actividad de la neurona, y que matemáticamente equivale a la agregación ponderada de las señales individuales que recibe de otras neuronas. Se dice que una neurona es activa, si su estado de activación es positivo, y pasiva, si es negativo.

- Funciones de activación.

Denominadas también *funciones de salida* o *funciones de transferencia*, transforman el estado actual de activación de una neurona en una señal de salida. Existen varios tipos: escalón o umbral, lineal, mixta (combinación de las dos anteriores), sigmoideal (logísticas y tangente-hiperbólica), gaussiana, polinómica, y combinaciones de ellas, entre otras.

- Conexiones sinápticas.

Usualmente referidas sólo como conexiones, son los vínculos entre las neuronas, a través de los cuales se transmite la información de unas a otras. A cada conexión de la red le corresponde un peso sináptico, cuya magnitud representa la intensidad o fuerza de la conexión; dichos pesos almacenan el conocimiento de la red, es decir, constituyen una parametrización de las relaciones matemáticas entre las variables presentadas a la red.

Según la conectividad, esto es, la forma en que están canalizadas las salidas de unas neuronas hacia otras, se dice que la red está alimentada hacia adelante (feedforward) o que es una red recurrente; en el primer caso, ninguna salida de las neuronas es entrada de neuronas de la misma capa o de capas anteriores; en el segundo caso, existen.

- Umbral de activación.

Parámetro asociado a cada neurona, representa el desfase entre la entrada neta a ella y su respectiva función de activación, y que suele tratarse como un peso sináptico más de la red neuronal. Para ello se adiciona en cada capa de ésta, excepto la de salida, una neurona denominada ficticia, que no recibe ninguna señal de entrada y cuya señal de salida es unitaria.

4.1.1 Modelo matemático de las redes neuronales determinísticas.

El modelo matemático de una neurona artificial es la base de la estructura de cualquier tipo de red neuronal artificial. En la Figura 4.2 se presenta un esquema del modelo en cuestión para una neurona de una red determinística, en el que se distinguen tres partes constituidas por una cierta información de entrada, su procesamiento, y la respectiva salida.

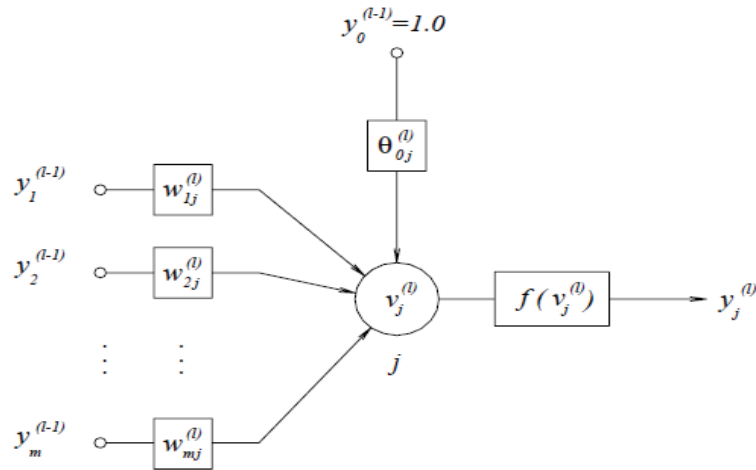


Figura 4.2: Esquema de una neurona de una red determinística.

Fuente: Ochoa-Rivera.2002

Después de que las neuronas de una capa dada de la red (capa $l-1$) emiten sus señales de salida, $y_i^{(l-1)}$, cada una de éstas es modificada o modulada según la intensidad o peso sináptico, $w_{ij}^{(l)}$, de su respectiva conexión con la neurona a la que llegan dichas señales (neurona j de la siguiente capa l); éstas se agregan para generar la entrada neta, $v_j^{(l)}$, a la neurona j , incluyendo en dicha agregación el umbral de activación, $\theta_{0j}^{(l)}$:

$$v_j^{(l)} = \sum_{i=0}^{m_{l-1}} w_{ij}^{(l)} y_i^{(l-1)} + \theta_{0j}^{(l)} \quad (4.1)$$

En esta ecuación, el umbral de activación $\theta_{0j}^{(l)}$, puede asemejarse al peso sináptico correspondiente a la conexión que llega a la neurona j desde una neurona ficticia (neurona 0) de salida unitaria, $y_0^{(l-1)} = 1.0$, lo cual hace que la ecuación anterior adopte la siguiente forma:

$$v_j^{(l)} = \sum_{i=1}^{m_{l-1}} w_{ij}^{(l)} y_i^{(l-1)} \quad (4.2)$$

En la cual, $\theta_{oj}^{(l)}$ toma la notación $w_{oj}^{(l)}$, análoga a la de los pesos de las demás conexiones.

Una vez se tiene evaluada la entrada neta $v_j^{(l)}$, la función de activación, f , de la neurona j se encarga de transformar dicha entrada en una señal de salida, $y_j^{(l)}$, la cual se emite a las siguientes neuronas en forma de entrada a estas:

$$y_j^{(l)} = f(v_j^{(l)}) \quad (4.3)$$

4.1.2 Proceso de aprendizaje.

El *aprendizaje*, también conocido como entrenamiento o calibración del modelo matemático subyacente a la red, puede efectuarse mediante diferentes paradigmas, así como distintos algoritmos matemáticos. Los más comunes se detallan a continuación:

4.1.2.1 Aprendizaje supervisado.

En este tipo de aprendizaje se comprueba la salida de la red y , en caso de que no coincida con la deseada, se modifican los pesos de las conexiones hasta lograrlo. Es el más utilizado en las aplicaciones hidrológicas (ASCE, 2000a). Dentro del aprendizaje supervisado se pueden diferenciar tres clases:

- Aprendizaje por corrección de error: Los pesos sinápticos de la red se modifican en función de la diferencia entre los valores deseados y los valores obtenidos en la salida de la red neuronal, es decir, en función del error cometido por ésta. Algunos algoritmos que se enmarcan dentro de este tipo de aprendizaje son la regla delta y la retropropagación del error (error-backpropagation). Siendo éste último una modificación de la primera regla de aprendizaje y el más empleado en los estudios de hidrología.
- Aprendizaje por refuerzo: En este caso se emplea una señal de refuerzo para decirle a la red neuronal si la salida de que de ella se obtiene es exitosa o no. Algunos ejemplos de este tipo de aprendizaje lo constituyen el algoritmo lineal con recompensa y penalización (linear

reward-penalty), de forma generalizada aparece el algoritmo asociativo con recompensa y error (associative rewardpenalty) entre otros.

- Aprendizaje estocástico: Consiste en realizar cambios aleatorios en los valores de los pesos y evaluar su efecto a partir del valor deseado y de distribuciones de probabilidad preestablecidas. Un tipo de red que emplea este aprendizaje es la máquina de Boltzman (Boltzman machine), en la cual siempre se dispone de neuronas ocultas que permiten introducir un ruido, decreciente durante el aprendizaje, para escapar de los mínimos locales de la función de energía, y favorecer la obtención del mínimo global. Otra RNA basada en este aprendizaje es la máquina de Cauchy (Cauchy machine).

4.1.2.2 Aprendizaje no-supervisado.

En este caso la red neuronal artificial no recibe información del entorno que le indique si la salida es o no correcta, y por ello se dice que este tipo de redes son capaces de auto-organizarse. De este aprendizaje existen dos clases:

- Aprendizaje hebbiano: El ajuste del peso de cada conexión se realiza según la correlación de los estados de activación (o señales de salida) correspondientes a las dos neuronas conectadas: si ambas son neuronas activas, su conexión se refuerza, pero si una es activa y otra pasiva, la conexión se debilita.
- Aprendizaje competitivo. Consiste en que las neuronas de la capa de salida de una RNA compiten por activarse, de tal forma que sólo una neurona puede estar activa en un instante dado. De esta manera se pueden descubrir rasgos estadísticos sobresalientes en los datos, útiles para su clasificación, la cual se puede hacer mediante correlaciones entre ellos, realizadas por la propia red.

En este trabajo se ha empleado el aprendizaje supervisado y el algoritmo de retropropagación del error. Este tipo de entrenamiento provee al sistema con información de entrada al igual que se proveen las salidas esperadas, de este modo el sistema tiene las salidas como punto de referencia y en función de la diferencia entre las salidas observadas y las calculadas se modifican los parámetros libres hasta encontrar sus mejores valores en sucesivas iteraciones..

4.1.3 El perceptron multicapa.

En el perceptron simple, la delimitación de las regiones de decisión se hace mediante una frontera cuya ecuación queda determinada en función de los pesos de las conexiones sinápticas y del valor del umbral de activación de la neurona de procesamiento. Esta característica, denominada *separabilidad lineal*, hace que dicho modelo no sea aplicable a problemas con regiones de decisión convexa o más compleja, tal como suele ocurrir en muchos problemas de Hidrología.

La *separabilidad lineal* se supera añadiendo una nueva capa de neuronas al perceptrón simple, lo cual da lugar a las *redes perceptron multicapa* (multilayer perceptron, MLP), cuya arquitectura tipo aparece mostrada en la figura 4.3. Así se pueden generar regiones de decisión arbitrariamente complejas aumentando de esta manera la capacidad de la red para resolver problemas más complejos. Algunos autores como Lapedes y Farber (1988) muestran que en las redes MLP no se requieren más de dos capas ocultas para resolver la mayoría de los problemas, y Hornik et al. (1989) demuestran, mediante el *teorema de aproximación universal*, que con dichas redes se puede representar el comportamiento de cualquier función continua, razón por la que también se las conoce como *aproximadores universales*.

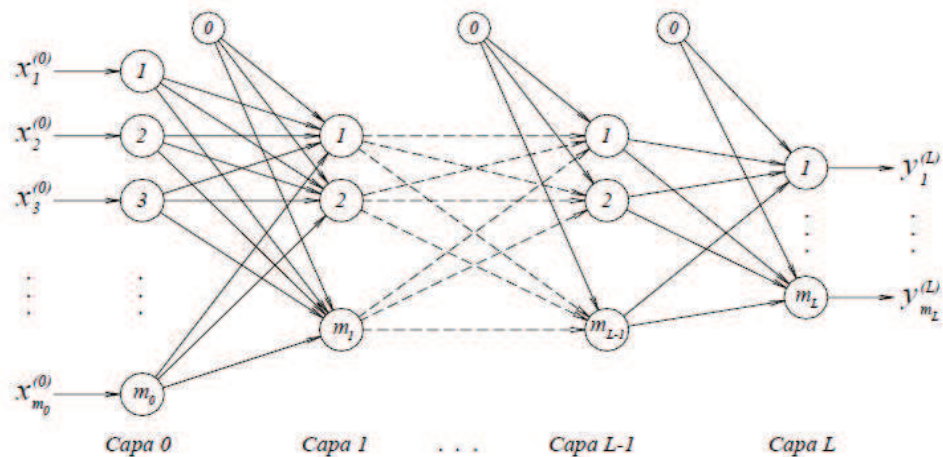


Figura 4.3: Configuración tipo de una red perceptron multicapa.

Fuente: Ochoa-Rivera.2002

4.1.4 Funciones de activación empleadas.

Aunque existen un gran número de funciones de activación, en este estudio solo se muestran las que se emplean mayoritariamente en el entrenamiento y operación de redes neuronales.

Si se atiende al punto de vista matemático se pueden encontrar funciones de activación continuas (generalmente aparecen función escalón y función lineal) o discontinuas. En el ámbito de la Hidrología y de los sistemas de recursos hídricos, las discontinuas son poco empleadas puesto que en este campo muchas de las funciones de activación de estructuran a partir de funciones derivadas, si bien es cierto que debe ser monótona creciente y definida en el intervalo de posibles valores de entrada (Haykin, 1999). Un tipo de funciones que se adaptan a estos criterios son las funciones sigmoidales, que de hecho, suelen ser las más comúnmente empleadas en estudios hidrológicos (Maier y Dandy, 2000).

a) Funciones escalón.

Este tipo de funciones, indicadas en la Figura 4.4, y denominadas también funciones umbral (threshold), se aplican a redes con salidas binarias, es decir, con dos posibles valores: 1 y 0, ó 1 y -1; cuando la suma de las entradas es igual ó mayor que el umbral de activación de la neurona (0 en el caso que se representa en la Figura 4.4), su estado de activación es 1, y si es menor, dicho estado es 0 (ó -1).

$$f(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases} \quad f(v) = \begin{cases} 1, & v \geq 0 \\ -1, & v < 0 \end{cases} \quad (4.5)$$

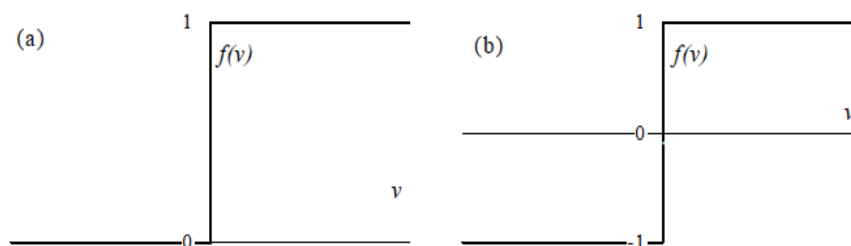


Figura 4.4: Funciones escalón a) rango [0; 1] b) rango [-1; 1]

Fuente: Ochoa-Rivera.2002

b) Funciones mixtas.

Estas funciones surgen como el resultado de la combinación entre una función lineal y una función escalón (figura 4.5). En este tipo de funciones, la señal de salida de la neurona es 0 (ó -1), si la entrada neta es menor que un cierto límite inferior; es 1, si la entrada neta es mayor que un límite superior; o es un valor entre 0 (o -1) y 1, que es función lineal de la entrada neta, si ésta se encuentra entre los límites inferior y superior. Ello hace que la forma de esta función tenga un carácter aproximadamente no-lineal.

$$f(v) = \begin{cases} 1, & v \geq 0 \\ v, & -a \leq v \leq a \\ 0, & v < 0 \end{cases} \qquad f(v) = \begin{cases} 1, & v \geq 0 \\ v, & -a \leq v \leq a \\ -1, & v < 0 \end{cases} \quad (4.6)$$

Cabe anotar que la función lineal sola también puede emplearse como función de activación, especialmente en las neuronas de la capa de salida.

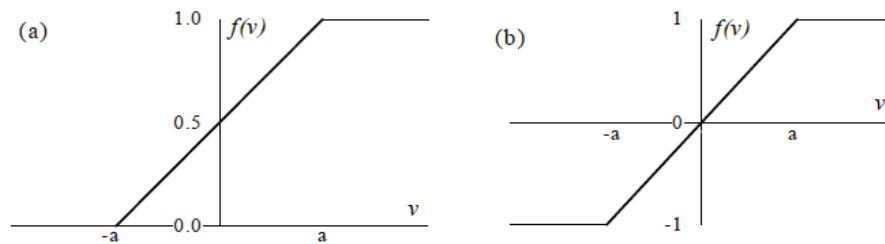


Figura 4.5: Funciones mixtas a) rango [0; 1] b) rango [-1; 1]
Fuente: Ochoa-Rivera.2002

c) Función sigmoideal.

Funciones sigmoideas son, como ya se ha dicho al principio de este apartado, las más populares, debido a que son funciones no-lineales que se pueden aproximar a la función escalón elevando el valor de su pendiente, y muy especialmente porque son funciones derivables (Hilera y Martínez, 1995). Dentro de esta categoría se encuentra la función logística (Figura 4.6(a)), dada por:

$$f(v) = \frac{c1}{1+\exp(-c3v)} + c2 \quad (4.7)$$

Cuando $c_1 = 2$ y $c_2 = -1$, la función logística adopta como rango el intervalo $(-1; 1)$ esta función es denominada por algunos autores función sigmoideal bipolar (Imrie et al., 2000). Otra forma muy común de la función logística es aquella en la que los valores de c_1 y c_2 son 1 y 0 respectivamente, caso en el cual las salidas quedan comprendidas en el rango $(0,1)$. La constante c_3 se denomina parámetro de pendiente y su efecto sobre la función es hacer que la variación de su pendiente sea mayor cuanto más alto sea el valor de dicho parámetro, es decir, cuando el valor de c_3 crece, la función logística tiende a la función escalón, siendo también la función mixta un caso particular de ella.

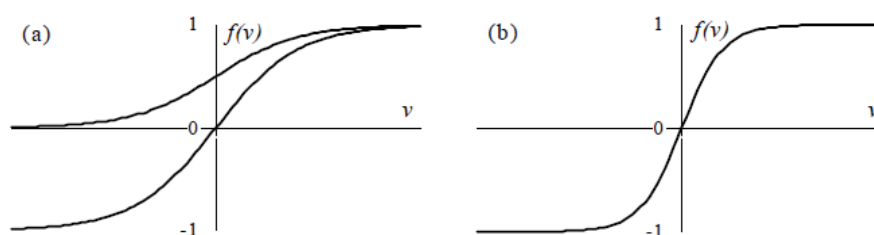


Figura 4.6: Funciones sigmoideas (a) logística; (b) tangente hiperbólica.
Fuente: Ochoa-Rivera.2002

Otro tipo particular de función sigmoideal es la función *tangente hiperbólica* (figura 4.6b), la cual, la constante c_4 es un parámetro que amplifica o reduce el intervalo de las salidas de la función, y cuyo valor suele ser 1, debido a que las redes neuronales artificiales normalmente se entrenan en el rango $[-1,1]$ u otros inferiores; en este último caso, c_4 adopta el valor absoluto de los límites de dicho rango. El parámetro de pendiente, c_5 , tiene el mismo comportamiento que en las funciones logísticas, en esta función, dicha variación es mayor que en la función logística cuando se considera en ambas el mismo valor para el parámetro de pendiente (i.e., para $c_3 = c_5$), tal como se aprecia en la Figura 4.6 (b), en la que $c_3 = c_5 = 1$.

$$f(v) = c_4 \tanh(c_5 v) = c_4 \frac{1 - \exp(-2 c_5 v)}{1 + \exp(-2 c_5 v)} \quad (4.8)$$

La mayoría de los autores recogen en sus trabajos el ensayo prueba-error para conocer la función de activación más idónea para el estudio que se está llevando a cabo, aunque pueden encontrarse algunas premisas ya definidas:

- Si las salidas de la red corresponden a variables continuas, las funciones tipo escalón no se consideran.
- En el caso de necesitar en una capa en particular que las señales de salida sean iguales a las entradas netas, se emplean funciones lineales.

4.1.5 El algoritmo de retropropagación del error (EBLA).

Como ya se ha comentado brevemente en el apartado 4.1.2.1, el EBLA es el algoritmo más importante para el entrenamiento supervisado de redes MLP de alimentación hacia delante, como es el caso de este estudio. Este algoritmo fue presentado por Rumelhart, Hinton y Williams (1986).

Este algoritmo propaga los errores hacia atrás en la red, es decir, atravesando en sentido inverso las distintas capas que la conforman. Para cada uno de los valores presentados a la red, ésta genera unos valores *de salida* que a su vez se comparan con el valor objetivo. Por tanto, el algoritmo EBLA irá modificando adecuadamente en sucesivas iteraciones los pesos de la red para reducir el error.

La forma más habitual de trabajar con éste algoritmo es emplear como función de error el error cuadrático medio (MSE por sus siglas en inglés). En el caso de que la salida de la red fuese un vector de varias componentes, se evaluaría el error cuadrático medio para cada ejemplar, definiéndose, en este caso, la función de error como el promedio extendido a todo el conjunto de patrones de entrenamiento.

La estrategia de búsqueda del EBLA se basa en el conocido método del gradiente, con la función del error cuadrática tal como se muestra en la figura 4.7. En cualquier punto donde se esté habrá distintas pendientes en la superficie y los valores de los pesos ayudarán a seguir la dirección de máxima pendiente, por tanto, se puede decir que este algoritmo emplea el método del gradiente aunque no garantiza encontrar el mínimo global.

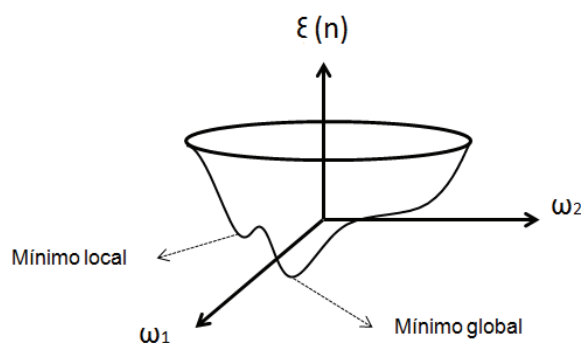


Figura 4.7: Imagen de una superficie de error de la red MLP.

El algoritmo contempla dos fases de cálculo. La primera fase hacia delante, realizando los cálculos de la red para generar las *salidas* a partir de los datos de entrada y la segunda fase, hacia atrás donde el error encontrado en la última capa se propaga hacia atrás modificando los valores de los pesos.

La variación del error respecto del peso $w_{ij}^{(l)}$ viene dada por la derivada parcial. De acuerdo con el criterio del gradiente, se adopta ahora un cambio en el peso $w_{ij}^{(l)}$ dado por $\Delta w_{ij}^{(l)} = \eta \delta x_i$ Donde η es un parámetro del algoritmo denominado *tasa de aprendizaje*. Un valor pequeño del mismo implica una aproximación lenta hacia el límite, por el contrario, un valor grande acelera el proceso. Este valor oscila entre 0 y 1. δ hace referencia al error del nodo que muestra la diferencia entre los valores producidos y los valores deseados para una red de nodos logísticos.

Una vez conocido el error, el algoritmo EBLA es capaz de definir el grado de contribución de los pesos situados en la capa oculta para proceder a modificarlos para obtener unos *outputs* adecuados. La fórmula empleada por el EBLA hace necesario comentar un nuevo parámetro α denominado *momentum* cuya función es favorecer la inercia del sistema para encontrar el mínimo global. De este modo, el cambio de peso queda como $\Delta w_{ij}^{(l+1)} = \eta \delta_j x_i + \alpha \Delta w_{ij}^{(l)}$

Ambos parámetros η y α condicionan el proceso de entrenamiento de forma muy significativa.

Este algoritmo EBLA puede implementarse o bien en *modo batch*, es decir, cuando se termine una época y se hayan extraído los valores de los errores cuadráticos medios, el algoritmo modificará los pesos y volverá a comenzar de nuevo hasta alcanzar el error deseado o hasta llegar a un número determinado de épocas. La otra forma, conocida como *modo secuencial*, se realiza la modificación de los pesos cada vez que se ha presentado un patrón y se ha comprobado el error respecto de la *salida* con el valor deseado.

Para establecer los *criterios de parada* al algoritmo se puede indicar un valor objetivo que una vez superado el entrenamiento de la red se detenga o bien fijar un número máximo de épocas para realizar el entrenamiento.

4.2 Construcción del modelo de red neuronal con la componente estocástica.

Para la construcción del modelo, se emplearán los pasos seguidos en la tesis doctoral de Ochoa-Rivera (2002) *Modelo estocástico de redes neuronales para la síntesis de caudales aplicados a la gestión probabilística de sequías*.

4.2.1 Exploraciones iniciales.

A continuación se resumen las exploraciones realizadas en el estudio de Ochoa-Rivera (2002).

Las primeras en ser probadas fueron algunas funciones univariadas puramente determinísticas, correspondientes a ecuaciones analíticas bien establecidas. Dentro de estas exploraciones iniciales, uno de los experimentos consistió en mapear una función parabólica, $y = f(x) = x^2$, en el intervalo $x^2 [1, 100]$. Para ello se llevaron a cabo los siguientes pasos:

1. Se evaluaron 100 valores de la variable dependiente, y , en el intervalo $[1,100]$, los cuales se agruparon en dos conjuntos de 50 valores cada uno: el primero se empleó para el entrenamiento de la red, y el segundo, para su validación.
2. Se escalaron los valores de y , en términos de su valor máximo y mínimo, para que adoptaran los rangos de las funciones de activación.
3. Se ensamblaron los patrones de entrenamiento a partir de los valores escalados del primer conjunto, considerando que cada uno de ellos, z_i , era función de los tres anteriores, z_{i-1} , z_{i-2} , y z_{i-3} , en virtud de que son tres los puntos necesarios para definir una parábola. De esta forma se tenían 47 patrones, cada uno de los cuales estaba constituido por tres predictores y un criterio, lo que implicaba considerar tres neuronas en la capa de entrada y una en la capa de salida de la red a entrenar.
4. Se entrenó la red neuronal probando distintas tipologías: se empezó con una red sin capas ocultas, luego se probó una de una capa oculta con una sola neurona, después con dos neuronas, y así sucesivamente mientras el error de entrenamiento, RRMSD, fuera disminuyendo hasta que ya no se redujera más; en este punto se almacenó la tipología correspondiente con sus respectivos pesos, la cual se sometería posteriormente al proceso de validación. En cada caso, los parámetros de entrenamiento, definidos en el Apartado 4.2.4 se variaron uniformemente; así, la tasa de aprendizaje tomó

valores en el intervalo (0,1) y la constante de momentum se hizo variar en el intervalo [0,1]. Como funciones de activación se consideraron las sigmoides indicadas en las Ecs. (4.7) y (4.8), función sigmoideal y función parabólica y funciones lineales.

- Finalmente se llevó a cabo el proceso de validación, para el cual se ensamblaron 47 patrones a partir del segundo conjunto de 50 valores de la variable dependiente, no empleados en la fase de entrenamiento. La tipología seleccionada fue aquella que arrojó el menor error, RRMSD, tanto en la fase de entrenamiento como en la de validación.

Los mejores resultados se obtuvieron para una tipología con una sola capa oculta de dos neuronas (Figura 4.8), esto es, una tipología 3 – 2 – 1, y con funciones de activación logísticas. Los resultados de esta red aparecen reflejados en la figura 4.9.

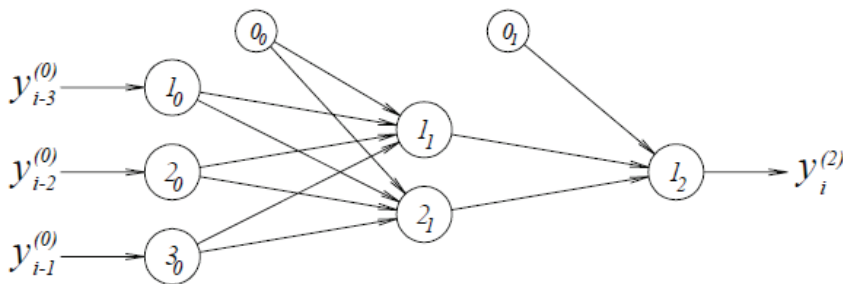


Figura 4.8: red MLP 3-2-1 para el mapeo de la función $y = x^2$
Fuente: Ochoa-Rivera. 2002

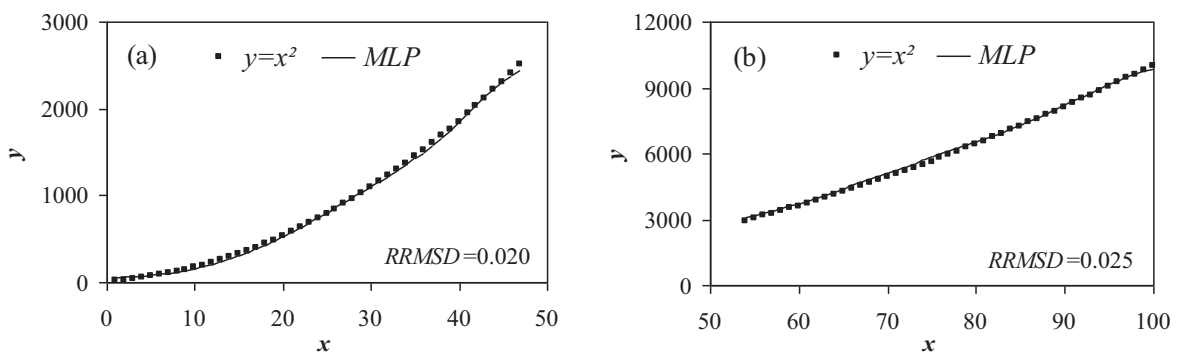


Figura 4.9: Mapeo de la función $y = x^2$ con una red MLP 3-2-1 (a) entrenamiento (b) validación.

Fuente: Ochoa-Rivera. 2002

Los resultados conseguidos en esta primera serie de experimentos muestran que las redes MLP pueden ser eficientes en el mapeo de funciones analíticas univariadas, lo cual es consecuente con el carácter de aproximadores

universales que se les atribuye a dichas redes (Hornik et al., 1989). Ello alentó la realización de una nueva serie de pruebas, en las cuales se incluyeron algunas funciones de series temporales caóticas, entre ellas una serie de Henon, dada por $y_{t+1} = y_t + 3[y_t(1 - y_t)]$. En este caso se siguió un procedimiento análogo al aplicado a la función parabólica. La configuración y los resultados de la red aparecen en la figuras 4.10 y 4.11.

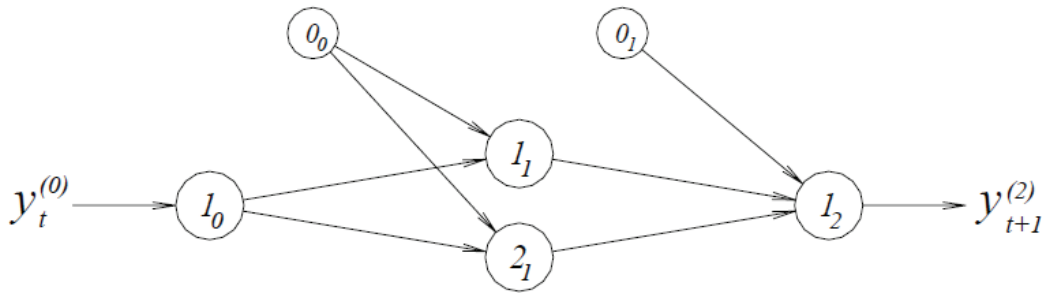


Figura 4.10: red MLP 1-2-1 para el mapeo de una serie de Henon.
Fuente: Ochoa-Rivera.2002

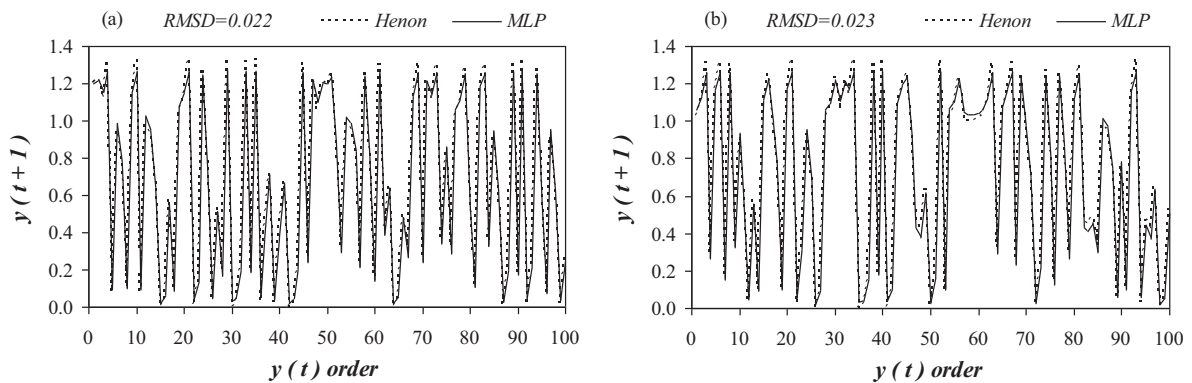


Figura 4.11: Mapeo de la función $y_{t+1} = y_t + 3[y_t(1 - y_t)]$ con una red MLP 1-2-1 (a) entrenamiento (b) validación.
Fuente: Ochoa-Rivera.2002

Cómo se obtuvieron resultados satisfactorios, se decidió realizar un nuevo grupo de experimentos sobre series temporales definidas por modelos autorregresivos univariados. La idea era mapear funciones estocásticas mediante redes MLP, para ver hasta qué punto éstas alcanzaban a reproducir series temporales con ruido, es decir, series con una componente aleatoria. Ello, debido a que en la literatura de las redes neuronales artificiales es común encontrar que se les atribuya a dichas redes la capacidad de tratar datos con ruido (Lehtokangas et al., 1996; ASCE, 2000a; Maier y Dandy, 2000), aunque no se especifica, qué tipo de ruido ni la importancia de éste dentro de la serie de datos.

En primer lugar se estudiaron el comportamiento de un modelo clásico AR(1) con la red neuronal artificial (figura 4.12) y se observó que la red no alcanzaba a captar plenamente el comportamiento aleatorio de la serie temporal. Posteriormente, dentro de este conjunto de pruebas se realizó un segundo experimento, en el cual se consideró una función estocástica autorregresiva AR(2), siguiendo el mismo procedimiento (figura 4.13 y 4.14). Los resultados a los que se llegó son de la misma calidad que los que se obtuvieron en el caso del modelo AR(1).

Los resultados obtenidos en esta fase de experimentación llevaron a concluir que las redes neuronales artificiales requieren la inserción de una componente aleatoria dentro su formulación, para poder reproducir procesos estocásticos en los que la varianza de los ruidos sea significativa, como es el caso de las series de caudales mensuales.

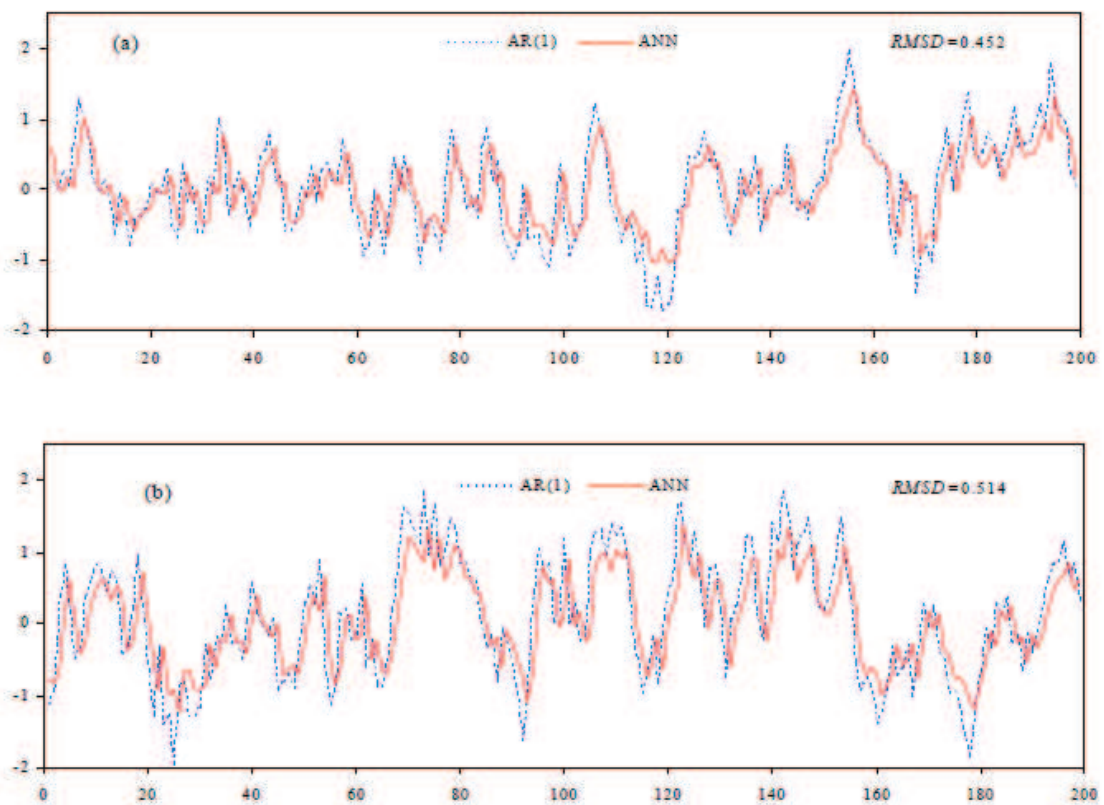


Figura 4.12: Mapeo de la función AR(1), $z_t = 0.796z_{t-1} + 0.481\varepsilon_t$ con una red MLP 1-2-1 (a) entrenamiento (b) validación.

Fuente: Ochoa-Rivera.2002

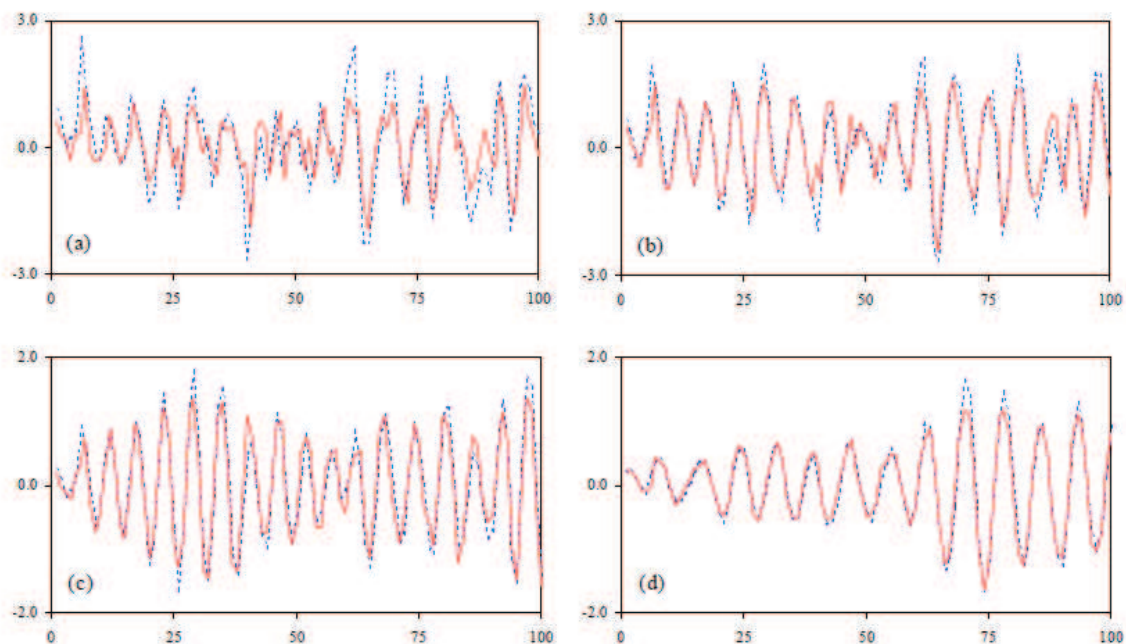


Figura 4.13: Mapeo de modelos AR(2) (línea discontinua) con redes MLP 2 – 1 – 1 (línea continua) en fase de entrenamiento (a) AR(2)1 (b) AR(2)2 (c) AR(2)3 (d) AR(2)4

Fuente: Ochoa-Rivera.2002

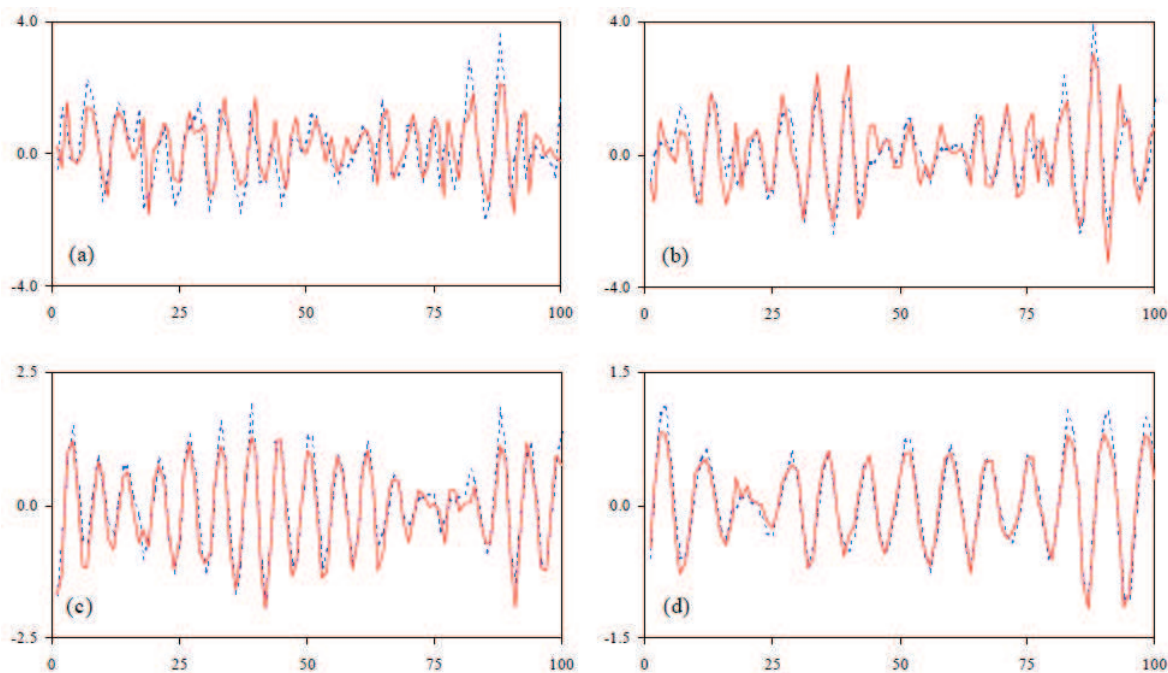


Figura 4.14: Mapeo de modelos AR(2) (línea discontinua) con redes MLP 2 – 1 – 1 (línea continua) en fase de validación (a) AR(2)1 (b) AR(2)2 (c) AR(2)3 (d) AR(2)4

Fuente: Ochoa-Rivera.2002

En el estudio previo a este trabajo y para tratar de corroborar la hipótesis de la importancia del componente aleatorio, se analizó un nuevo experimento con el modelo autorregresivo AR(2) siguiendo el método citado por Stern (1996). 6

modelos AR(2) fueron mapeados con distintas configuraciones de redes MLP. Con este estudio se ha podido mostrar, que cuanto mayor es la componente aleatoria, el RRMSD aumentaba tal y como aparece reflejado en la figura 4.15, demostrando así lo indicado en el estudio de Ochoa-Rivera (2002).

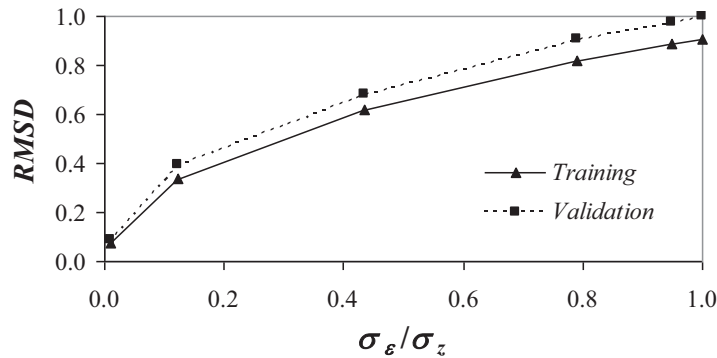


Figura 4.15: Error de mapeo de los procesos AR(2).

Fuente: Ochoa-Rivera et al. (En revisión).

4.2.2 Preprocesamiento de caudales.

El primer paso para la construcción del modelo consiste en preprocesar las series de caudales que se emplean en el entrenamiento de la red neuronal. Éste es un procedimiento que tiene como fin hacer más eficiente el desempeño de la red, y está constituido básicamente por tres etapas: normalización, tipificación y escalamiento. En este sentido, las redes neuronales artificiales no son diferentes a los modelos clásicos de series temporales, y al igual que en éstos, en aquéllas sus resultados están bajo la influencia –significativa en muchos casos– del preprocesamiento de las series de datos (Maier y Dandy, 2000).

Con la normalización se trata de que los datos adopten –si no las tienen– distribuciones de probabilidad normal, ya que de esta forma la red capta más rápidamente los patrones que se le presentan.

Para comprobar que las series sintéticas que se vayan generar presenten una distribución normal, se calcula el intervalo de normalidad del sesgo (Snedecor y Cochran, 1967. Citados por Salas et al., 1980).

$$[-3,9601 N^{-0,4598}; +3,9601N^{-0,4598}] \quad \text{para } N < 150 \quad (4.9)$$

$$[-1,96 \frac{6}{N}; +1,96 \frac{6}{N}] \quad \text{para } N > 150 \quad (4.10)$$

Se considera estadísticamente que si el valor del coeficiente del sesgo de caudales está dentro del intervalo, Ecs. (4.9) y (4.10) la serie está distribuida normalmente.

Las funciones de normalización con las que se ha trabajado han sido las siguientes:

- Normalización raíz cuadrática.

$$X_{vt} = \sqrt{Q_{vt}} \quad (4.11)$$

- Normalización logarítmico-neperiana simple.

$$X_{vt} = \text{Ln}(Q_{vt} + 1) \quad (4.12)$$

- Normalización logarítmico-neperiana doble.

$$X_{vt} = \text{Ln}[\text{Ln}(Q_{vt} + 1) + 1] \quad (4.13)$$

- Normalización potencial.

$$X_{vt} = (Q_{vt} - a_1)^{a_2} \quad (4.14)$$

- transformación normalizadora, Raman y Sunilkumar (1995).

$$X_{vt} = \text{Log}(Q_{vt} + C_T \bar{Q}_T) \quad (4.15)$$

$$C_T = \frac{a}{g_T^2} \quad (4.16)$$

En la ecuación (4.15) y (4.16), Q_{vt} es el caudal en su escala original, correspondiente al mes t del año v ; \bar{Q}_T es la media de los caudales del mes t ; a es un parámetro adimensional a calibrar, que puede tomar valores entre 0 y 1, siendo 0,30 y su entorno una zona de partida recomendable; g_T es el coeficiente de asimetría de los caudales originales del mes t .

Por otro lado, el ajuste de Fourier se aplica con el objetivo de reducir el número global de parámetros del modelo estocástico. Por tanto, con la aplicación de esta técnica se pretende respetar el principio de parsimonia estadística.

El ajuste en series de Fourier debe llevarse a cabo con cuidado cuando se trata de series mensuales (Salas et al., 1980), ya que si no se emplea adecuadamente se corre el riesgo de obtener un modelo que dé lugar a caudales sintéticos cuyos estadísticos difieran significativamente de los históricos.

En este paso se calculan las ecuaciones 4.17 a 4.20 para obtener los coeficientes A_j y B_j de Fourier para todos los armónicos. Como este caso utiliza series mensuales, el número de armónicos será igual a $\omega/2 = 12/2 = 6$ armónicos.

$$A_j = \frac{2}{\omega} \sum_{t=1}^{\omega} U\tau \cos\left(\frac{2\pi j\tau}{\omega}\right) \quad j=1, \dots, h-1 \quad (4.17)$$

$$A_j = \frac{1}{\omega} \sum_{t=1}^{\omega} U\tau \cos\left(\frac{2\pi j\tau}{\omega}\right) \quad j=h \quad (4.18)$$

$$B_j = \frac{2}{\omega} \sum_{t=1}^{\omega} U\tau \sin\left(\frac{2\pi j\tau}{\omega}\right) \quad j=1, \dots, h \quad (4.19)$$

con
$$U\tau \begin{cases} X\tau \\ S_{X\tau} \end{cases}$$

donde:

ω : Número de periodos del año (12 para series mensuales).

τ : Orden del mes del año.

$U\tau$: Media (ó desviación típica) mensual de la serie normalizada para el mes τ .

j : Orden del armónico.

h : Número total de armónico. (Es $\frac{\omega}{2}$ dado que ω es par).

A continuación se calcula la representación en series de Fourier para la media y la desviación típica mensual de la serie normalizada, en la cual se consideran todos los armónicos.

$$V\tau = u + \sum_{j=1}^h \left[A_j \cos\left(\frac{2\pi j\tau}{\omega}\right) + B_j \sin\left(\frac{2\pi j\tau}{\omega}\right) \right] \quad \tau=1, \dots, \omega \quad (4.20)$$

con
$$u = \frac{1}{\omega} \sum_{t=1}^{\omega} u_{\tau}$$

Con los coeficientes de Fourier se calcula la varianza explicada por cada armónico con el fin de construir el periodograma acumulado y decidir, a partir de este, el número de armónicos significativos para la media y la desviación típica mensual de la serie normalizada.

$$C_j^2 = A_j^2 + B_j^2 \quad (4.21)$$

donde: $\frac{C_j^2}{2}$ es la varianza explicada por el armónico j .

Con la varianza explicada y la varianza total se obtiene el porcentaje de la varianza explicada por cada uno de los armónicos. Ec. (3.12).

$$\%Vexp = \frac{C_j^2}{2s_u^2} \quad (4.22)$$

con
$$s_u^2 = \frac{\sum_{t=1}^{\omega} (u\tau - u)^2}{\omega}$$

donde:

$\%Vexp$: Porcentaje de la varianza total explicada de las medias (ó desviaciones típicas) mensuales de la serie normalizada antes del ajuste de Fourier.

s_u^2 : Varianza de las medias (ó desviaciones típicas) mensuales de las series normalizadas.

Una vez definido el porcentaje de varianza que se acepta sea explicada por los estadísticos ajustados (en nuestro caso el 90%) se entra en el periodograma acumulado y se calcula la media y la desviación típica.

$$\widehat{V}_\tau = u + \sum_{j=1}^h \left[A_j \cos\left(\frac{2\pi j\tau}{\omega}\right) + B_j \operatorname{sen}\left(\frac{2\pi j\tau}{\omega}\right) \right] \quad j = h_1, \dots, h_s \quad (4.23)$$

$$\widehat{V}_\tau \left\{ \begin{array}{l} \widehat{X}_\tau \\ \widehat{S}_{X\tau} \end{array} \right.$$

donde:

\widehat{X}_τ : Media mensual del mes τ de la serie normalizada.

$\widehat{S}_{X\tau}$: Desviación típica mensual del mes τ de la serie normalizada.

h_1 : primer armónico significativo.

h_s : Último armónico significativo.

Por otro lado, la tipificación consiste en la remoción de las tendencias y periodicidades de las series de caudales (Ecuación 4.24). La razón de esta transformación estriba fundamentalmente en que el proceso de entrenamiento de la red cobra mayor agilidad cuando se realiza con las variables tipificadas, (Haykin, 1999). De este modo, permite evaluar las correlaciones entre las distintas series, las cuales sirven para establecer el tipo de modelación a partir de las matrices de correlación cruzada y de las funciones de autocorrelación.

$$\text{Serie mensual: } z_{v\tau} = \frac{X_{v\tau} - \widehat{x}_\tau}{\widehat{s}_{x\tau}} \quad \text{Serie anual: } z_v = \frac{X_v - \widehat{x}}{\widehat{s}_x} \quad (4.24)$$

En algunas investigaciones se reporta el uso de la transformación Wilson-Hilferty, sobre las series normalizadas (Fernández y Salas, 1986), sobre las series normalizadas y tipificadas (Raman y Sunilkumar, 1995), con el fin de reducir más aún su coeficiente de asimetría. Si bien es cierto que hace aumentar el número de parámetros, hecho que se ha tenido en cuenta a la hora de realizar este estudio.

El escalamiento tiene como finalidad transformar las series de caudales, en series que adopten una escala análoga al rango de las funciones de activación.

4.2.3 Componente determinística multivariada.

La primera capa de la red (capa 0) transmite las señales, $x_p^{(0)}$, que recibe la red desde el medio externo hasta la primera capa de neuronas de procesamiento, en este estudio, la primera capa oculta. En la capa de entrada no se realizan transformaciones de ningún tipo, ya que no existen funciones de activación, sus señales de salida, $y_p^{(0)}$, son iguales a las señales de entrada. Esta primera capa permite definir las señales que intervienen en el cálculo de la de entrada neta de cada neurona de la primera capa oculta.

$$y_p^{(0)} = x_p^{(0)}; p = 1, 2, \dots, m_0 \quad (4.25)$$

El procesamiento de la información se inicia en las neuronas de la primera capa oculta (capa 1): sus entradas netas, $v_{j1}^{(1)}$, equivalen a la suma ponderada de las señales de salida de la capa 0, siendo los factores de ponderación, $w_{pj1}^{(1)}$, los pesos de las conexiones sinápticas entre las neuronas de ambas capas.

$$v_{ji}^{(1)} = \sum_{p=0}^{m_0} w_{pj1}^{(1)} y_p^{(0)}; j_1 = 1, 2, \dots, m_1 \quad (4.26)$$

Esta entrada neta puede expresarse en función de las señales de entrada a la red, sustituyendo la Ec. (4.25), en la Ec. (4.26), la cual queda como:

$$v_{ji}^{(1)} = \sum_{p=0}^{m_0} w_{pj1}^{(1)} x_p^{(0)} \quad (4.27)$$

Acto seguido, las entradas netas de la primera capa oculta se transforman en señales de salida, $y_{j1}^{(1)}$, mediante la función de activación, f_{j1} , de cada neurona.

$$y_{j1}^{(1)} = f_{ji}(v_{ji}^{(1)}) = f_{ji}\left(\sum_{p=0}^{m_0} w_{pj1}^{(1)} x_p^{(0)}\right) \quad (4.28)$$

A su vez, estas señales se transmiten a la siguiente capa oculta (capa 2), adonde llegan combinadas linealmente para formar las entradas netas, $v_{j_2}^{(2)}$, de sus respectivas neuronas. Los factores de ponderación de estas entradas están constituidos por los pesos de las conexiones entre las neuronas de las capas 1 y 2, $w_{j_1 j_2}^{(2)}$. Por tanto, las señales de salida de las neuronas de la segunda capa oculta, $y_{j_2}^{(2)}$, que se obtienen mediante sus correspondientes funciones de activación, f_{j_2} , se pueden expresar en términos de las señales de entrada a la red, $x_p^{(0)}$.

$$y_{j_2}^{(2)} = f_{j_2} \left(\sum_{p=0}^{m_1} w_{j_1 j_2}^{(2)} f_{j_1} \left(\sum_{p=0}^{m_0} w_{p j_1}^{(1)} x_p^{(0)} \right) \right); j_2 = 1, 2, \dots, m_2 \quad (4.29)$$

Para las siguientes capas ocultas, el proceso es similar a las anteriores, hasta llegar a la última capa de la red (capa L). Las señales salida, $y_r^{(L)}$, están determinadas por las funciones de activación de las neuronas de dicha capa, fr, las cuales se expresan en términos de las funciones de activación de las neuronas de todas las capas ocultas, f_{j_l} (con $l = 1, 2, \dots, L - 1$), y de las señales de entrada a la red, $x_p^{(0)}$.

La ecuación 4.29 puede representarse de forma matricial como:

$$\{y_r^{(L)}\} = F_1 \left(\{x_p^{(0)}\} \right) \quad (4.30)$$

Siendo F_1 la función que comprende todas las operaciones matemáticas involucradas en la RNA. La ecuación 4.30 representa el proceso de predicción de caudales escalados.

La componente determinística multivariada del modelo de generación sintética, con la cual se predice un caudal Z' para un intervalo de tiempo t en varias estaciones j (con $j = 1, 2, \dots, N_s$), en función de N_{qa} valores antecedentes de caudal en dichas estaciones se define como:

$$\{Z'_j\}_t = \left(\{Z'_j\}_\psi \right); \psi = t - N_{qa}, t - N_{qa} + 1, \dots, t - 1 \quad (4.31)$$

4.2.4 Entrenamiento de la red neuronal.

Para el entrenamiento se ha utilizado el algoritmo de retropropagación del error (apartado 4.1.5). La limitación relativa a la velocidad de convergencia de este algoritmo cobra verdadera importancia sólo cuando se tienen tipologías con un gran número de conexiones, caso que es bastante infrecuente en el contexto de esta investigación.

La estimación de los pesos sinápticos, mediante el algoritmo de retropropagación del error, puede verse desde el punto de vista matemático como un problema de optimización no-lineal sin restricciones, para el cual, la función objetivo a utilizar es el error cuadrático medio (MSE).

$$MSE = \frac{1}{2Np} \sum_{n=1}^{Np} \sum_{j=1}^{mL} [e_j(n)]^2 = \frac{1}{2Np} \sum_{n=1}^{Np} \sum_{j=1}^{mL} [d_j(n) - y_j(n)]^2 \quad (4.32)$$

donde:

Np : El número de patrones que se le presentan a la red.

$e_j(n)$: El error de señal de la neurona j para el patrón n .

$d_j(n)$: La señal de salida deseada.

$y_j(n)$: La señal de salida calculada.

El uso generalizado de la función MSE en la mayoría de las aplicaciones de las redes neuronales artificiales radica en sus ventajas: es fácil de calcular, penaliza los errores de mayor magnitud, y su derivada parcial con respecto a los pesos se obtiene de forma sencilla. (Warner y Misra, 1996; Maier y Dandy, 2000).

Para empezar, hace falta darle un valor semilla a los pesos de las conexiones sinápticas. Por lo que se necesita seleccionar un conjunto aleatorio de valores con una distribución de probabilidad uniforme en un intervalo definido. En cualquier caso, lo más recomendable (Maier y Dandy, 2000) es emplear distintos conjuntos de valores aleatorios iniciales, con el fin de ver si los resultados que se obtienen son consistentes. Para la generación de los valores iniciales de los pesos se puede emplear un algoritmo congruencial lineal (Fishman, 1996), como el dado por:

$$\varrho_t = \frac{\Lambda_t}{M(\omega_u - \omega_l)} + \omega_l \text{ con } \Lambda_t = \text{MOD}(b\Lambda_{t-1} + C, M) \quad (4.33)$$

donde:

$$b = 16.807$$

$$C = 0$$

$$M = 2^{31} - 1$$

Seleccionando como semilla Λ_0 , un valor apropiado.

En este trabajo se efectúan las exploraciones iniciales en el intervalo $[-1; 1]$

Una vez establecidos los valores iniciales de los pesos sinápticos, se comienza la fase del entrenamiento propiamente dicha, por lo que se procede a realizar tantos reajustes en la configuración de la red hasta que se encuentre la configuración óptima cuando alcancen algunos de los criterios de parada establecidos en la fase de entrenamiento, que en este estudio son:

- Obtener un MSE igual o inferior a 0,00001.
- Alcanzar 1.000 épocas en el proceso de entrenamiento.

Además, para cada configuración de la red *MLP* se ajustan los valores de la constante de momentum y de la tasa de aprendizaje que se obtienen por ensayo y error.

Por último, las funciones de activación pueden variar para cada capa de procesamiento, y por ello constituyen un recurso más para la búsqueda de la red óptima.

4.2.5 Red neuronal con componente estocástica.

Para poder obtener una red neuronal artificial capaz de predecir caudales mensuales, se necesita incorporar una componente aleatoria dentro de su arquitectura. De este modo, la red neuronal conseguirá generar series sintéticas equiprobables a las series históricas. La adición de esta componente a las redes perceptron multicapa ha sido propuesta por Ochoa-Rivera (2002).

La componente aleatoria en cuestión está dada en función de la parametrización de las series residuales neuronales y de un generador multivariado de ruidos blancos. Las primeras corresponden a las diferencias entre las predicciones de la red determinística (sin componente aleatoria) y los valores observados, y el generador multivariado de sendos blancos es un algoritmo para la obtención de números aleatorios distribuidos normalmente.

Para obtener las series residuales neuronales se pone a trabajar la red ya entrenada en modo predicción a un mes vista, y se evalúan las diferencias entre las predicciones y las observaciones. Posteriormente se comprueba si la media de la serie residual es estadísticamente igual a cero, si se distribuye normalmente, si es independiente en el tiempo y si está correlacionada espacialmente o no con las demás series. Una vez comprobada, se pasa a obtener la media, coeficiente de asimetría, la función de autocorrelación y su matriz de covarianzas y, si se cumplen las anteriores condiciones, se puede expresar la componente aleatoria de la red neuronal para la generación sintética como:

$$\{\epsilon'_j\}_t = B\{\epsilon_j\}_t \quad (4.34)$$

con:

$$BB^T = \Sigma \quad (4.35)$$

Siendo $\{\epsilon'_j\}_t$, las series residuales neuronales sintéticas; $\{\epsilon_j\}_t$ las series de ruidos blancos, de media cero, varianza unitaria e independientes en el tiempo; B, la matriz de parámetros de la componente aleatoria de la red neuronal.

Esta ecuación puede resolverse de igual forma que la correspondiente al modelo AR(1) si se asume que B es matriz triangular inferior y su matriz Σ sea definida positiva o al menos, semidefinida positiva.

Una vez establecidas la componente determinística multivariada Ec. (4.31) y la componente aleatoria del modelo Ec.(4.34) se puede reagrupar la ecuación para incluirlas en una sola expresión y de poder definir el modelo estocástico de red neuronal empleado para la generación sintética: los caudales obtenidos mediante la Ec. (4.31) son variables escaladas que pueden transformarse en variables tipificadas, si se les aplica una función F_2 , que es precisamente la función inversa al escalamiento, esto permite que ambas componentes puedan sumarse, para obtener los caudales sintéticos tipificados.

$$\{Y'\}_t = F_2(\{Z'\}_t) + \{\epsilon'\}_t \quad (4.36)$$

Si los valores de la ecuación anterior se *destipifican* (F_3), se obtienen los caudales sintéticos normalizados.

$$\{X'\}_{vT} = F_3(F_2(\{Z'\}_t) + \{\epsilon'\}_t) \quad (4.37)$$

A los valores obtenidos con esta ecuación, finalmente se les aplica una función F_4 , que comprende las inversas de las funciones empleadas en la fase de normalización, para obtener los caudales sintéticos en la misma escala de los observados. De tal forma que el modelo puede expresarse entonces como:

$$\{Q'\}_t = F(\{Z'\}_t, \{\epsilon'\}_t) \quad (4.38)$$

Donde F incluye a F_2 , F_3 y F_4 .

Quedando así establecida la formulación explícita del modelo estocástico de red neuronal que se emplea para generar las series sintéticas de caudales mensuales.

4.3 Modelación con la red neuronal.

La obtención de series sintéticas hecha en este estudio, empleando el modelo estocástico de las redes MLP, puede resumirse en dos etapas, tal y como aparecen descritas a continuación:

4.3.1 Ajuste del modelo.

Esta fase está constituida por la construcción del modelo, la cual se ha comentado en el apartado anterior. Para poder desarrollar una idea previa de este estudio, puede resumirse el procedimiento realizado de la siguiente forma:

1. **Pre-procesamiento de caudales históricos.** En este primer punto, los valores de los caudales históricos se han normalizado, tipificado y escalado. Además se ha decidido aplicar la transformación de Wilson-Hilferty, después de la tipificación porque se han observado mejoras en los resultados del modelo.
2. **Determinación de la tipología inicial.** Se ha establecido el número de neuronas de las capas de entrada y salida de la red. Esta última tiene tantas neuronas como caudales a predecir (uno por cada estación); el número de neuronas de la capa de entrada es igual al número de caudales antecedentes que determinan los caudales a predecir. Después de que se tiene la tipología inicial definida, se ensamblan los

patrones con los que se entrena la red neuronal, de tal forma que la sección de predicción esté formada por las señales de entrada a las neuronas de la primera capa, y la sección de criterios esté constituida por las señales de salida de la red. Para este caso de estudio se emplearon dos caudales antecedentes, lo que ha permitido comparar los resultados con los proporcionados por el modelo tradicional AR(2).

3. **Entrenamiento de la red.** Esta fase conduce a la determinación de la tipología y de los valores de las conexiones sinápticas. La tipología se hace crecer hasta que se encuentre un valor óptimo. Si bien es cierto que en este estudio se ha buscado la tipología más simple posible que proporcione los mejores resultados posibles, para ello se ha tratado de explorar un gran número de configuraciones posibles, incluyendo la mayoría de los intervalos de tasas de aprendizaje y de la constante momentum.
4. **Estimación de la matriz de parámetros de la componente aleatoria.** Se hace predicción con la red entrenada y los valores obtenidos se comparan con la serie histórica, dando así lugar a las series residuales neuronales observadas. Se verifica que éstas estén normalmente distribuidas, tengan media cero y sean independientes temporalmente. De ser así, se obtiene la matriz de parámetros de la componente aleatoria del modelo estocástico de red neuronal, en función de la matriz de covarianzas de las series residuales neuronales. Si estas series no cumplen las propiedades estadísticas antes señaladas, es probable que la red entrenada pueda mejorarse, es decir, puede que exista una tipología que arroje un menor MSE que el que se consigue con la red obtenida en el entrenamiento. Otra razón por la cual las series residuales neuronales pueden no ajustarse a las propiedades estadísticas indicadas, es la falta de normalización de las series de caudales empleadas en el entrenamiento de la red.

Cuando ya se tiene establecida la configuración de la red neuronal para la predicción de caudales, se procede a su validación.

4.3.2 Validación del modelo.

El modelo una vez ajustado, ya está preparado para la obtención de series sintéticas que reproduzcan los estadísticos relevantes para la gestión de un sistema de recursos hídricos, como son los estadísticos de sequía y almacenamiento. Cuando se trabaja en esta dirección, el proceso de validación, consiste en comparar los estadísticos sintéticos con sus homólogos históricos (Stedinger y Taylor, 1982).

El procedimiento tradicional, en el que la muestra de caudales se divide en tres conjuntos para poder entrenar con el primer conjunto, realizar la validación cruzada con el segundo conjunto y validar finalmente la red con el tercer conjunto, es apropiado cuando la finalidad de la red es realizar tareas de predicción, donde toma especial relevancia la precisión de los valores a predecir. Por el contrario, en la generación de series sintéticas, lo más importante es el comportamiento de sus descriptores estadísticos, resultando más indicado hacer la validación sobre dichos descriptores. Además, de esta forma se puede contrastar el comportamiento del modelo estocástico tradicional con las redes neuronales.

Capítulo 5

Casos de estudio

En este capítulo se muestran los resultados obtenidos al emplear la red neuronal estocástica y los modelos autorregresivos. El objetivo es mostrar las diferencias de rendimiento entre ambos tipos de modelos y su comportamiento al generar la escenarios univariados o bivariados futuros.

Los análisis de resultados se han centrado en la capacidad de preservación de los estadísticos básicos (media, desviación típica, sesgos y coeficiente de correlación) proporcionados por los diferentes modelos y algunos estadísticos relevantes de sequía y almacenamiento.

Se han analizado dos casos de estudio, por un lado, el análisis univariado se ha centrado en la cuenca del Tajo empleando una única serie histórica de aportaciones, mientras que el análisis bivariado se ha realizado en la cuenca del Júcar empleando dos series de aportaciones históricas de igual longitud per separadas geográficamente. El proceso utilizado en ambos casos de estudio ha sido el siguiente:

- Ajuste de un modelo estocástico lineal de uso generalizado (modelos autorregresivos), según la metodología descrita en el apartado 4.2.2. Para ello se aplicaron los programas MODESTE y GENESIS. El primero se ha empleado para la modelación de las series temporales mensuales y el segundo programa se ha utilizado para la generación de las series sintéticas.
- Ajuste del modelo estocástico de red neuronal, de acuerdo con la metodología descrita en el capítulo 4. Los resultados se han obtenido aplicando el programa SERENA.
- Generación de series sintéticas con los modelos analizados para los diferentes casos.
- Comparación de los estadísticos de las series sintéticas generadas mediante las diferentes técnicas empleadas con sus homólogos de las series históricas.
- Aplicación de las series sintéticas bajo condiciones de reducción de aportaciones al sistema reflejadas en las tablas 5.1 y 5.2.

- Comparación de los estadísticos de sequía y almacenamiento (apartado 5.1) a partir de las series sintéticas generadas.

5.1 Consideraciones previas.

Se ha trabajado con distintos umbrales de sequía (tablas 5.1 y 5.2), definidos como porcentaje de las aportaciones medias mensuales históricas para la definición de las sequías históricas del embalse de Bolarque, Alarcón y Contreras. De esta manera, al aplicarle el porcentaje de reducción a considerar a la serie de aportación mensual (tanto histórica como sintética), en caso que la diferencia entre el valor de la aportación mensual respecto el porcentaje de reducción de caudal muestre un valor negativo durante al menos dos meses consecutivos se le ha considerado un episodio de sequía, cuando esta diferencia es positiva se considera situación de disponibilidad.

Tabla 5.1: Umbrales de sequía para Bolarque (Escenario univariado).

Bolarque	Hm ³
Valor medio serie histórica (Qmedio)	26,60
50% Q medio	13,30
60% Q medio	15,96
70% Q medio	18,62
80% Q medio	21,28
90% Q medio	23,94
100% Q medio	26,60

Tabla 5.2: Umbrales de sequía para Alarcón y Contreras. (Escenario bivariado).

Alarcón	Hm ³
Valor medio serie histórica (Qmedio)	33,47
50% Q medio	16,73
60% Q medio	20,08
70% Q medio	23,43
80% Q medio	26,77
90% Q medio	30,12
100% Q medio	33,47

Contreras	Hm ³
Valor medio serie histórica (Qmedio)	29,12
50% Q medio	14,56
60% Q medio	17,47
70% Q medio	20,38
80% Q medio	23,30
90% Q medio	26,21
100% Q medio	29,12

5.1.1 Estadísticos de sequía y almacenamiento.

A continuación se definen de forma breve los estadísticos de almacenamiento y sequía empleados en los casos de estudio de este capítulo.

- *Umbral*: Fracción del caudal medio de toda la serie de caudales históricos, que se emplea para determinar una situación de *sequía* o para determinar la *capacidad de almacenamiento* de un embalse. Un umbral se puede asimilar a una línea horizontal superpuesta sobre el hidrograma de toda la serie de caudales a una altura igual a esa fracción del caudal medio antes mencionada.
- *Déficit*: Cada una de las diferencias entre el umbral y los caudales menores que éste.
- *Sequía*: Sucesión de períodos (meses o años) consecutivos en estado de déficit. Está caracterizada básicamente por tres elementos: *duración*, *intensidad* y *magnitud*, que son propiamente los estadísticos que se calculan dentro de esta categoría.

Los estadísticos de sequía empleados se describen a continuación y pueden observarse en la figura 5.1.

- **Duración**: Número de períodos consecutivos en estado de déficit.
- **Intensidad**: Máximo déficit de todos los que conforman la sequía.
- **Magnitud**: Suma total de los déficits que constituyen la sequía.

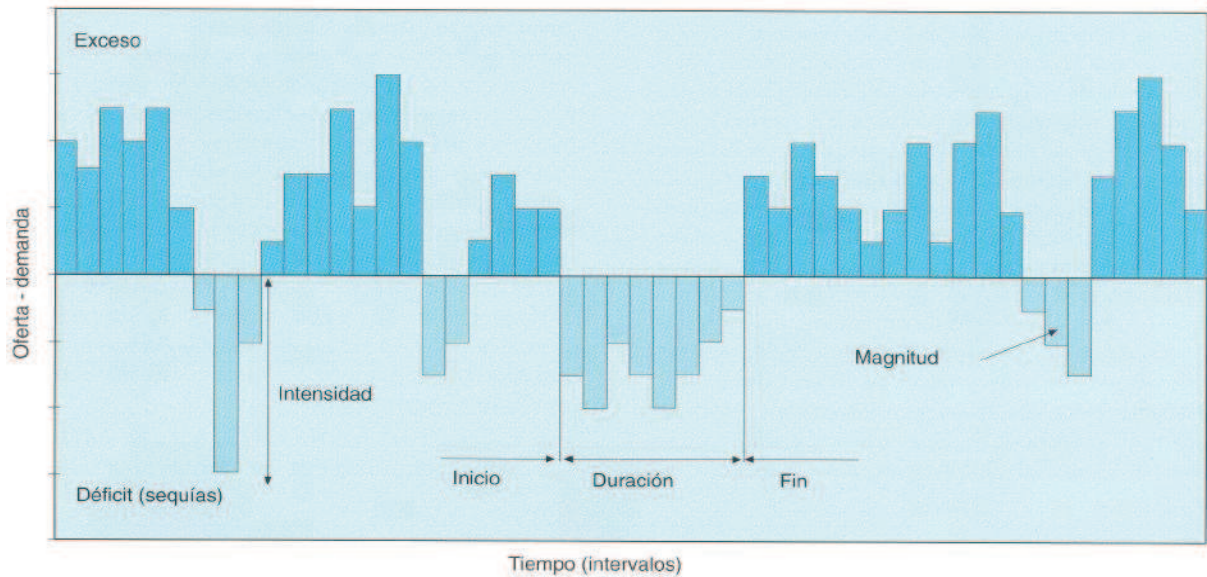


Figura 5.1: Serie tipo para la caracterización de las propiedades de las sequías. Intensidad, magnitud y duración.

Fuente: Fernández 1997

Mediante los programas informáticos GENESIS y SERENA, a las diferentes series de sequías se les ha calculado las medias, máximos y desviaciones típicas de *duración*, *intensidad* y *magnitud*.

$$\bar{\gamma}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \gamma_i ; \gamma_{m\acute{a}x} = M\acute{a}x (\gamma_1, \gamma_2, \dots, \gamma_{n_j}) ; S_{\gamma_j} = \sqrt{\frac{\sum_{i=1}^{n_j} (\gamma_i - \bar{\gamma}_j)^2}{n_j - 1}} \quad (5.1)$$

$$j = 1, 2, 3, \dots, N_s$$

$$n_j > 1$$

donde:

N_s : Número de series sintéticas generadas en cada estación.

n_j : Número de sequías de la j -ésima serie sintética.

Luego, los programas calculan la cantidad media, máxima y mínima de sequías,

$$\bar{n}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} n_i \quad n_{m\acute{a}x_j} = M\acute{a}x (n_1, n_2, \dots, n_{n_j}) \quad n_{m\acute{i}n_j} = M\acute{i}n (n_1, n_2, \dots, n_{n_j}) \quad (5.2)$$

Una vez calculados los estadísticos anteriores, se obtienen los valores promediados de duración, intensidad y magnitud medias y máximas.

$$\bar{\gamma} = \frac{1}{N_s} \sum_{j=1}^{N_s} \bar{\gamma}_j; \quad \overline{\gamma_{m\acute{a}x}} = \frac{1}{N_s} \sum_{j=1}^{N_s} \gamma_{m\acute{a}x j}; \quad \overline{S_\gamma} = \frac{1}{N_s} \sum_{j=1}^{N_s} S_{\gamma_j} \quad (5.3)$$

Finalmente el intervalo de confianza se realiza en base al 95% de probabilidad para cada una de los estadísticos medios Ec. (5.4) y máximos Ec.(5.5).

$$\left[\bar{\gamma} - 1,96 \frac{\overline{S_\gamma}}{\bar{n}}; \bar{\gamma} + 1,96 \frac{\overline{S_\gamma}}{\bar{n}} \right] \quad (5.4)$$

$$\left[\overline{\gamma_{m\acute{a}x}} - 1,96 \frac{S_{\gamma_{m\acute{a}x}}}{N_s}; \overline{\gamma_{m\acute{a}x}} + 1,96 \frac{S_{\gamma_{m\acute{a}x}}}{N_s} \right] \quad (5.5)$$

con:

$$S_{\gamma_{m\acute{a}x}} = \sqrt{\frac{\sum_{j=1}^{N_s} (\gamma_{m\acute{a}x j} - \overline{\gamma_{m\acute{a}x}})^2}{N_s - 1}} \quad (5.6)$$

La *capacidad de embalse* se define como el volumen necesario para almacenar la cantidad de agua requerida por una demanda constante equivalente al valor del umbral. El cálculo de la capacidad de almacenamiento se realiza mediante el algoritmo de los *Picos Secuenciales*, según el cual, dicha capacidad de embalse es el máximo de los valores calculados con la Ec. (5.7).

$$S_t = \begin{cases} U_t - Q_t + S_{t-1} & \text{si } S_t > 0 \\ 0 & \text{si } S_t < 0 \end{cases} \quad \text{para } t = 1, 2, 3, \dots, 2N \quad (5.7)$$

donde:

S_t : Volumen de embalse en el periodo t ($S_0 = 0$).

U_t : Umbral en el periodo t. Es constante y no puede superar el valor del caudal medio de toda la serie.

Q_t : Caudal en el periodo. De N+1 a 2N la serie se repite, con el fin de tener en cuenta aquellos casos en los que la secuencia de caudales críticos (menores que el umbral) están al final de la serie.

N: Número de periodos (meses o años) de la serie.

Por último, aparece el coeficiente de *Hurst*. Este coeficiente se considera como un indicador de la presencia o ausencia de dependencias a largo plazo a largo plazo en series hidrológicas. Este índice permite detectar elementos no periódicos de baja frecuencia (rara ocurrencia) dentro de las series. Este coeficiente (h) se expresa de la siguiente forma:

$$h = \frac{\ln(R^{**})}{\ln\left(\frac{N}{2}\right)} \quad (5.8)$$

donde:

$$R^* = \text{Máx}(S_0, S_1, \dots, S_N) \quad (5.9)$$

$$R^{**} = \frac{R^*}{S_N} \quad (5.10)$$

con:

$$S_t = S_{t-1} + (Q_t - Q_m) \quad t= 1, 2, 3, \dots, N \quad (5.11)$$

donde:

$$S_N = \sqrt{\frac{\sum_{t=1}^N (Q_t - Q_m)^2}{N}} \quad (5.12)$$

siendo:

Q_m : Caudal medio de toda la serie.

R^* : Rango ajustado.

R^{**} : Rango ajustado reescalado.

S_N : Desviación típica de toda la serie.

Cuando el valor resultante de h es mayor de 0,5 indica que existe una persistencia de largo plazo en la serie.

5.2 Caso Univariado. Cabecera del Tajo.

La primera aplicación del modelo de redes neuronales consistió en la generación univariada de caudales sintéticos. Para ello se utilizó la serie de aportaciones del embalse de Bolarque, el cual se sitúa en la cabecera de la cuenca hidrográfica del Tajo. Este sistema presenta una gran importancia por ser el río más largo de la península ibérica, es uno de los ríos más regulados en España y además discurre por dos países, España y Portugal.

5.2.1 Descripción del sistema.

El embalse de Bolarque encuentra situado entre las provincias de Cuenca y Guadalajara en España, ocupando un área total de 7.171 Km². Las precipitaciones medias anuales de esta zona están en torno a 685 mm y su temperatura media se sitúa en 13° C. Desde este sistema parte el trasvase Tajo-Segura permitiendo enviar agua al sureste peninsular.

El embalse de Bolarque fue construido en el año 1910 y tuvo que sufrir numerosas intervenciones y remodelaciones, hasta que en el año 1954 se dio por concluido. Tiene aproximadamente 45 metros de altura, embalsa 30,70 Hm³ y se extiende por una superficie de 490 hectáreas. El promedio histórico de las reservas de este embalse en el periodo 1940 a 1993 ha sido de 26,6 Hm³.

Este embalse junto con los embalses de Entrepeñas y Buendía forma el sistema denominado de Cabecera tal y como puede apreciarse en la figura 5.2.

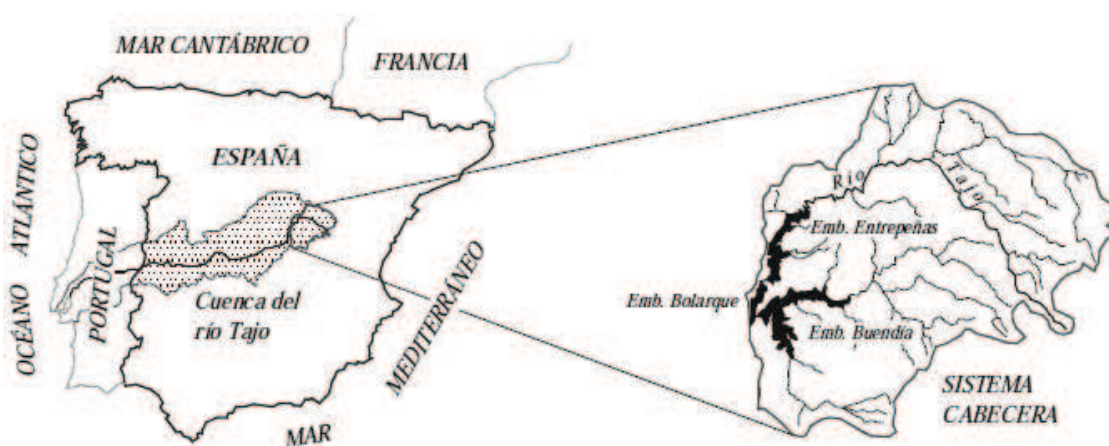


Figura 5.2: Localización del embalse de Bolarque y del sistema de Cabecera en la Cuenca hidrográfica del Tajo.

Fuente: Ochoa-Rivera, 2002

5.2.2 Ajuste del modelo estocástico de red neuronal.

La normalización que se empleó es la indicada en la ecuación (4.15), empleando 12 valores del parámetro c , cada uno de ellos calibrado manualmente por ensayo y error, es decir, se prescindió de la Ec. (4.16). Los caudales normalizados se tipificaron mediante la Ec. (4.24), luego se les aplicó la transformación de Wilson-Hilferty, y finalmente, los valores transformados se escalaron en el rango (0,1).

El entrenamiento de la red empezó comenzando empleando la topología mínima de dos capas con dos neuronas en la primera y una en la segunda. De esta forma se pudieron formar 634 patrones a partir de los 53 años de datos mensuales (1940- 1993). La tipología inicial se hizo crecer mediante la inclusión de una capa de neuronas ocultas, en la que el número de dichas neuronas se iba aumentando en los sucesivos entrenamientos de la red, en la medida en que con ello se redujera significativamente el error MSE, dado en la Ec. (4.32). Durante todo el proceso, se probaron varias funciones de activación, también distintos valores para la tasa de aprendizaje y la constante de momentum. Otras variables que también se modificaron fueron el rango de escalamiento, el número de épocas, y se aplicaron distintos conjuntos de valores iniciales a los pesos sinápticos. Los mejores resultados se obtuvieron para una red MLP 2-1-1, con funciones logísticas de rango (0,1) en las dos capas de procesamiento, con una tasa de aprendizaje de 0.11, y una constante de momentum igual a 0.03; el mínimo error que se obtuvo fue $MSE = 0.155$, manteniéndose más o menos estable a partir de las 1.500 épocas (figura 5.3). El modelo que define la red neuronal propuesta es la siguiente (Ec. 4.36):

$$y_t^{(2)} = \frac{2 \left\{ 1 + \exp \left[-2.195 \left(2 \left\{ 1 + \exp \left[- \left(0.293x_{t-1}^{(0)} + 2.171x_{t-2}^{(0)} - 1.931 \right) \right] - 1 \right) + 1.850 \right] \right\}^{-1}}{-1 + 0.559\xi_t} \quad (5.13)$$

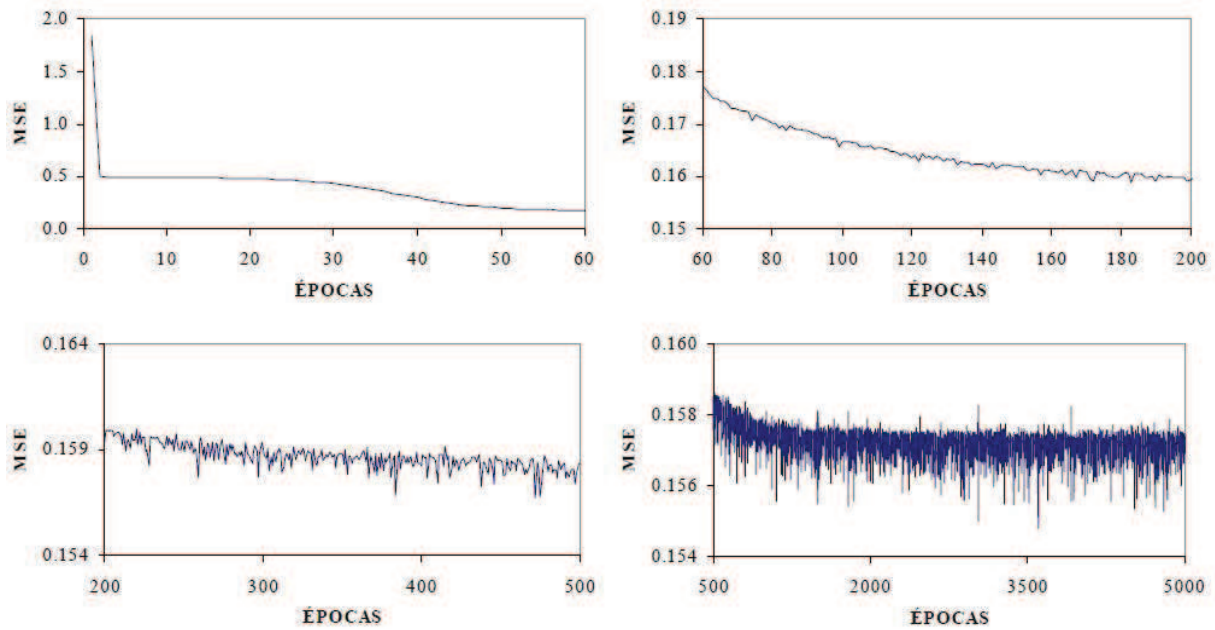


Figura 5.3: Evolución del error cuadrático medio para la RNA 2-1-1.

Una vez entrenada la red neuronal se procedió a su validación siguiendo los pasos indicados en el apartado 4.3.2. Para ello se generaron 200 series de caudales mensuales con una longitud similar a la serie histórica (636 meses). El periodo de calentamiento que se empleó para evitar la influencia de los valores iniciales de caudal en el resto de valores de las series, fue de 120 meses.

5.2.3 Rendimiento de la red neuronal propuesta frente a los modelos autorregresivos.

Debido a que en la estructura de la red neuronal concebida para este caso de estudio cada caudal estaba dado en función de los dos valores antecedentes, el modelo autorregresivo aplicado es de segundo orden, es decir, un modelo AR(2); con este se llevó a cabo un proceso de calibración y validación, tal como se indica en el Capítulo 4. El preprocesamiento de los datos, así como la fase validación, se llevaron a cabo empleando las mismas funciones de normalización y escalamiento de forma exactamente igual a como se procedió en la modelación con la red neuronal, con el propósito de que las comparaciones entre los dos enfoques pudieran establecerse sobre una misma base. Como resultado de la calibración, los valores estimados para los parámetros autorregresivos de este modelo son $\phi_1 = 0,735$ y $\phi_2 = 0,102$, y el valor del parámetro de la componente aleatoria es $\theta_0 = 0,573$.

En la Figura 5.4 aparecen los resultados de los estadísticos básicos, en esta aparecen reflejados la media, desviación típica, coeficiente de sesgo y la función de correlación mensual para 15 desfases temporales; también se presentan los límites de la prueba de independencia temporal de Anderson para un nivel de confianza del 95%. En todos ellos, se observa un mejor ajuste en los resultados de la red neuronal. La figura 5.5 muestra el intervalo de confianza del 95% para los valores de las aportaciones medias mensuales del embalse de Bolarque. El resto de intervalos de confianza se muestran en el anexo.

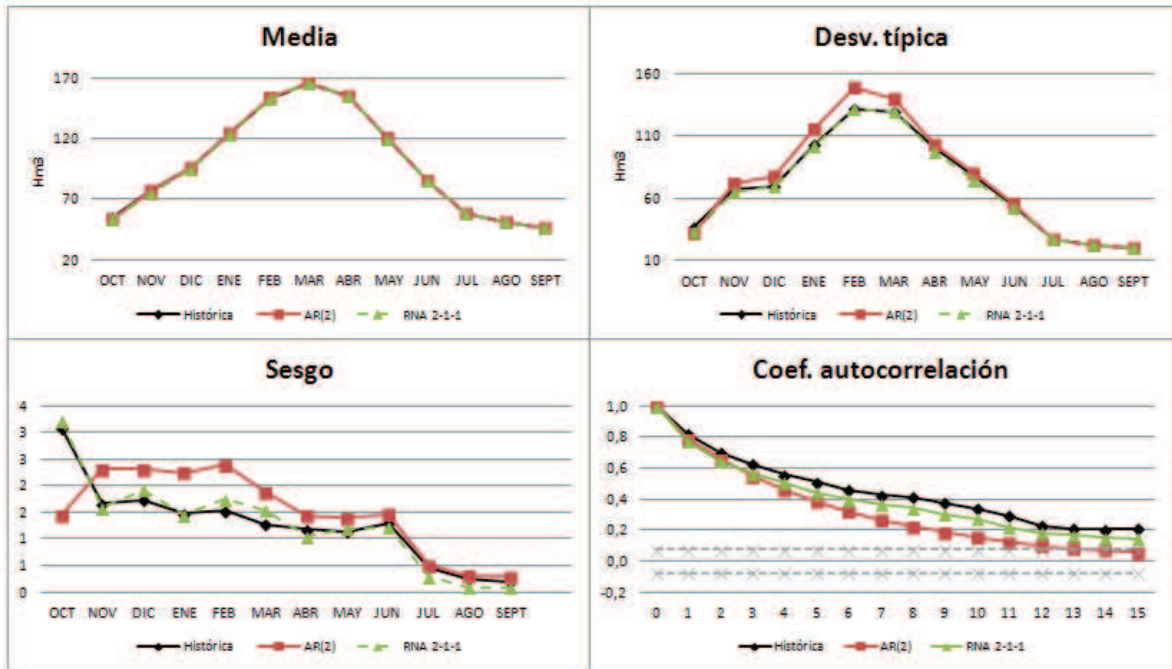


Figura 5.4: Comparación modelos estadísticos básicos del modelo univariado (Bolarque).

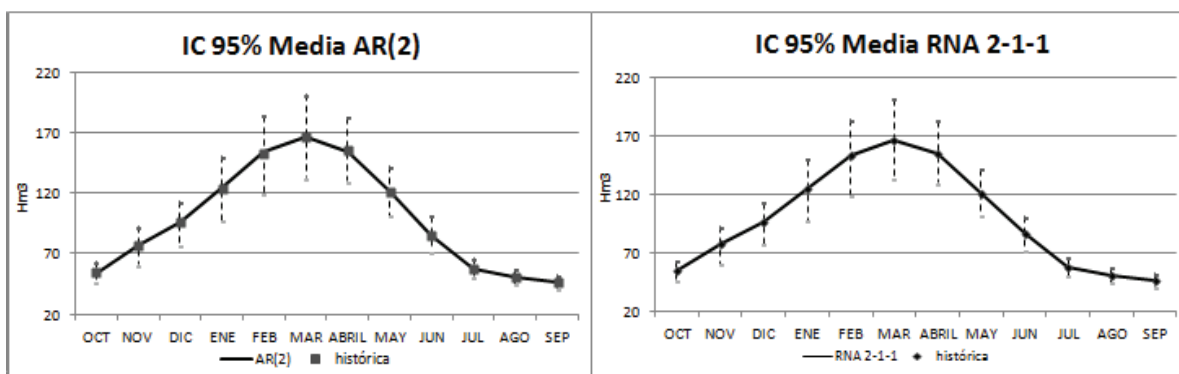


Figura 5.5: Intervalos de confianza del 95% para el valor del caudal medio mensual del modelos univariado (Bolarque).

Para la evaluación numérica de los modelos, se emplearon los índices descritos en el capítulo 3. Los resultados del modelo AR(2) y de la RNA 2-1-1 aparecen mostrados en la tabla 5.3.

Tabla 5.3: Resultados de los índices empleados para la comparación de los diferentes modelos.

<i>AR(2)</i>	r^2	RRMSD	$d'1$	CE	KGE	PI
Media	0,997	0,010	0,995	0,998	0,947	0,997
Desviación típica	0,975	0,111	0,921	0,959	0,798	0,896
Sesgo	0,389	0,521	0,586	0,201	0,929	0,573
Función autocorrelación	0,987	1,097	0,706	0,758	0,651	0,729

<i>RNA 2-1-1</i>	r^2	RRMSD	$d'1$	CE	KGE	PI
Media	0,999	0,010	0,999	0,998	0,916	0,998
Desviación típica	0,998	0,023	0,978	0,997	0,913	0,992
Sesgo	0,977	0,125	0,889	0,957	0,847	0,977
Función autocorrelación	0,996	0,503	0,847	0,938	0,854	0,948

Con los resultados de las figuras 5.4; 5.5 y de la tabla 5.3 se pone de manifiesto el mejor rendimiento de la red neuronal frente a los modelos autorregresivos para la generación de caudales. En todos los índices empleados para el análisis de los resultados se obtienen mejores resultados para la red neuronal (RNA 2-1-1).

Por último, se presentan las frecuencias medias, mínimas y máximas de las sequías estimadas sobre las series anuales sintéticas (anuales), las duraciones y magnitudes de las sequías máximas anuales, las capacidades mínimas de embalse de las series sintéticas anuales y el coeficiente de Hurst.

En la figura 5.6 aparecen los resultados describiendo el porcentaje de error relativo de cada uno de los modelos respecto a los valores de la serie histórica. Este porcentaje de error se define como:

$$\% \text{ Error relativo} = \frac{|generado-observado|}{observado} * 100 \quad (5.14)$$

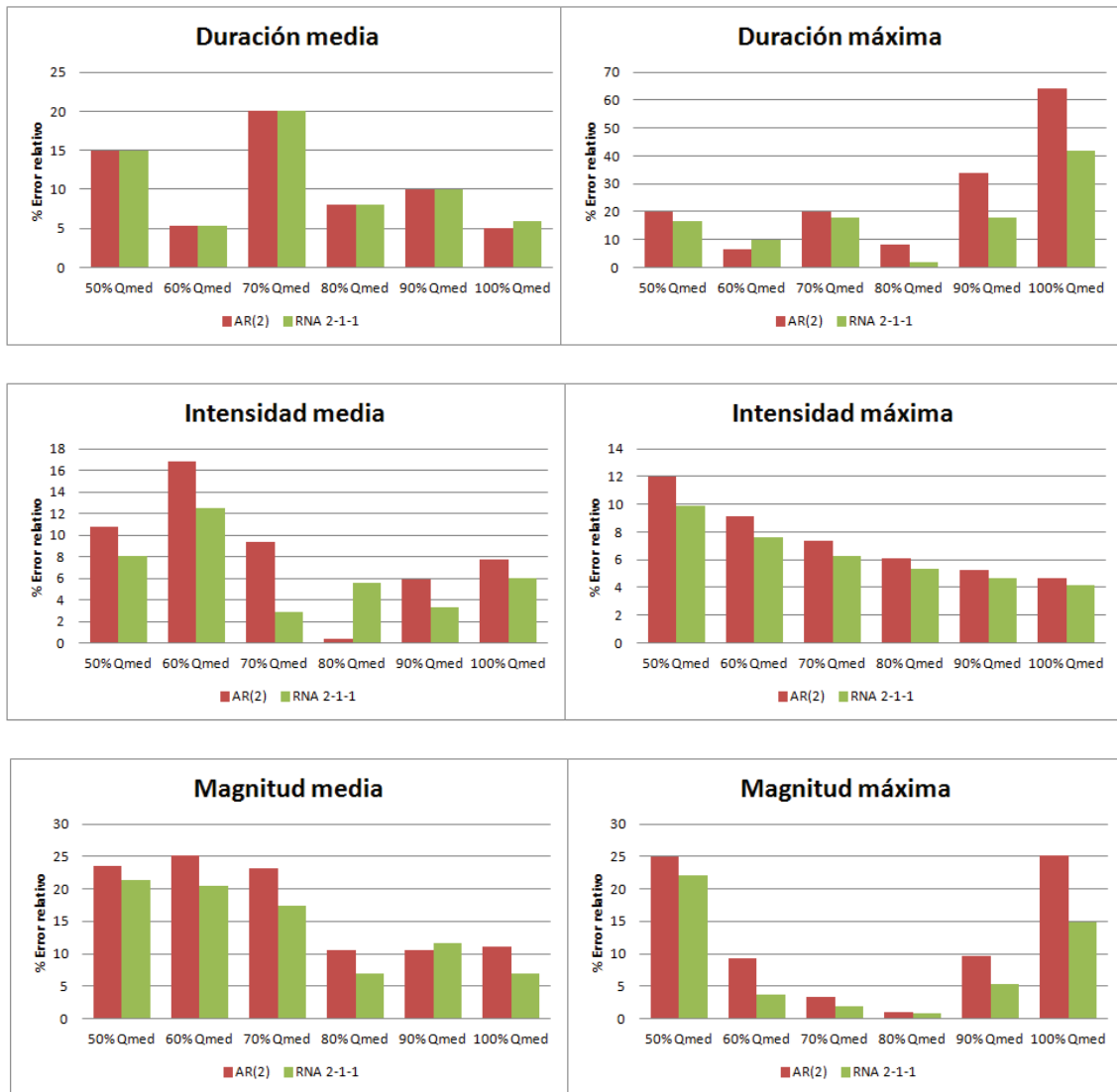


Figura 5.6: Comparación error de predicción de los estadísticos de sequía del modelo univariado (Bolarque).

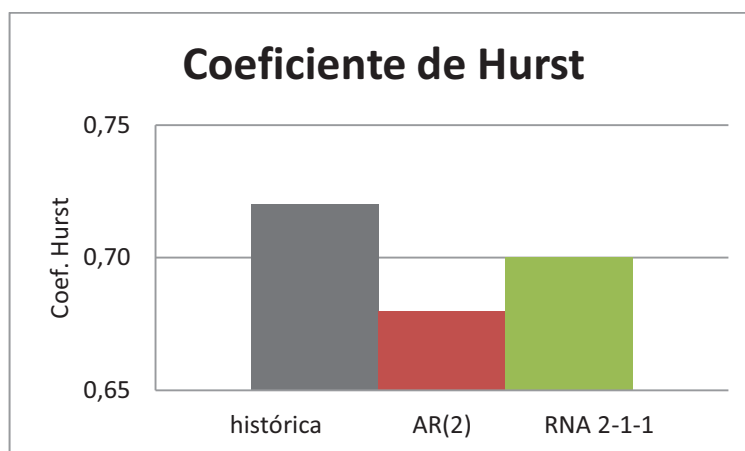


Figura 5.7: Coeficiente de Hurst para el escenario univariado.

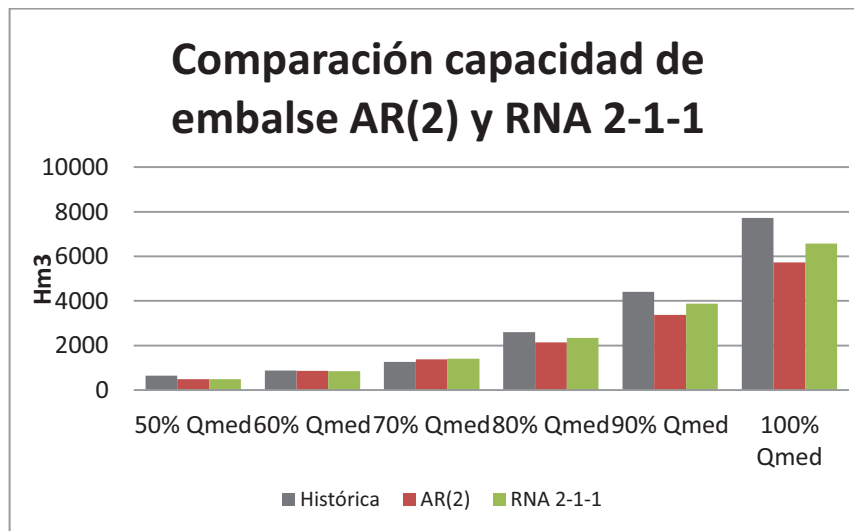


Figura 5.8: Capacidades mínimas de embalse para el escenario univariado.

De las figuras anteriores (5.6) y (5.7), puede deducirse que la red neuronal estocástica presenta un mayor rendimiento que el modelo AR(2); la red neuronal artificial logra una mejor preservación de los estadísticos de las series históricas, especialmente de aquellos relacionados con sus características de sequía a escala anual. Estos resultados son debidos a que se logra una mejor preservación de la estructura de correlación temporal de la serie histórica en la red que en el modelo AR(2), razón por la cual puede decirse que la red constituye un modelo de mayor memoria hidrológica que el modelo AR(2). Característica atribuible al carácter no lineal de la red neuronal. Por último, se muestra la figura 5.7 donde el coeficiente de Hurst en ambos modelos se encuentra por encima de 0,665, indicando que existe persistencia a largo plazo dentro de la serie. El coeficiente de Hurst es más próximo al valor de la serie histórica para la RNA 2-1-1, por esta razón se observa en la figura 5.8 valores de capacidad mínima de embalse más altos que la RNA 2-1-1 para suministrar las aportaciones medias.

La modelación global de los modelos es igualmente parsimoniosa, ya que la relación número de datos/número de parámetros, denominada índice general de parsimonia en parámetros (Salas et al., 1980), es igual a 10 para los dos modelos, tal como se demuestra en la Sección 5.4, en la cual se trata más ampliamente este aspecto.

5.3 Caso bivariado. Cuenca del Júcar.

Una vez obtenidos los resultados del caso univariado, se trató aplicó el mismo tipo de modelo estocástico de red neuronal al caso bivariado, para mostrar su capacidad de generalización. En este caso se estudiaron los embalses de Alarcón y Contreras situados en la cuenca del Júcar. La finalidad era estudiar el comportamiento de las redes neuronales estocásticas en un sistema de recursos hídricos más complejo que el escenario univariado para comprobar la idoneidad de este tipo de modelos.

5.3.1 Descripción del sistema.

Los embalses de Alarcón y Contreras, indicados en la figura 5.9, se sitúan en la provincia de Cuenca el primero y en los límites de la provincia de Cuenca y Valencia el segundo. Alarcón recoge las aguas del primer tramo del río Júcar con una capacidad máxima de 1.118 Hm^3 , mientras que el segundo embalse cuya capacidad máxima se sitúa en 943 Hm^3 hace lo mismo con las aguas del río Cabriel.

La extensión de la parte de la cuenca del Júcar empleada en este estudio es de aproximadamente unos 22.000 Km^2 tal y como aparece en la figura 5.9. La precipitación media anual se sitúa en torno a 510 mm/año y su temperatura media anual es de $13,6^\circ\text{C}$. El promedio histórico de las reservas en el periodo 1940 a 2006 ha sido de $33,46 \text{ Hm}^3$ para el embalse de Alarcón y $29,12 \text{ Hm}^3$ para el caso del embalse de Contreras.

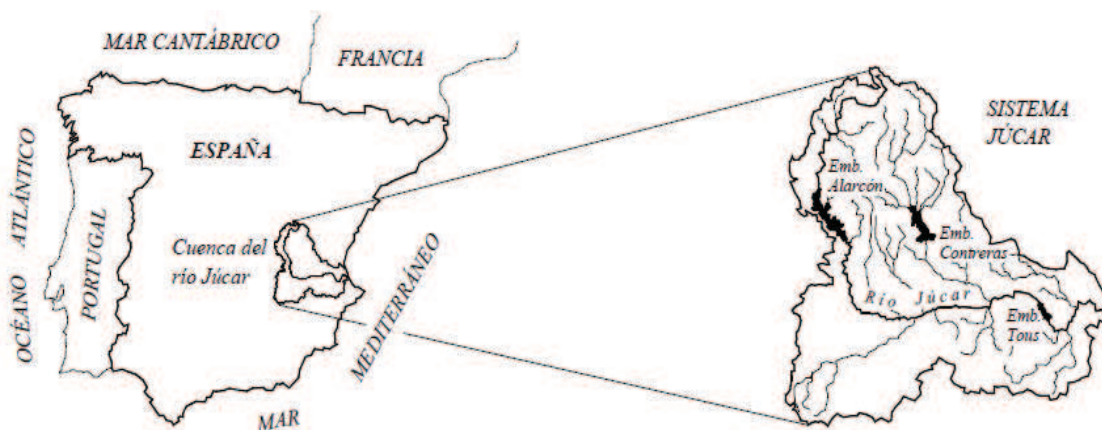


Figura 5.9: Localización de los embalses de Alarcón y Contreras en la Cuenca hidrográfica del Júcar.

Fuente: Ochoa-Rivera.2002

5.3.2 Ajuste del modelo estocástico de red neuronal.

La normalización utilizada para la serie de caudales del embalse de Alarcón es la siguiente:

- Para los meses de octubre, noviembre, diciembre, enero, febrero, marzo, mayo y junio: ecuación (4.12)
- Los meses de abril, julio, agosto y septiembre: ecuación (4.11).

En el embalse de Contreras, las funciones de normalización fueron las siguientes:

- Meses de octubre, noviembre, diciembre, enero, febrero y marzo: ecuación (4.12)
- Los meses de abril, mayo y julio: ecuación (4.11).
- Para los meses de junio, agosto y septiembre no se les aplicó ninguna función de normalización.

Para la normalización de los caudales históricos de las diferentes series utilizadas, se observó la idoneidad de emplear el ajuste de Fourier. Con los resultados observados en la figura 5.10 se decidió comprobar la influencia del número de armónicos con el porcentaje de varianza explicada.

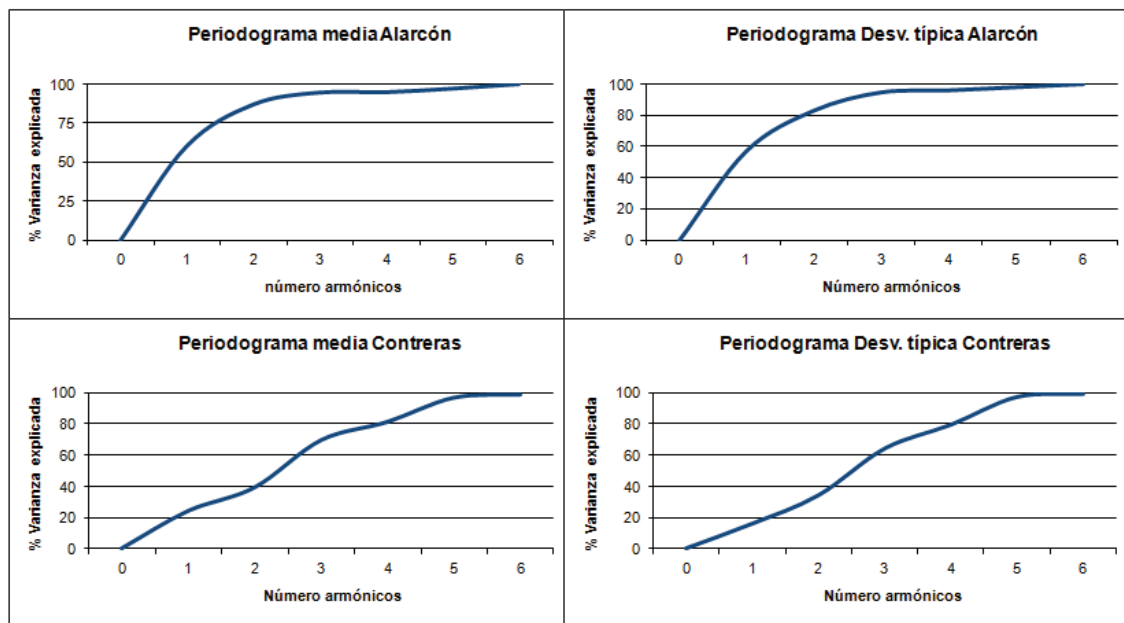


Figura 5.10: Porcentaje de la varianza explicada para la media y desviación típica de las series históricas de Alarcón y Contreras.

En la figura 5.11 se refleja la relación entre el número de armónicos utilizados y el ajuste obtenido para la media y la desviación típica en los modelos AR(1) y

AR(2). La inclusión del modelo AR(1) se realiza para comprobar si la influencia del número de armónicos y el ajuste obtenido se comporta de la misma manera en un modelo autorregresivo de diferente orden. El porcentaje de error relativo aparece definido en la ecuación (5.14).

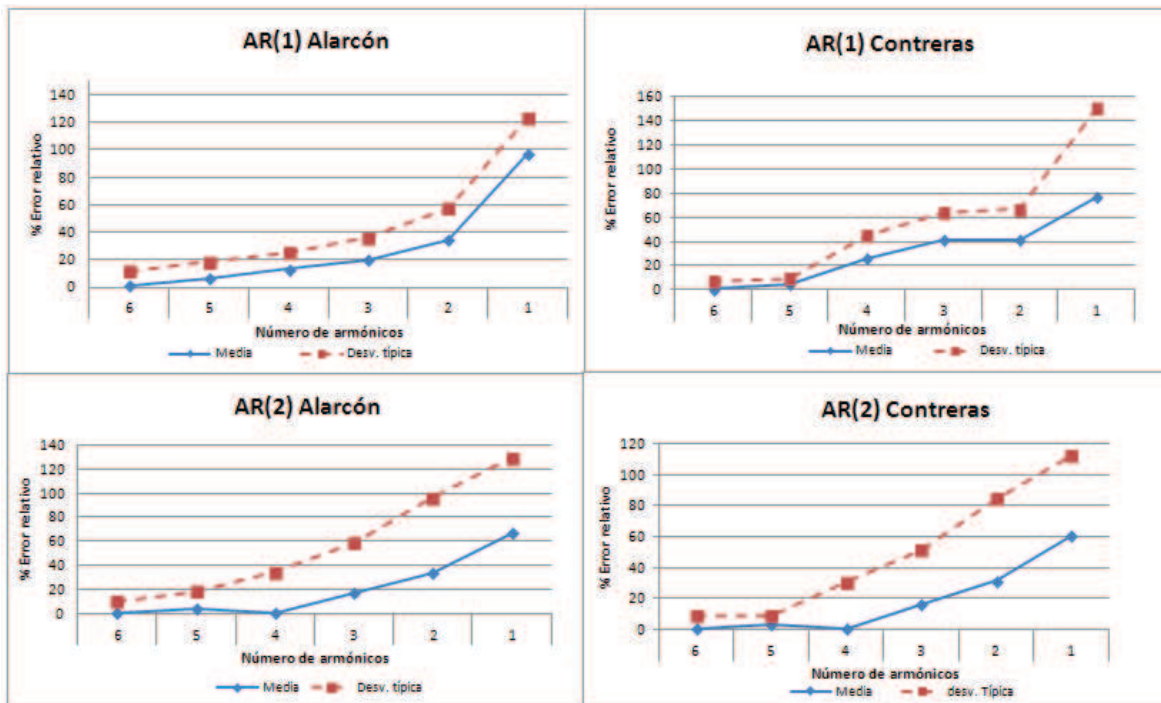


Figura 5.11: Influencia del número de armónicos y su relación con el error asociado. Modelos AR(1) y AR(2) para la media y desviación típica de las series históricas de Alarcón y Contreras.

En la figura 5.11 se observa una ligera mejora en los valores de la media de los modelos AR(2) al emplear cuatro armónicos. Cabe señalar que, cuando la serie de tiempo es demasiado corta, el correlograma puede mostrar variaciones sustanciales en función de la muestra empleada y por tanto, puede ser que oculte su forma real (Ven Te Chow, 1969).

Si se trata de relacionar el número de armónicos utilizados con la media y la desviación típica obtenida se pueden apreciar varios hechos significativos:

- 1- Tanto en los modelos AR(1) y AR(2), conforme decrece el número de armónicos empleados, los errores aumentan.
- 2- En todos los casos, si se emplean cuatro armónicos, los índices muestran que el modelo para el caso de la desviación típica no es adecuado ya que se obtienen resultados negativos (tablas 5.4 y 5.5). De lo que se desprende que serían mejor los valores medios que los generados por el modelo (CE) y además, unos valores negativos del índice de persistencia (PI) indica que el modelo no está captando bien los valores originales.

3- Los armónicos pueden ser tomados como un modelo que representa unas formas regulares u oscilatorias de las variaciones, tales como cambios estacionales que ocurren con frecuencia en los fenómenos hidrológicos (Salas, 1980). Tales variaciones dependen de los valores observados y en el caso de emplearse deberían ser para emplear a lo sumo uno o dos armónicos, hecho que no se corresponde con nuestro caso de estudio.

Por tanto, todas las razones anteriormente citadas justifican el hecho de trabajar con las series originales normalizadas pero sin aplicarles Fourier.

Tabla 5.4: Valores de la media y la desviación típica de Alarcón en función del número de armónicos empleados.

AR(1) media	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,997	0,941	0,996	0,975	0,906	0,996
5 armónicos	0,900	2,810	0,891	0,863	0,866	0,932
4 armónicos	0,645	6,731	0,377	0,716	0,597	0,426
3 armónicos	0,435	11,511	-0,713	0,594	0,191	-0,571
2 armónicos	0,0009	15,250	-2,199	0,379	-0,114	-1,940
1 armónico	0,047	63,690	-54,816	0,193	-5,401	-50,404

AR(1)Desv. típica	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,945	2,956	0,935	0,871	0,835	0,909
5 armónicos	0,805	5,161	0,799	0,773	0,802	0,721
4 armónicos	0,587	9,030	0,386	0,619	0,637	0,146
3 armónicos	0,347	15,620	-0,836	0,571	0,191	-1,552
2 armónicos	0,0004	17,540	-1,317	0,421	0,003	-2,202
1 armónico	0,127	119,421	-106,37	0,189	-8,013	-148,223

AR(2) media	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,997	0,513	0,996	0,978	0,911	0,998
5 armónicos	0,953	1,85	0,953	0,978	0,895	0,956
4 armónicos	0,996	0,491	0,995	0,978	0,914	0,966
3 armónicos	0,458	9,367	-0,262	0,653	0,398	-0,111
2 armónicos	0,137	19,87	-4,444	0,433	-0,589	-4,001
1 armónico	0,060	38,54	-19,44	0,279	-2,502	-17,882

AR(2)Desv. típica	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,943	2,812	0,941	0,884	0,863	0,917
5 armónicos	0,747	6,067	0,724	0,782	0,790	0,615
4 armónicos	0,398	12,131	-1,017	0,635	0,468	-0,538
3 armónicos	0,007	25,775	-3,99	0,489	-0,452	-5,943
2 armónicos	0,008	40,383	-11,278	0,231	-1,526	-16,606
1 armónico	0,001	56,160	-22,743	0,274	-2,756	-31,998

Tabla 5.5: Valores de la media y la desviación típica de Contreras en función del número de armónicos empleados.

AR(1) media	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,995	0,412	0,994	0,974	0,911	0,997
5 armónicos	0,92	1,73	0,907	0,858	0,862	0,951
4 armónicos	0,39	8,800	-1,41	0,471	-0,006	-0,278
3 armónicos	0,102	16,38	-7,35	0,370	-0,986	-3,431
2 armónicos	0,101	16,83	-7,82	0,372	-1,045	-3,680
1 armónico	0,115	38,66	-45,556	0,234	-4,489	-23,697

AR(1)Desv. típica	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,943	0,761	0,931	0,881	0,839	0,921
5 armónicos	0,903	2,391	0,901	0,857	0,861	0,885
4 armónicos	0,290	11,101	-1,146	0,465	0,144	-1,480
3 armónicos	0,167	18,261	-4,803	0,425	-0,605	-5,723
2 armónicos	0,172	19,380	-5,530	0,421	-0,737	-6,571
1 armónico	0,110	59,660	-60,997	0,226	-5,523	-70,782

AR(2) media	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,996	0,37	0,995	0,977	0,909	0,997
5 armónicos	0,945	1,35	0,943	0,899	0,889	0,969
4 armónicos	0,996	0,36	0,995	0,977	0,914	0,997
3 armónicos	0,396	7,23	-0,062	0,598	0,255	0,136
2 armónicos	0,109	15,33	-6,318	0,37	-0,923	-2,888
1 armónico	0,047	29,23	-25,606	0,247	-3,071	-13,113

AR(2)Desv. típica	r^2	RRMSD	CE	d'_1	KGE	PI
6 armónicos	0,939	2,06	0,926	0,871	0,817	0,914
5 armónicos	0,933	2,04	0,927	0,871	0,839	0,915
4 armónicos	0,338	8,08	-0,271	0,604	0,421	-0,317
3 armónicos	0,054	18,72	-4,531	0,445	-0,541	-5,401
2 armónicos	0,006	27,74	-12,398	0,3	-1,651	-14,517
1 armónico	0,0001	38,42	-24,698	0,241	-2,915	-28,768

Una vez resuelto el problema de la normalización de las series, estas se tipificaron mediante la Ec. (4.24) y se les aplicó la transformación de Wilson-Hilferty, y finalmente, los valores transformados se escalaron en el rango (0,1).

El entrenamiento de la red comenzó empleando la topología mínima de dos capas con dos neuronas en la primera y dos en la capa de salida, tratando de mantener en todo momento el principio de parsimonia (la necesidad de dos neuronas en la capa de salida reside en extraer un valor sintético para la serie de Alarcón y otro para la serie de Contreras), para ello se trataba de explorar

todo el espacio existente en la constante de momento y la tasa de aprendizaje (rango 0,1). Cuando se agotaba todo el espacio existente se añadía una neurona más en la segunda capa (capa oculta) y se repetía de nuevo el procedimiento.

Esta primera configuración de red neuronal con dos neuronas en la capa de entrada se desechó rápidamente, dado que al tener dos series históricas diferentes, la red neuronal en cada entrenamiento solo cogía un valor por cada serie, mostrando un comportamiento equivalente a los modelos autorregresivos de orden 1. Es por esta razón que se pasó a emplear cuatro neuronas en la capa de entrada, para poder comparar los resultados con los modelos AR(2).

De nuevo, la topología inicial con cuatro neuronas en la capa de entrada se hizo crecer mediante la inclusión de nuevas neuronas en la capa oculta. Tras los sucesivos entrenamientos de la red se buscó aquella configuración que más redujese el error MSE (Ec. 4.32). Durante todo el proceso se probaron varias funciones de activación, también distintos valores para la tasa de aprendizaje y la constante de momentum.

El mejor resultado se alcanzó para una red MLP 4-4-2, descrita en la figura 5.12, con los detalles que aparecen descritos a continuación:

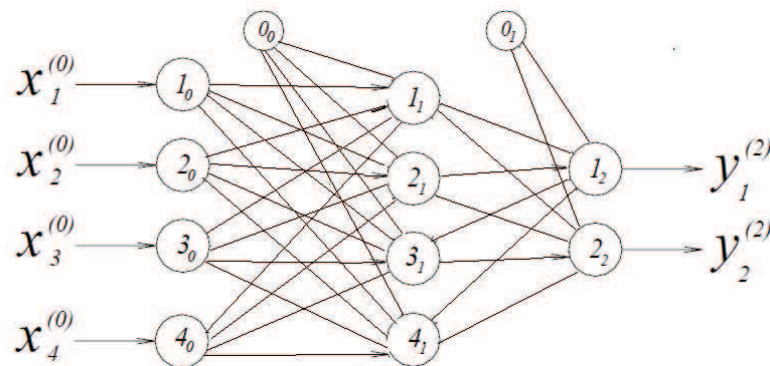


Figura 5.12: Configuración de la red neuronal propuesta.

Descripción de la red neuronal artificial empleada para el análisis bivariado:

- Número de capas ocultas: 1
- Número de neuronas de la capa de entrada: 4
- Número de neuronas en cada capa oculta: 4
- Número de neuronas de la capa de salida: 2
- Funciones de activación de las capas ocultas Sigmoidal [-1; 1]
- Funciones de activación de las capas de salida Sigmoidal [0; 1]
- Parámetros de las funciones de activación :1
- Tasa de aprendizaje (η): 0,15

- Constante de Momentum (α): 0,05
- Número de épocas: 5000
- Error medio cuadrado admisible: 0,00001
- Presentación de patrones: aleatoria.
- Límites inferior y superior de los valores aleatorios iniciales de los pesos [-1.0; 1.0]

El mínimo error que se obtuvo fue $MSE = 0.276$, manteniéndose más o menos estable a partir de las 1000 épocas tal y como se observa en la figura 5.13.

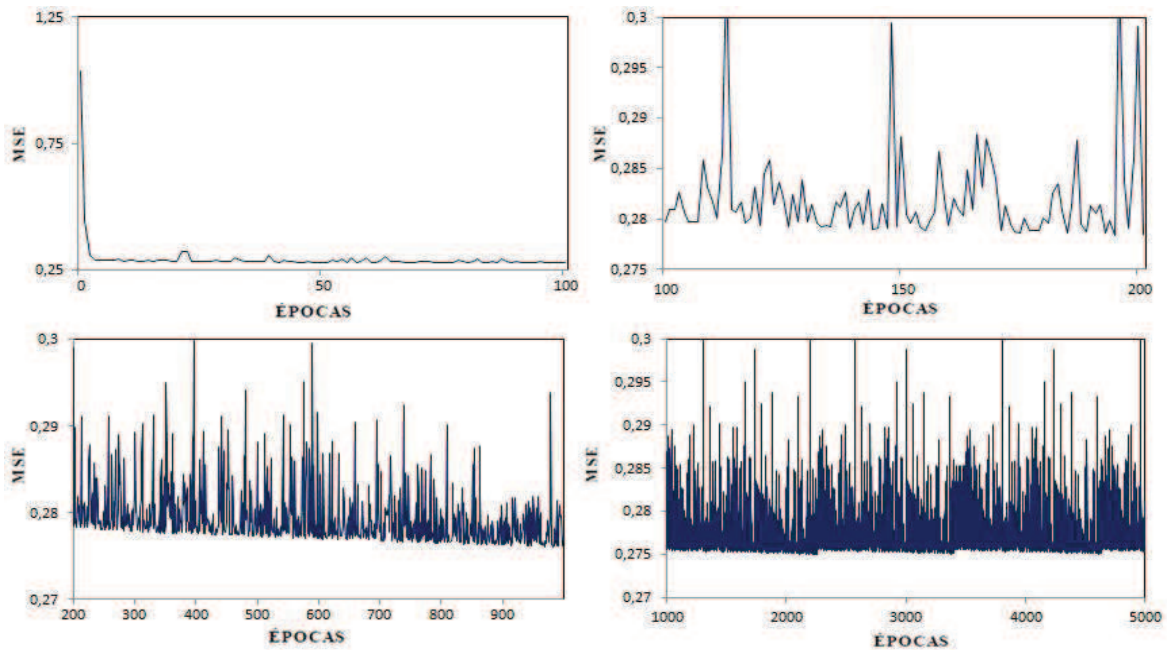


Figura 5.13: Evolución del error cuadrático medio para la RNA 4-4-2.

Los pesos de las conexiones aparecen reflejados en la tabla 5.6:

Tabla 5.6: Pesos de las conexiones de la red neuronal MLP 4-4-2.

Neurona Salida	Capa Salida	Neurona Entrada	Capa de Entrada	Peso de la conexión
0	0	1	1	-0,5441518
0	0	2	1	0,4409833
0	0	3	1	-0,3937535
0	0	4	1	0,1844449
1	0	1	1	-0,034493
1	0	2	1	0,2766614
1	0	3	1	-0,151917
1	0	4	1	0,7313046
2	0	1	1	-0,9628224
2	0	2	1	-0,1567035
2	0	3	1	0,2847438
2	0	4	1	-0,3111911
3	0	1	1	-0,1885152
3	0	2	1	-0,3747363
3	0	3	1	-0,1931462
3	0	4	1	-0,2086673
4	0	1	1	0,9290538
4	0	2	1	0,6069379
4	0	3	1	0,805038
4	0	4	1	0,2732511
0	1	1	2	0,5305095
0	1	2	2	0,2726941
1	1	1	2	-0,8300786
1	1	2	2	0,8689389
2	1	1	2	0,2564769
2	1	2	2	0,6069074
3	1	1	2	0,292129
3	1	2	2	-0,1872287
4	1	1	2	-0,752419
4	1	2	2	0,0939527

A partir de los resultados de la red 4-4-2, se continuó añadiendo neuronas en la capa oculta y nuevas capas sin observarse mejoras significativas, por lo que se descartaron el resto de configuraciones realizadas.

Una vez entrenada la red neuronal, se validó siguiendo los pasos indicados en el apartado 4.3.2. Para ello se generaron 200 series de caudales mensuales con una longitud similar a la serie histórica (792 meses). El periodo de calentamiento que se fue de 120 meses.

5.3.3 Rendimiento de la red neuronal propuesta frente a los modelos autorregresivos.

Al igual que en el caso de estudio univariado, el modelo de red neuronal propuesto es similar al comportamiento de un modelo AR(2); los valores estimados para los parámetros autorregresivos de este modelo son $\phi_1 = 0,660$ y $\phi_2 = 0,148$, y el valor del parámetro de la componente aleatoria es $\theta_0 = 0,524$ para la serie de Alarcón. En el caso de Contreras se obtuvieron parámetros $\phi_1 = 0,352$ y $\phi_2 = 0,027$, y el valor del parámetro de la componente aleatoria es $\theta_0 = 0,471$.

Los resultados obtenidos aparecen en la figuras 5.14 y 5.15. Mientras que en la figura 5.16 se muestra el intervalo de confianza del 95% para el caso de las medias de los embalses de Alarcón y Contreras.

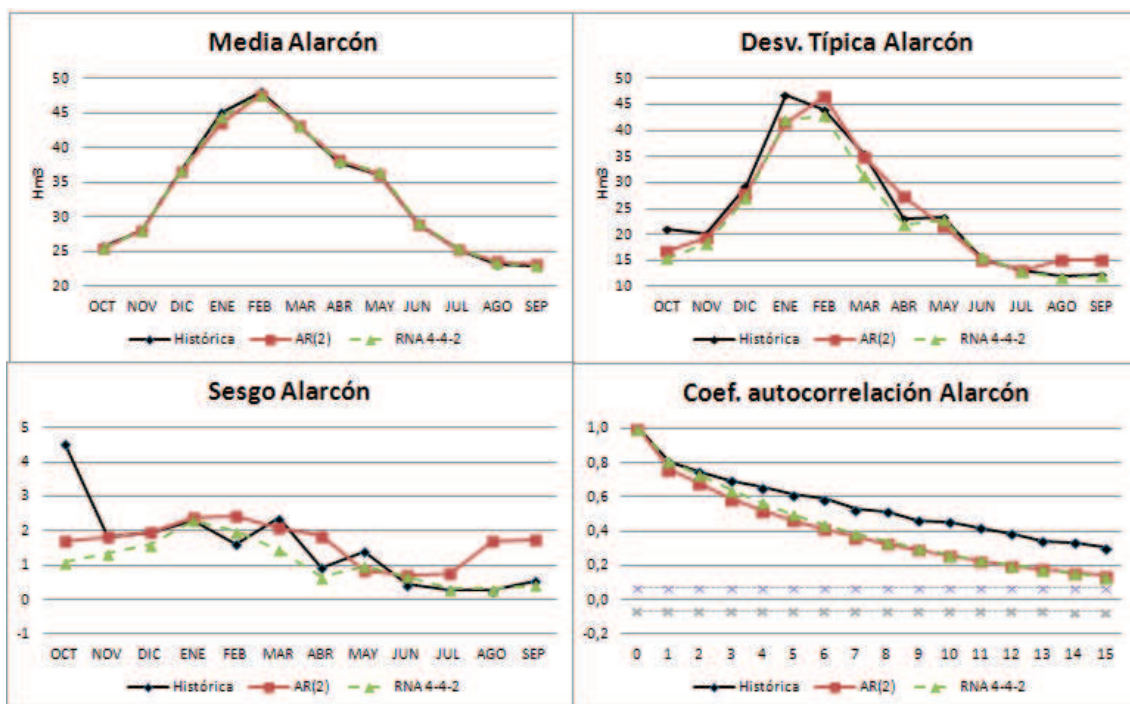


Figura 5.14: Estadísticos básicos para el embalse de Alarcón en el modelo bivariado.

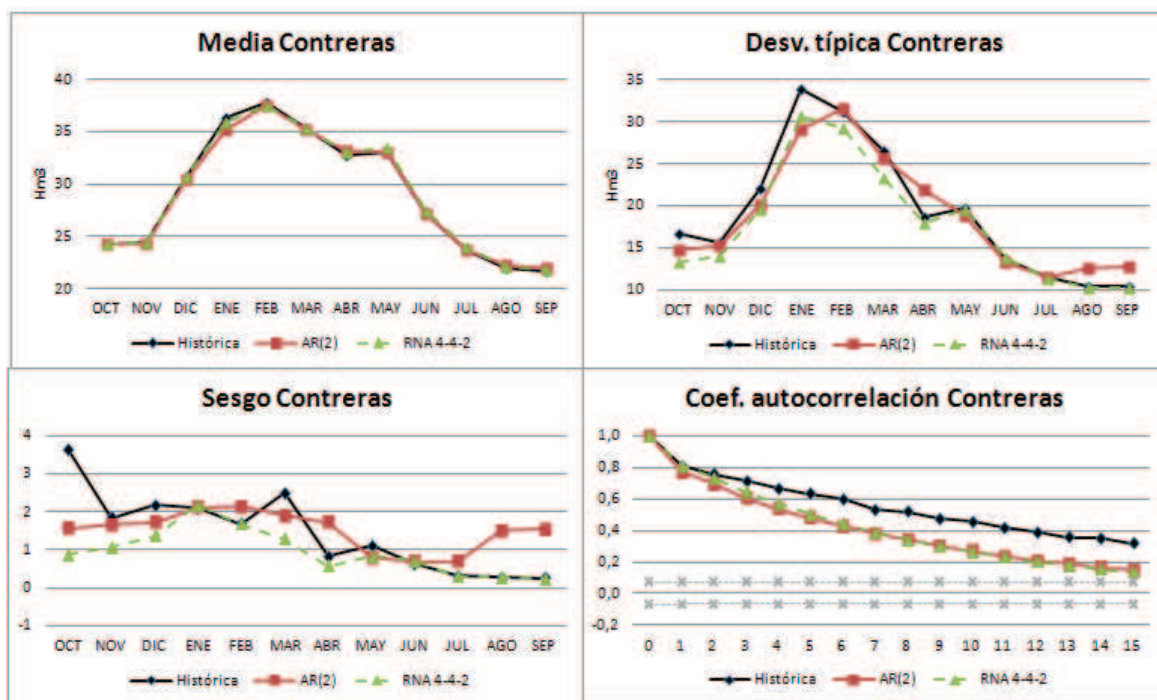


Figura 5.15: Estadísticos básicos para el embalse de Contreras en el modelo bivariado.

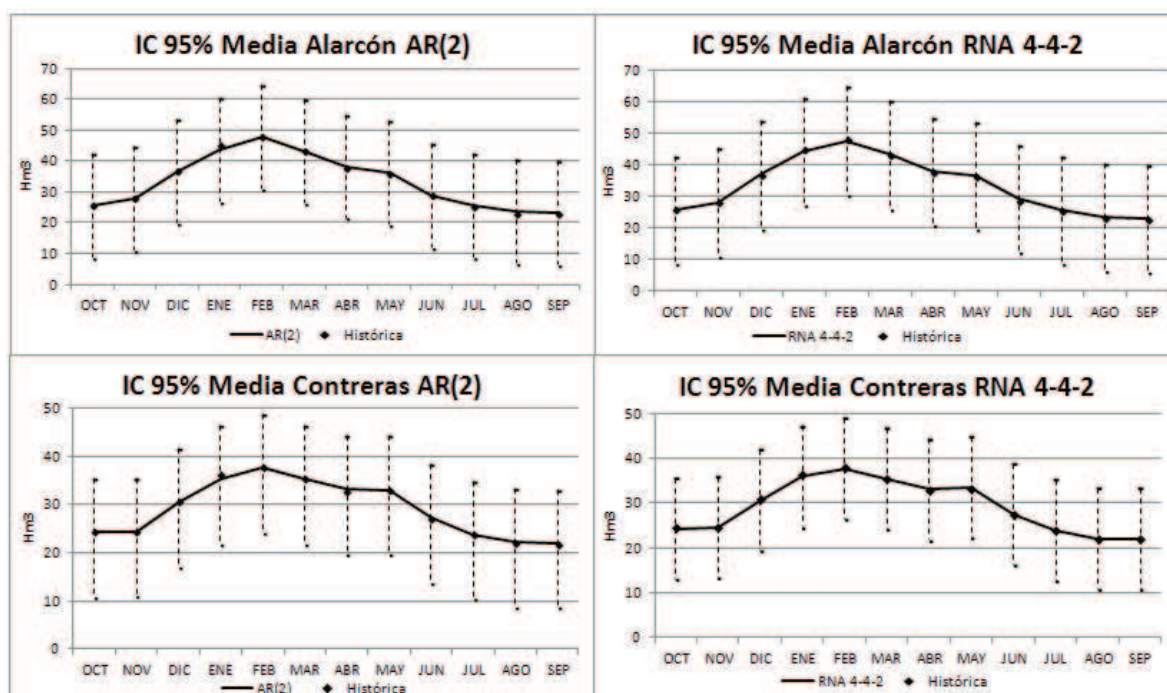


Figura 5.16: Intervalos de confianza del 95% para el valor de los caudales medios mensuales de los embalses de Alarcón y contreras.

Los estadísticos muestran que las series sintéticas generadas por la red neuronal con configuración 4-4-2 se ajustan mejor que en el caso de las AR(2). Los resultados de los índices aparecen en las tablas 5.7 y 5.8:

Tabla 5.7: Resultados de los índices empleados para la comparación de los diferentes modelos Alarcón.

<i>AR(2) Alarcón</i>	r^2	RRMSD	d'1	CE	KGE	PI
Media	0,996	0,500	0,977	0,997	0,910	0,996
Desviación típica	0,941	2,811	0,880	0,941	0,872	0,917
Sesgo	0,161	1,064	0,473	0,177	0,194	0,579
Función autocorrelación	0,935	0,987	0,641	0,109	0,618	0,641

<i>S3LP 4-4-2 Alarcón</i>	r^2	RRMSD	d'1	CE	KGE	PI
Media	0,998	0,300	0,984	0,999	0,914	0,999
Desviación típica	0,946	2,680	0,896	0,977	0,851	0,924
Sesgo	0,161	1,060	0,639	0,312	0,268	0,579
Función autocorrelación	0,935	0,990	0,687	0,109	0,589	0,699

Tabla 5.8: Resultados de los índices empleados para la comparación de los diferentes modelos Contreras.

<i>AR(2) Contreras</i>	r^2	RRMSD	d'1	CE	KGE	PI
Media	0,995	0,402	0,975	0,996	0,907	0,997
Desviación típica	0,923	2,110	0,864	0,934	0,822	0,910
Sesgo	0,242	0,886	0,487	0,245	0,247	0,546
Función autocorrelación	0,960	0,989	0,623	0,525	0,619	0,657

<i>S3LP 4-4-2 Contreras</i>	r^2	RRMSD	d'1	CE	KGE	PI
Media	0,998	0,220	0,983	0,998	0,913	0,999
Desviación típica	0,935	1,931	0,880	0,981	0,835	0,925
Sesgo	0,144	0,941	0,658	0,390	0,298	0,488
Función autocorrelación	0,960	0,992	0,659	0,601	0,576	0,680

Al igual que en el análisis univariado, el coeficiente de sesgo es muy difícil de ajustar en ambos modelos. Esta dificultad se observa bastante bien en las gráficas, donde el mes de octubre es difícil de reproducir por la variabilidad tan fuerte de precipitación que ocurre durante esos meses en las cuencas de ámbito mediterráneo. Si se omite el valor del mes de octubre, el coeficiente de sesgo ve mejorados sus estadísticos.

A partir del gráfico resultante para el coeficiente de sesgo donde se observa que el mes de Octubre presenta una diferencia muy grande respecto al resto de valores y además, los modelos no son capaces de reproducirlo. Se ha tratado de estudiar este fenómeno para poder justificar si esta irregularidad está asociada a la precipitación o si por el contrario, esta se debe a un problema de calentamiento del modelo. Para ello se ha generado una nueva

serie sintética comenzando en el mes de diciembre, los valores de sesgo para cada uno de los embalses del modelo bivariado aparecen en la figura 5.17:

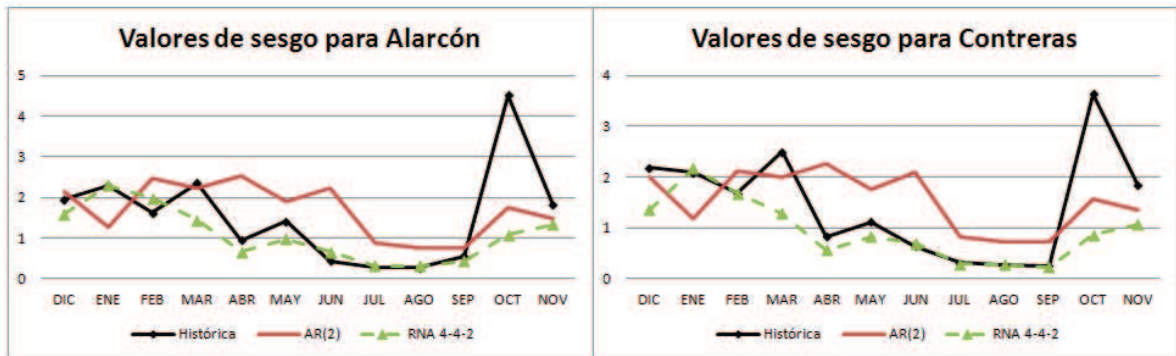


Figura 5.17: Valores de sesgo para los Alarcón y Contreras en el modelo bivariado iniciando la generación de series sintéticas en Diciembre.

Con los resultados mostrados en la figura 5.17, se plasma nuevamente la variabilidad que existe en octubre respecto al resto de meses del año y por tanto, se descarta un problema de calentamiento de los modelos.

Para finalizar con los resultados del análisis bivariado, se presentan las frecuencias medias, mínimas y máximas de las sequías estimadas sobre las series anuales sintéticas (sequías anuales), las duraciones y magnitudes de las sequías máximas anuales, y las capacidades mínimas de embalse de las series sintéticas anuales y el coeficiente de Hurst. En las figuras 5.18 y 5.19 aparecen los resultados describiendo el porcentaje de error relativo de cada uno de los modelos respecto a los valores de la serie histórica, en la figura 5.20 los valores del coeficiente de Hurst para ambos embalses y en la figura 5.21 los valores mínimos de capacidad de embalse.

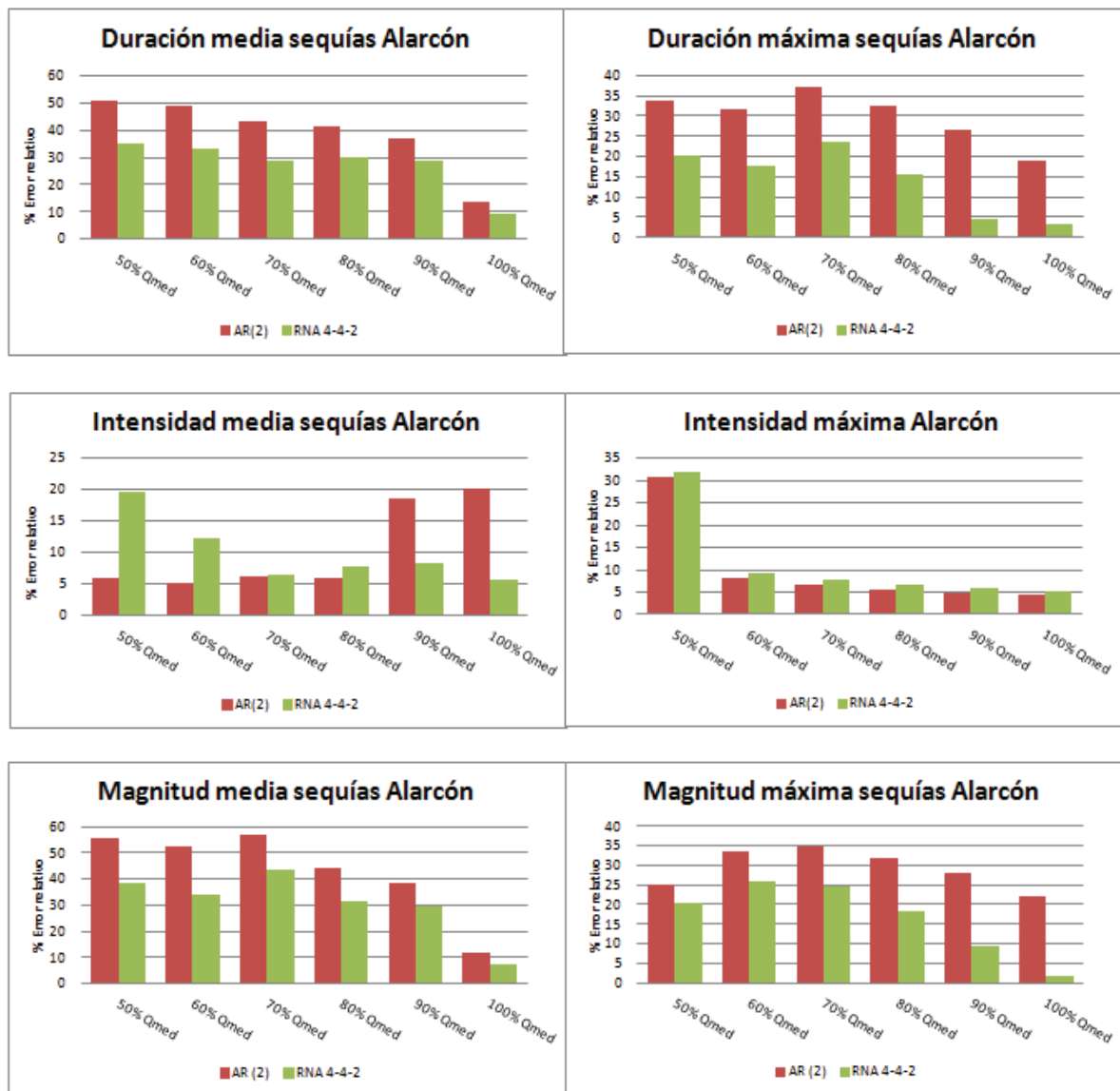


Figura 5.18: Porcentaje de error de predicción de los estadísticos de sequía del modelo bivariado para el embalse de Alarcón.

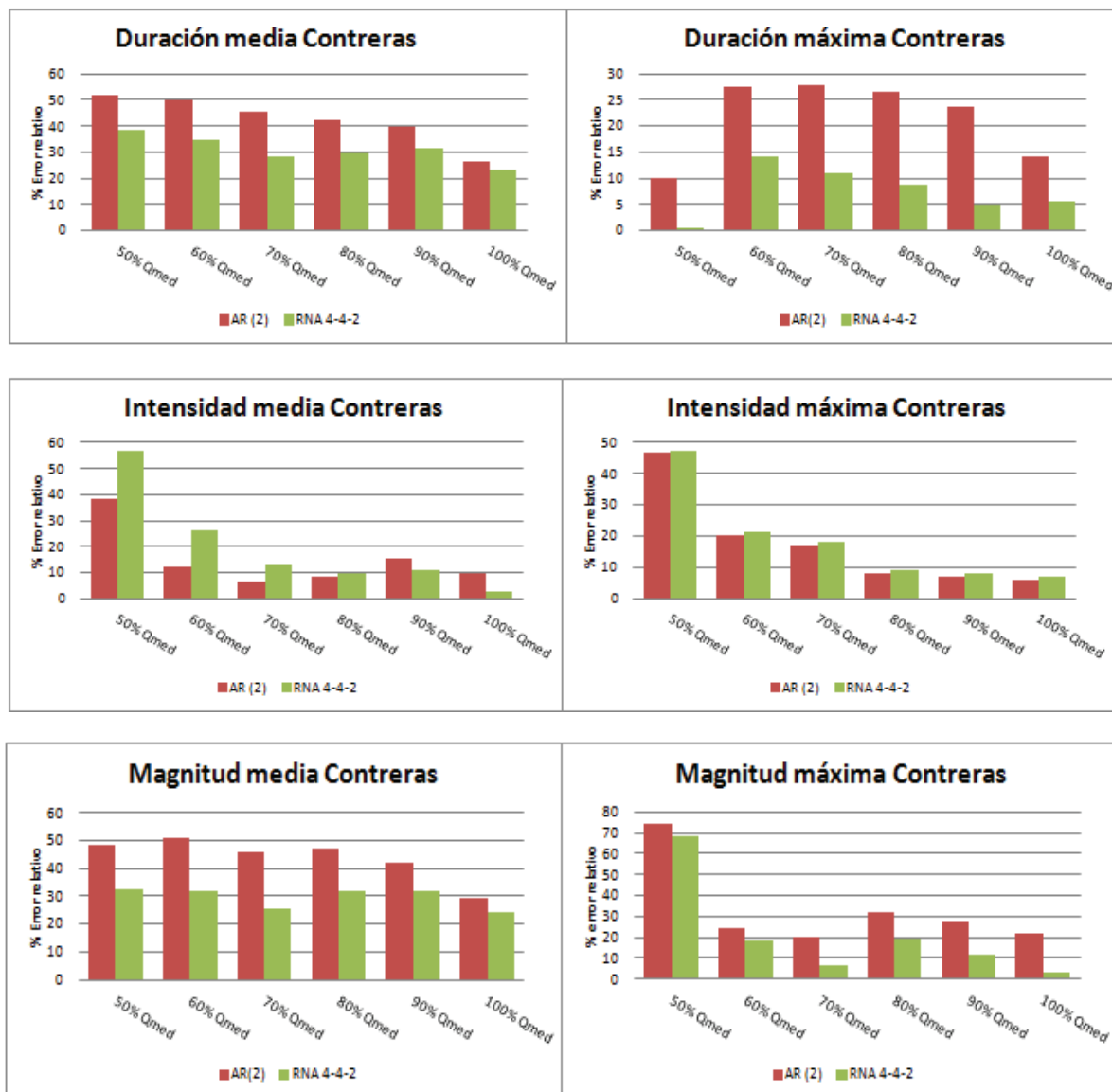


Figura 5.19: Porcentaje de error de predicción de los estadísticos de sequía del modelo bivariado para el embalse de Contreras.

Los resultados de los gráficos de sequía tanto para la serie de Alarcón y Contreras indican, para el caso de la magnitud y la duración de las sequías medias y máximas, que la red neuronal es capaz de reproducirlas mejor, mientras que, para la intensidad media y máxima, la red neuronal reproduce mejor los resultados cuando se trabaja con reducciones de caudal cercanos al caudal medio y, para umbrales menores, presenta mayores errores relativos en comparación con el modelo AR(2). Esta peculiaridad no ocurre en el caso de la predicción de las intensidades máximas, donde los resultados son prácticamente similares a los que se correspondería con la serie histórica salvo en el caso de de trabajar con una reducción de las aportaciones igual a la mitad

del caudal medio histórico donde este error es mayor. Esta anomalía puede deberse a la dificultad de los modelos a la hora de reproducir el sesgo, hecho que dificulta la generación de picos de caudales máximos en las series sintéticas similares a los registrados en la serie histórica.

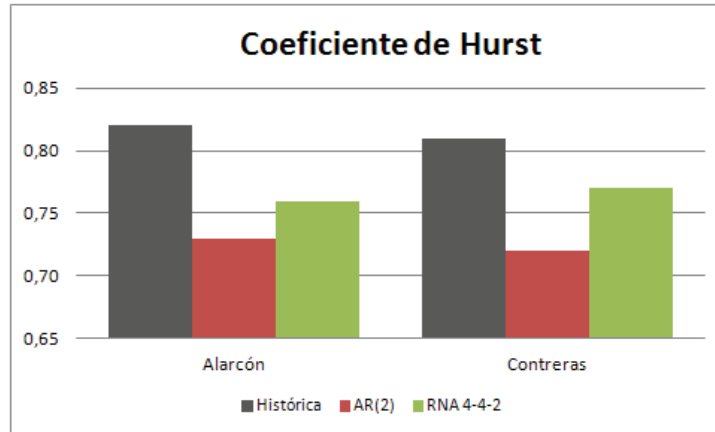


Figura 5.20: Coeficiente de Hurst para el escenario bivariado

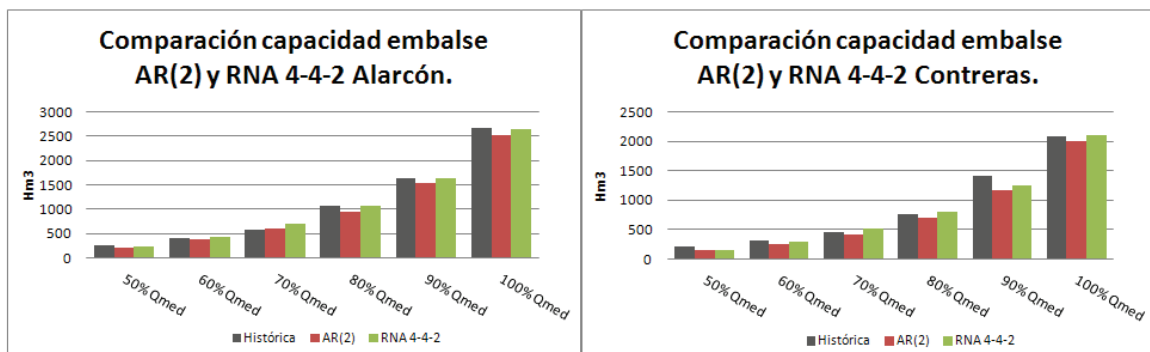


Figura 5.21: Capacidades mínimas de embalse para el escenario bivariado.

Tal como puede observarse en la figura 5.20 los valores del coeficiente de Hurst son en todos los casos superiores a 0,75, poniendo de manifiesto que existe persistencia a largo plazo en las series. Por otro lado, en la figura 5.21 se puede observar que los valores obtenidos a partir del modelo de red neuronal son más altos que para el modelo AR(2), por esta razón, la red neuronal presenta unos valores de capacidad de embalse más próximos a los valores de la serie histórica.

5.4 Parsimonia de los modelos aplicados.

El número de parámetros a emplear en el modelo de generación de series sintéticas depende de forma significativa de los estadísticos de las series históricas que se desee reproducir.

A mayor precisión a la hora de preservar los estadísticos, mayor es el número de parámetros que se requiere y menos generalista resulta el modelo, por el contrario, un modelo con menor número de parámetros puede dar como resultado series significativamente diferentes a las históricas. En general, se adopta el criterio de conservar los estadísticos de primer y segundo orden, como la media y la desviación típica, al igual que la estructura de la correlación temporal, ya que estos descriptores representan una gran parte de la variabilidad natural de las series (Salas et al., 1980). En este trabajo se ha tenido en cuenta el principio de parsimonia, por lo que la configuración de la red ha sido la mínima posible que garantice una mejora significativa en los resultados respecto a los modelos tradicionales.

Para comprobar que se cumple este principio se han empleado dos criterios diferentes.

En primer lugar, el criterio propuesto por Weigend et al. (1990) definido por:

$$N_w \leq \frac{N_p}{10} \quad (5.15)$$

donde:

N_w : Número de pesos sinápticos de la red.

N_p : Número de patrones de entrenamiento.

Este criterio, según Lachtermacher y Fuller (1994), es fuertemente restrictivo y por ello estos autores han presentado otro criterio, adaptado del anterior de la siguiente forma:

$$\frac{1,1P}{10} \leq H(I + 1) < \frac{3P}{10} \quad (5.16)$$

donde:

P : Número de patrones utilizados en el entrenamiento de la red neuronal.

H : Número de neuronas ocultas.

I : Número de neuronas de la capa de entrada.

Atendiendo a la modificación propuesta, Ochoa-Rivera (2002) concluye en su tesis doctoral que la implicación de un límite inferior para el número de pesos de la red es, en cierto modo contradictorio, puesto que, en el estudio que realizó se obtuvieron mejores resultados con configuraciones de red más pequeñas que el límite inferior de pesos a emplear.

$$H(I + 1) < \frac{3P}{10} \quad (5.17)$$

Por tanto, al seguir el modelo propuesto por la tesis doctoral de Ochoa-Rivera (2002), se omitirá el límite inferior de la regla de Lachtermacher y Fuller.

En la tabla 5.9 se indican los resultados de la verificación de las dos reglas de parsimonia, para los casos de estudios analizados en este capítulo:

Tabla 5.9: Verificación de las reglas de parsimonia de Weigend et al. (1); Lachtermacher y Fuller (2) y Ochoa-Rivera (3).

RNA	N _p o P	N _w	N _p /10	H	I	H(I+1)	3P/10	Regla 1	Regla 2	Regla 3
2-1-1	634	5	63,4	1	2	3	190,2	Cumple	Cumple	Cumple
4-4-2	1584	30	158,4	4	4	20	475,2	Cumple	No cumple	Cumple

Puede apreciarse como la tipología de la red muestra que se cumplen los criterios de Weigend y Ochoa-Rivera en ambos modelos. La regla de Lachtermacher y Fuller no se cumple para el caso de la red neuronal del escenario bivariado porque su límite inferior es muy restrictivo para la red propuesta, donde se ha tratado de buscar la misma base para poder comparar la red neuronal y el modelo estocástico AR(2).

$$\text{Índice parsimonia} = \frac{N}{Np} \quad (5.18)$$

Donde:

N: Número de datos.

Np: número total de parámetros.

Las tabla 5.10 y 5.11 muestran los índices de parsimonia para los diferentes modelos autorregresivos empleados para los diferentes casos de estudio:

Tabla 5.10: Índice de parsimonia globales de los modelos aplicados en los casos de estudio del modelo univariado.

	Número total de parámetros	Número de datos	Índice de parsimonia
AR(1)	61	636	10
AR(2)	63	636	10

Tabla 5.11: Índice de parsimonia globales de los modelos aplicados en los casos de estudio del modelo bivariado.

	Número total de parámetros	Número de datos	Índice de parsimonia
<i>AR(1)</i>	100	1584	16
<i>AR(2)</i>	109	1584	15

Con la información anteriormente mostrada, no puede afirmarse desde el punto de vista de la parsimonia, que la red neuronal sea más idónea que el modelo autorregresivo correspondiente, sólo se han mostrado las reglas que se emplean generalmente para estudiar si las redes neuronales son parsimoniosas o no. (Ecuaciones 5.15 a 5.17). En cambio, para los modelos autorregresivos algunos autores (Salas et al., 1980) sugieren que cuando el índice de parsimonia de un modelo tiene valores iguales o superiores a 15, dicho modelo tiene un grado de parametrización adecuado, aunque en muchas ocasiones esta regla es difícil que se cumpla, en especial cuando las series de las que se disponen no son de gran longitud.

Como aportación del presente trabajo, se tratará de aplicar el criterio AIC propuesto por Akaike (1974) para determinar cuál es el modelo más adecuado en el caso de estudio de Alarcón y contreras (modelo bivariado).

Tabla 5.12: Resultados del AIC de los modelos aplicados en los casos de estudio del modelo bivariado.

<i>AIC Alarcón</i>	<i>AR(1)</i>	<i>AR(2)</i>	<i>RNA 4-4-2</i>
Media	0,126	0,135	0,167
Desv. Típica	0,124	0,133	0,165
Sesgo	0,126	0,135	0,165
Coef. Autocorrelación	0,126	0,135	0,165

<i>AIC Contreras</i>	<i>AR(1)</i>	<i>AR(2)</i>	<i>RNA 4-4-2</i>
Media	0,126	0,136	0,167
Desv. Típica	0,124	0,133	0,165
Sesgo	0,126	0,135	0,165
Coef. Autocorrelación	0,126	0,135	0,166

Como puede apreciarse en la tabla 5.12, para el caso de la red neuronal aparecen los mayores valores, hecho que, en caso de emplear objetivamente el criterio Akaike tendríamos que descartarla como modelo idóneo. Pero, en realidad, lo que ocurre es que el criterio AIC no sirve para comparar modelos que emplean bases matemáticas diferentes, es decir, a la hora de comparar un modelo autorregresivo con una red neuronal, esta última por su configuración

va a presentar un mayor número de parámetros y, en consecuencia, va a estar muy penalizada en su resultado final.

Si se aplica el criterio FPE (Akaike, 1969) basándose en la estimación del error cuadrático medio de predicción a un paso descrito en el capítulo 3 (Ec. 3.10) se obtienen los resultados mostrados en la tabla 5.13:

Tabla 5.13: Resultados del FPE de los modelos aplicados en los casos de estudio del modelo bivariado.

FPE Alarcón	AR(1)	AR(2)	RNA 4-4-2
Media	0,759	0,289	0,107
Desv. Típica	13,308	9,129	8,605
Sesgo	1,38	1,299	1,341
Coef. Autocorrelación	1,054	0,421	0,310

FPE Contreras	AR(1)	AR(2)	RNA 4-4-2
Media	0,358	0,184	0,057
Desv. Típica	8,467	5,14	4,463
Sesgo	0,769	0,895	1,058
Coef. Autocorrelación	1,104	0,383	0,349

Con este criterio se obtienen mejores resultados para la red neuronal frente a los modelos autorregresivos. La justificación radica en que su función de pérdida está basada en la diferencia al cuadrado entre los valores observados y generados. De este modo, sí que permite comparar modelos de diferente base matemática porque el utilizar mayor número de parámetros no implica que se obtengan peores resultados, sino que lo que penaliza este índice es el número de caudales mensuales previos (n ; $n-1$; $n-2$;...) empleados. Por tanto, la solución óptima será aquel modelo que proporcione un menor FPE.

En el caso del sesgo, se observa como el sesgo más adecuado no es el correspondiente a la red neuronal, sino que oscila entre el AR(2) y AR(1) debido a la peculiaridad del mes de octubre comentada en el apartado 5.3.3. Lo que dificulta que este comportamiento tan peculiar sea más difícil de captar por la red neuronal artificial. Pese a este inconveniente los valores no difieren en gran medida.

Capítulo 6

Conclusiones y líneas futuras de investigación

El presente estudio surge de la idea de conseguir una mejor preservación de los estadísticos básicos, de sequía y de almacenamiento en la generación de series sintéticas empleando redes neuronales con componente estocástica. Para ello se ha tratado de profundizar en la capacidad de generalización de este tipo de redes empleando diferentes casos de estudio. Aunque la idea no es nueva, ya que este uso de las redes neuronales es propuesto en la tesis doctoral de Ochoa-Rivera (2002), se pretende aportar mayor conocimiento acerca del funcionamiento de las redes neuronales perceptron multicapa con componente estocástica en la generación de series sintéticas. La aplicación de la red S3LP ha sido ilustrada con 2 casos de estudio de generación de series estocásticas de aportaciones mensuales; un caso univariado, en la cabecera del Tajo, y uno bivariado, en el río Júcar (aportaciones en Alarcón y Contreras). La generación de aportaciones es un problema esencial en la hidrología operacional, pues permite trabajar con múltiples escenarios futuros potenciales, equiprobables al histórico. Estos escenarios tienen diferentes aplicaciones, como por ejemplo, proporcionar importantes inputs para el análisis de la operación de embalses o la planificación y gestión de sistemas de recursos hídricos.

Los resultados obtenidos en este trabajo indican que la aplicación de la técnica descrita ha sido satisfactoria para el mapeo de procesos autorregresivos y por tanto, es una técnica adecuada para la modelación de series sintéticas de caudales (u otras variables hidrológicas) y la predicción de caudales futuros. Esto permite abrir nuevas vías y/o herramientas a explorar en este ámbito.

En este capítulo se pretende reflejar las conclusiones, aportaciones y las posibles líneas futuras de investigación derivadas del presente trabajo final de máster.

6.1 Conclusiones generales.

Se ha podido observar el comportamiento de la red neuronal estocástica a la hora de trabajar con dos series de aportaciones (bivariado) frente el análisis con una única aportación (univariado). Los resultados que se han obtenido en dos regiones diferentes nos permiten afirmar la capacidad de generalización de este tipo de redes a la hora de generar escenarios hidrológicos futuros.

En la mayoría de los casos, las redes neuronales se consideran herramientas a las que no se les conoce muy bien su funcionamiento interno. Por ese motivo, ejercicios de este tipo permiten analizar el comportamiento de las redes neuronales en diferentes escenarios de generación hidrológica. El modelo de generación de escenarios hidrológicos futuros a través de las redes neuronales estocásticas empleadas en este estudio es, en general, mejor que los modelos tradicionales. Esta mejora en los resultados es debido, en gran parte a la componente aleatoria del modelo, la cual se estima a partir de las series residuales que resultan de poner a trabajar la red entrenada en modo predicción a un paso de tiempo. Este hecho impide que aparezca la subjetividad y/o manipulación del modelador, puesto que esta serie se genera de forma automática en función de los datos.

A la hora de trabajar con la red neuronal, la validación del modelo se ha llevado a cabo de la misma forma que se procede con los modelos estocásticos tradicionalmente empleados en la síntesis de caudales, es decir, mediante la evaluación de las diferencias entre los estadísticos medios sintéticos y sus homólogos históricos. Pero para la red neuronal estudiada, la adición de un nuevo caudal antecedente en la formulación de la red implica la incorporación de una nueva neurona en su capa de entrada, sin que el proceso de entrenamiento sufra variación alguna. Esto representa otra de las ventajas del modelo aquí propuesto, ya que en los modelos multivariados tradicionales, la inclusión de un nuevo caudal antecedente en su formulación implica cambios importantes en el proceso matemático de la estimación de los parámetros.

Por último, se ha tratado de determinar la influencia del número de armónicos empleado en el análisis de Fourier para la reducción del número global de parámetros del modelo y su relación con el error asociado. Los resultados obtenidos analizando los modelos AR(1) y AR(2) muestran como el error relativo aumenta conforme se emplean un menor número de armónicos. Por esta razón, aunque el ajuste de Fourier es muy utilizado en estudios, cuando se puede trabajar con uno o dos armónicos los resultados finales están más distorsionados.

6.2 Comparación de los modelos empleados.

A la hora de comparar los diferentes modelos presentes en el ámbito de la hidrología no existe un criterio bien estándar, por lo que es fácil encontrar en la bibliografía que cada autor indique unos índices diferentes para presentar sus resultados y en consecuencia, existe una interesante discusión en el ámbito científico.

En este aspecto, en el presente TFM se proponen un conjunto de índices que pueden servir como base para poder comparar diferentes modelos de forma sencilla y comprensible. Por esta razón se ha decidido seleccionar un conjunto de indicadores estadísticos que permitan comparar de forma eficaz los diferentes modelos empleados.

De este modo, se han empleado cinco medidas de bondad de ajuste comúnmente empleadas en la bibliografía: el coeficiente de determinación (r^2); la raíz cuadrada del error cuadrático medio RMSD y su correspondiente valor adimensional RRMSD (Fernández y Salas, 1986); el coeficiente de eficiencia, CE (Nash y Sutcliffe, 1970), el índice de persistencia, PI (Kitanidis y Bras, 1980) y el índice de bondad modificado, $d'1$ (Legates y McCabe, 1999). Adicionalmente, se ha considerado muy útil el índice Kling-Gupta; KGE (Gupta et al., 2009), la razón reside en la necesidad de poder comparar el índice CE. Todos los índices seleccionados son adimensionales (a excepción del RMSD, de ahí la necesidad de aplicar su error relativo, RRMSD) hecho que facilita la comparación de la red neuronal con los modelos clásicos o con otras variantes de la red neuronal. Por último, se ha incluido el Final Prediction Error (FPE), índice propuesto por Akaike (1969), ya que gracias a este índice es posible comparar modelos con diferente base matemática como es el caso de los modelos empleados en este estudio.

Se han obtenido las siguientes conclusiones relativas al uso de los índices mencionados:

- Las redes neuronales artificiales con componente estocástica presenta mejores rendimientos en ambos escenarios. Al hablar de mejor rendimiento se hace referencia a que se ha conseguido una mejor reproducción de los estadísticos más relevantes dentro del ámbito de la gestión de sistemas de recursos hídricos, en especial para situaciones de sequía. Si bien en algunos casos las diferencias de rendimiento entre la red neuronal y los modelos tradicionales no son de gran magnitud, sí resulta significativo el hecho de conseguir mejores resultados, en especial en áreas como es el caso del mediterráneo donde la variabilidad de las precipitaciones es muy importante.

- Los estadísticos más relevantes en los problemas relacionados con la gestión y planificación del agua, como son los estadísticos relativos a las sequías y los de almacenamiento, están fuertemente vinculados a la correlación espacio-temporal y la memoria hidrológica de las series de caudales. El modelo de redes neuronales analizado preserva mejor las propiedades estadísticas básicas de las sequías históricas. Mediante el uso de estadísticos de almacenamiento (capacidad de embalse necesaria, coeficiente de Hurst) se observa que la red neuronal también presenta mejores resultados en la preservación de la persistencia mostrada por los caudales históricos que los modelos clásicos
- Empleando el índice FPE se ha podido comprobar que el modelo de red neuronal utilizado es igual o más parsimonioso que los modelos autorregresivos con los que se ha comparado.

6.3 Líneas futuras de investigación.

Como ya se ha comentado al inicio del capítulo, las redes neuronales constituyen una herramienta muy competente en el estudio de la Hidrología. Las nuevas aportaciones en este ámbito están dirigidas hacia la combinación de estas herramientas junto con otras nuevas. Por esta razón se proponen las siguientes líneas de desarrollo:

- Emplear las redes neuronales con otro tipo de modelos como los de lógica difusa (fuzzy models) o onditas (wavelets) para tratar de analizar la potencialidad de las técnicas híbridas resultantes frente a los modelos clásicos más empleados.
- Mejorar los resultados de las redes neuronales aplicadas en este estudio empleando máquinas vectoriales y diferentes tipos de RNA dentro de un mismo modelo.
- Desarrollar un método de estimación directa de los parámetros de la red neuronal para producir las series sintéticas y los estadísticos de almacenamiento y sequía de forma rápida sin necesidad de explorar todo el espacio estudiado durante la fase de entrenamiento de las mismas (como puede ser el caso de generar un algoritmo de barrido automático de todo el espacio de trabajo para encontrar el óptimo de forma automática en la fase de calibración).

Referencias

- Ackley, D. H., Hinton, G. E. y Sejnowski, T. J. 1985. A learning algorithm for Boltzmann machine, *Cognitive Science* 9. pp 147-169.
- Akaike, H. 1969. Fitting Autoregressive Models for Prediction. *Annals of The Institute of Statistical Mathematics*, 21(2),pp 243–247.
- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19 (6). pp 716–723.
- Anderson, P; Gebeyehu-Tesfaye, Y; Meerschaert, M. 2007. Fourier-PARMA Models and Their Application to River Flows. *Journal of Hydrologic Engineering*. pp 462-472. September/October.
- Andreu, J. 1993. *Análisis de sistemas y modelación, en Conceptos y métodos para la planificación hidrológica*, J. Andreu (ed.), CIMNE, Barcelona, pp. 25-33.
- Andreu, J., Capilla, J. y Sanchís, E. 1996. AQUATOOL, a generalized decision support system for water-resources planning and operational management. *Journal of Hydrology*, 177. pp 269-291.
- ASCE Task Committee on Application of Artificial Neural Networks (Govindaraju, R.). 2000a. Artificial neural networks in hydrology. 1: Preliminary concepts, *Journal of Hydrologic Engineering* 5 (2). pp 115-123.
- ASCE Task Committee on Application of Artificial Neural Networks (Govindaraju, R.). 2000b. Artificial neural networks in hydrology. II: Hydrologic applications, *Journal of Hydrologic Engineering* 5 (2). pp 124-137.
- Bacanli, U.G., Firat, M., Dikbas, E.F. 2009. Adaptive Neuro- Fuzzy Inference System for drought forecasting. *Journal Stochastic Environmental Resources Risk Assesment*. 23.pp 1143-1154.
- Bates, B.C., Kundzewic, Z.W., Wu, S. Palutikof, S. 2008. *Climate change and water*. Technical paper of the IPCC. Geneva.
- Bienenstock, E., Geman, S. 1994 Neural Networks: A Review from a Statistical Perspective, *Statistical Science*, 9. pp 36-38.
- Blenkinsop, S., Fowler, H.J. 2007. Changes in drought characteristics for Europe projected by the PRUDENCE regional climate models. *International Journal of Climatology*. 27. pp 1595-1610.
- Bonifacio, F., Salas, J.D., 1986. Periodic Gamma Autoregressive Processes for Operational Hydrology. *Water resources Research*. 22. (10) pp 1385-1396.
- Box, G. E. P. y Jenkins, G., 1970. *Time series analysis, forecasting and control*, 1st edition, Holden-Day Inc., San Francisco.
- Box, G. E. P. Jenkins G., Reinsel G.C., 1994. *Time series analysis. "Forecasting and Control"*. Prentice Hall. Englewood Cliffs, NJ.

- Bras, R. L. y Rodríguez-Iturbe, I., 1985. *Random functions and hydrology*. Addison-Wesley, Massachusetts.
- Bruins, H. J., 1993. Drought risk and water management in Israel: *planning for the future, en Drought assessment, management and planning: theory and case studies*, D. A. Wilhite (ed.), Kluwer Academic Publishers, Boston, pp. 133-155.
- Chang, T.J., Zheng, H., Kleopa, X.A., Teog C.B. 1996. Develop of an expert system for daily drought monitoring. *Journal Comput. Civil Engineering* 10. pp 20-24.
- Chaturvedi, M. C., 2000. Water for food and rural development: developing countries, *Water International*, 25 (1). pp 40-53.
- Confederación Hidrográfica del Júcar (CHJ), 2001. *Proyecto de implantación en la Confederación Hidrográfica del Júcar de herramientas de decisión en materia de explotación de recursos hídricos*. Área de Explotación, CHJ - Ministerio de Medio Ambiente, Valencia.
- Deo, M. C. y Thirumalaiah, K., 2000. *Real time forecasting using neural networks, en Artificial neural networks in hydrology*, R. S. Govindaraju y A. R. Rao (eds.), Kluwer Academic Publishers, Dordrecht. pp. 53-71.
- Chon, K.H., Cohen, R.J. 1997. Linear and nonlinear ARMA model parameter estimation using an artificial neural network. *IEEE Transactions on Biomedical Engineering* 44 (3). pp168–174.
- Directiva 2000/60/CE del Parlamento Europeo y del Consejo, de 23 de octubre de 2000.
- Dowla, F. U. y Rogers, L. L., 1995. *Solving problems in environmental engineering and geosciences with artificial neural networks*, MIT Press, Cambridge, Massachusetts.
- Dunn, S. M., Mackay, R., Adams, R. y Oglethorpe, D. R., 1996. The hydrological component of the NELUP decision-support system: an appraisal, *Journal of Hydrology* 177. pp 213-235.
- Fernández, B. y Salas, J. D., 1986. Periodic gamma autoregressive processes for operational hydrology, *Water Resources Research*, 22 (10). pp 1385-1396.
- Fernández, B. 1997. Identificación y caracterización de sequías hidrológicas en Chile central. *Ingeniería del Agua*, 4. pp 37-46.
- Fernández B., Vega J.A., Fonturbel T., Jiménez E. 2009. Streamflow drought time series forecasting: a case study in a small watershed in North West Spain. *Journal Stochastic Environmental Resources Risk Assesment*.23. pp 1063-1070.
- Firor, S. E., Finney, B. A., Willis, R. y Dracup, J. A., 1996. Disaggregation modeling process for climatic time series, *Journal of Water Resources Planning and Management*, 122 (3), pp 205-212.
- Fishman, G. S., 1996. *Monte Carlo - Concepts, algorithms, and applications*, Springer-Verlag. New York.

- Gabriel, K.R., Neuman, J., 1962. A Markov chain model for a daily rainfall occurrences at Tel Aviv. *Journal Royal Meteorology Society*. 88 pp 80-95.
- Gupta, H. V., Hsu, K. y Sorooshian, S., 2000. Effective and efficient modeling for streamflow forecasting. *Artificial neural networks in hydrology*, R. S. Govindaraju y A. R. Rao (eds.), Kluwer Academic Publishers, Dordrecht. pp 7-22.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F. 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *Journal of Hidrology* 337 (1-2). pp 80-91.
- Haykin, S., 1999. *Neural networks. A comprehensive foundation*, Prentice-Hall, New Jersey.
- Hill T., Marquez L., O'Connor M., Remus W. 1994. Artificial neural network models for forecasting and decision making. *International Journal of Forecasting* 10 (1). Pp5-15.
- Hornik, K., Stinchcombe, M. y White, H., 1989. Multilayer feedforward networks are universal aproximators, *Neural Networks* 2. pp 359-366.
- Imrie, C. E., Durucan, S. y Korre, A., 2000. River flow prediction using artificial neural networks: generalisation beyond the calibration range, *Journal of Hydrology* 233. pp138-153.
- Jackson, B. B., 1975. The use of streamflow models in planning, *Water Resources Research* 11 (1), pp 54-63.
- Kim, T., Valdes J.B, Yoo, C. 2003. Nonparametric approach for estimating return periods of droughts in arid regions. *Journal Hydrology Engineering*. 8. pp 237-246.
- Kirpatrick, S., Gelatt, Jr., C. D. y Vecchi, M. P., Optimization by simulated annealing, *Science* 220. pp 671-680. 1983.
- Kitanidis, P. K. & Bras, R. L. 1980. Adaptive filtering through detection of isolated transient errors in rainfall-runoff models. *Water. Resources Research*. 16(4), pp. 740-748.
- Kitanidis, P. K. & Bras, R. L. 1980. Real time forecasting with a conceptual hydrologic model, 1. Analysis of uncertainty. *Water. Resources Research*. 16(6), pp.1025-1033.
- Kitanidis, P. K. & Bras, R. L. 1980. Real lime forecasting with a conceptual hydrologic model, 2. Applications and results. *Water. Resources Research*. 16(6), pp.1034-1044.
- Kohonen, T., 1988. An introduction to neural computing, *Neural Networks*, 1. pp 3-16.
- Koutsoyiannis, D., 2000. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36 (6), 1519- 1533.
- Lane, W. L., 1979. *Applied stochastic techniques* (LAST computer package), user manual, Division of Planning Technical Services, Bureau of Reclamation, Denver, Colorado.

- Lapedes, A. y Farber, R., 1988. How neural net works, en *Neural information processing systems*, D. Z. Anderson (ed.), American Institute of Physics, New York. pp. 442- 456.
- Legates, D.R., Mc Cabe GJ., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*. 35. pp 233-241.
- Lehtokangas, M., Saarinen, J. y Kaski, K., 1996. A network of autoregressive processing units for time series modeling, *Applied Mathematics and Computation* 75. pp 151-165.
- Lettenmaier, D. P., 1993. Some thoughts about the state-of-the-art in stochastic hydrology and streamflow forecasting, *Proceedings of the NATO Advanced Study Institute on stochastic hydrology and its use in water resources systems simulation and optimization*, J. B. Marco et al. (eds.), Kluwer Academic Publishers, Dordrecht, pp. 209-215.
- Linsley, R. K., Kohler, M. A. y Paulus, J. C. H., 1975. *Hydrology for engineers*, McGraw-Hill, New York.
- Liu, W.T., Negrón-Juarez, R.I. 2001. ENSO drought onset prediction prediction In north-east Brazil using NDVI. *International Journal Remote Sensor*.17 . pp 3483-3501.
- Lohani, V.K. Loganathan, G.V., 1997. An early warning system for drought management using the Palmer drought index. *Journal American Water Resources Association*. 33 (6).pp 1375- 1386.
- Lohani, A.K. Kumar, R., Singh, R.D. 2012. Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autorregressive techniques. *Journal of Hydrology*, 442-443. Pp 23-25.
- Loucks, D. P., Stedinger, J. R. y Haith, D. A., 1981. *Water resources system planning and analysis*, Prentice Hall, New Jersey.
- Loucks, D. P., 2000. Sustainable water resources management, *Water International*, 25 (1). pp3-10.
- Maier, H.R., Graeme, C.D., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*.15.
- Maier, H. R. y 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. pp101-124.
- Mandelbrot, B. B. y Van Ness, J. W., 1968. Fractional brownian motions, fractional noises and applications, *SIAM Review*, 10 (4). pp 422-437.
- Mandelbrot, B. B. y Wallis, J. R., 1968. Noah, Joseph and Operational Hydrology, *Water Resources Research*, 4 (5). pp909-918.

- Mc Cuen, R. LL., Snyder W. M., 1975. A proposed index for computing hydrographs. *Water resources Research*. 11. pp 1021-1024.
- McCulloch, W. y Pitts, W., 1943. A logical calculus of the ideas immanent nervous activity, *Bulletin of Mathematical Biophysics*, 5.pp115-133.
- Mejía, J. M. y Rousselle, J., 1976. Disaggregation models in hydrology revisited, *Water Resources Research*, 12 (2). Pp 185-186.
- Menéndez, M. y Aguilera, L., 1996. Country paper of Spain, *Proceedings of management and preventions of crisis situations: floods, droughts and institutional aspects, Third technical review*, Euraqua, Roma, Oct. 23-25. pp. 171-180.
- Ministerio de Medio Ambiente, 1998. Libro blanco del agua en España, Madrid.
- Minsky, M. L. y Papert, S. A., 1969. *Perceptrons*, MIT Press, Cambridge, Massacusetts.
- Mishra, A.K., Desai, V.R., 2005. Drought forecasting using stochastic models. *Journal Stochastic Environmental Resources Risk Assesment*.19. pp 326-339.
- Mishra, A.K., Desai, V.R., 2006. Drought forecasting using feed forward recursive neural network. *International journal on Ecological Modelling*.198. pp 127-138.
- Mishra, A.K., Singh V.P. 2010. A review of droughts concepts. *Journal of Hydrology*. 391 (1-2). pp 202-216.
- Moreira, E.E., Paulo, A., Pererira, L.S., Mexia, J.T., 2006. Analysis of SPI drought class transitions using loglinear models. *Journal Hydrology*, 331. pp 349-359.
- Morid, S., Smakhtin, V., Bagherzadek, K., 2007. Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology*. 27 (15). 2103.
- Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part I. A discussion of principles. *Journal of Hydrology* 10 (3). pp 282–290.
- Ochoa-Rivera, J.C. 2002. Modelo estocástico de redes neuronales para la síntesis de caudales aplicados a la gestión probabilística de sequías. Tesis Doctoral. Universitat Politècnica de València.
- Ochoa-Rivera, J.C., Andreu, J., García-Bartual, R. 2007. Influence of Inflows Modeling on Management Simulation of Water Resources System. *Journal of water Resources planning and Management*. pp 106-116.
- Ochoa-Rivera, J.C. 2008. Prospecting droughts with stochastic artificial neural networks. *Journal of Hidrology* 352. pp 174-180.
- Organización meteorológica mundial. (OMM). 1971. Congreso Meteorológico Mundial 6a sesión. Ginebra, Suiza.
- Pallotino, S., Secchi, G.M., Zuddas P., 2005. A DSS for water resources management under uncertainty by scenario analysis. *Environ. Model Softw*. 20 (8). Pp 1031-1042.

- Paulo, A., Ferreira, E., Coelho, C., Pereira L.S., 2005. Drought class transition analysis through Markov and Loglinear models, an approach to early warning. *Agriculture water management* 77. pp 59-81.
- Raman, H. y Sunilkumar, N., 1995. Multivariate modelling of water resources time series using artificial neural networks, *Journal of Hydrological Sciences* 40. pp145-163.
- Reimold,R.J. 1998. *Watershed management practice, policies and coordination*. Mc Graw-Hill. New York.
- Reitsma, R. F., Zagona, E. A., Chapra, S. C. y Strzepeck, K. M., 1996. Decision support systems (DSS) for water resources management, en *Water resources handbook*, L. W. Mays (ed.). McGraw-Hill, New York.
- Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological Review* 95. pp 386-408.
- Rumelhart, D. E., Hinton, G. E. y Williams R. J., 1986. *Learning internal representations by error propagation*, Vol. 1, MIT Press, Cambridge, Massachusetts. pp 318-362.
- Salas, J. D., Delleur, J. W., Yevjevich, V. y Lane, W. L., 1980. *Applied modeling of hydrologic time series*. Water Resources Publications, Littleton, Colorado.
- Salas, J. D., Markus, M. y Tokar, A. S., 2000. Streamflow forecasting based on artificial neural networks, en *Artificial neural networks in hydrology*, R. S. Govindaraju y A. R. Rao (eds.), Kluwer Academic Publishers. Dordrecht. pp 23-51.
- Sánchez, S. T., Andreu, J. y Solera, A., 2001. *Gestión de recursos hídricos con decisiones basadas en estimación del riesgo*, Editorial Universitat Politècnica de València, Valencia.
- Sarle, W.S. (1994): *Neural Networks and Statistical Models*, Proceedings of the Nineteenth Annual SAS Users Group International Conference, April 1994.
- Simonovic, S. P., 2000. *Tools for water management: one view of the future*. *Water International*, 25 (1). pp 76-88.
- Siscoe, G.L., 1978. Solar-terrestrial influences on weather and climate. *Nature* 276. pp 1-5.
- Smith, M.B., Seo, D.-J., Koren, V.I., Reed, S.M., Zhang, Z., Duan, Q., Moreda, F., Cong, S. 2004. The distributed model intercomparison project (DMIP): motivation and experimental design. *Journal of Hydrology* 298. pp 4–26.
- Stedinger, J. R. y Taylor, M. R., 1982. Synthetic streamflow generation. 1. Model verification and validation, *Water Resources Research* 18 (4). pp 909-918.
- Stern, H. S., 1996. Neural networks in applied statistics, *Technometrics* 38 (3), pp 205-214.
- Tadesse, T., Wilhite, D.A., Harms, S.K., Hayes, M.J., Goddard, S., 2004. Drought monitoring using data mining techniques: a case study for Nebraska, USA. *Natural Hazards* 33. pp 1-21.

- Thomas, H. A. y Fiering, M. B., 1962. Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, *en Design of water resources systems*, A. Mass et al. (eds.), Harvard University Press, Cambridge, Massachusetts.
- Ula, T. A., 1992. Generation of multivariate autoregressive sequences with emphasis on initial values, *Journal of Hydrology*, 40. pp 209-233.
- Universitat Politècnica de València (UPV). 2000a. Generación de escenarios hidrológicos futuros. Programa MODESTE (Modelación estocástica de series temporales) – Manual de usuario, Departamento de Ingeniería Hidráulica y Medio Ambiente - UPV, Valencia.
- Universitat Politècnica de València (UPV). 2000b. Generación de escenarios hidrológicos futuros. Programa GENESIS (Generación estocástica de series sintéticas) – Manual de usuario, Departamento de Ingeniería Hidráulica y Medio Ambiente - UPV, Valencia.
- Universitat Politècnica de València (UPV). 2001. SERENA Software para el entrenamiento de redes neuronales artificiales - Manual, Departamento de Ingeniería Hidráulica y Medio Ambiente - UPV, Valencia.
- UNESCO.1999. Working group M.IV. *Sustainability criteria for water resources systems*, Cambridge University Press, Cambridge, UK.
- Valencia, D. y Schaake, Jr., J. C., 1973. Disaggregation processes in stochastic hydrology, *Water Resources Research*, 9 (3). pp 580-585.
- Von der Malsburg, C., 1973. *Self-organization of orientation sensitive cells in the striate cortex*, *Kybernetik*, 14. pp 85-100.
- Wagener, T., Lees, M.J., Wheeler, H.S. 2002. *A toolkit for the development and application of parsimonious hydrological models*. In: Singh, V.P., Frevert, D.K. (Eds.), *Mathematical Models of Large Watershed Hydrology* pp. 91–140.
- Werick, W. J. y Whipple, Jr., J. (eds.), 1994. *National study of water management during drought. Managing water for drought*, IWR Report 94-NDS-8, U. S. Army Corps of Engineers, Sept.
- Westphal, K. S., Laramie R. L., Borgatti, D., Stoops, R. 2007. Drought management planning with economic risk factors. *Journal of Water Resources Planning Management*. 133 (4). pp 351-362.
- Wilcox, B.P., Rawls, W.J., Brakensick, D.L., Wight, J.R. 1990. Prediction runoff from rangeland catchments: a comparison for two models. *Water Resources Research*. 26. pp 2401-2410.
- Wilhite, D. A., 1993. The enigma of drought, *en Drought assessment, management and planning: theory and case studies*, D. A. Wilhite (ed.), Kluwer Academic Publishers, Boston, pp. 3-15.
- Willmott, C.J., Wicks, D.E. 1980. An empirical method for the spatial interpolation of monthly precipitation within California. *Physical Geography* 1. pp 59-73.

Willmott, C.J. 1981. On the validation of models. *Physical Geography* 2. pp 184-194.

Willmott, C.J., Robeson, S.C., Matsuura, K. 2011. Short Communication: A refined index of model performance. *International Journal of Climatology*. Published online.

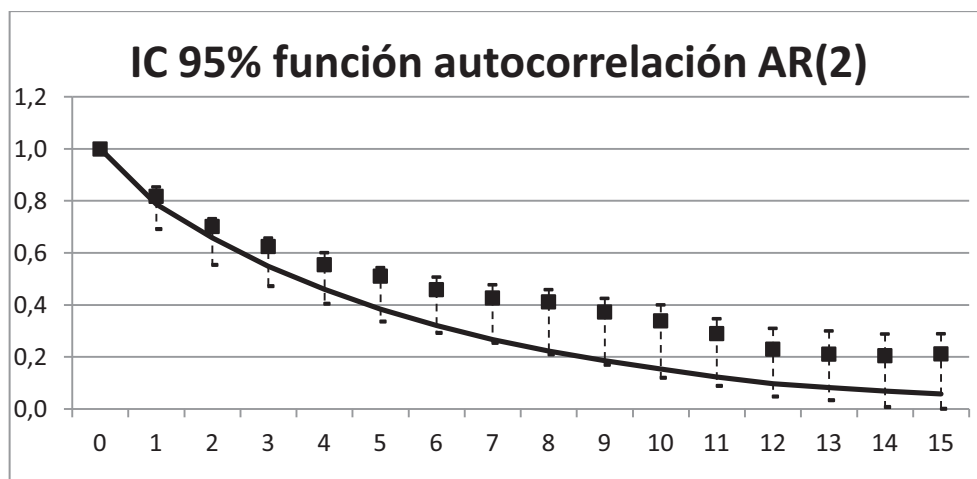
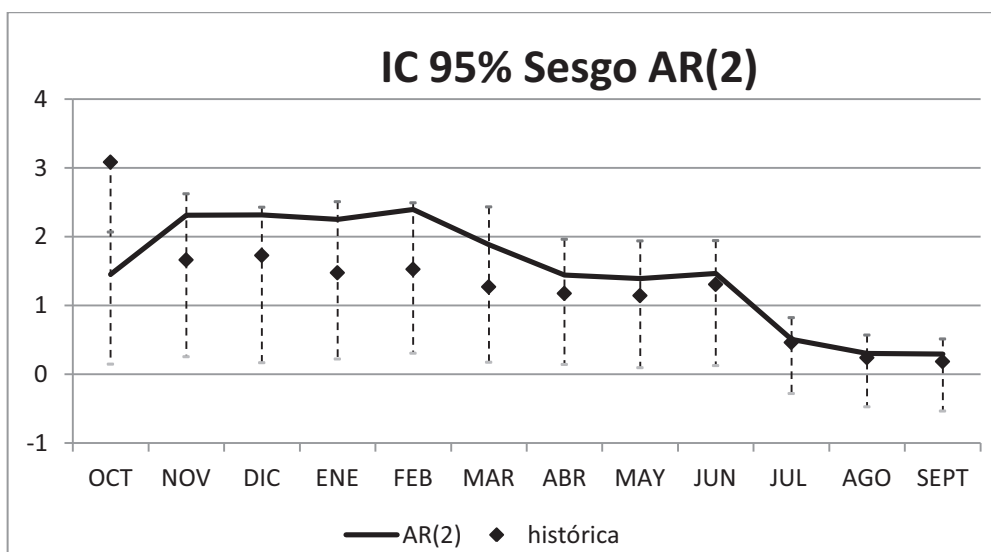
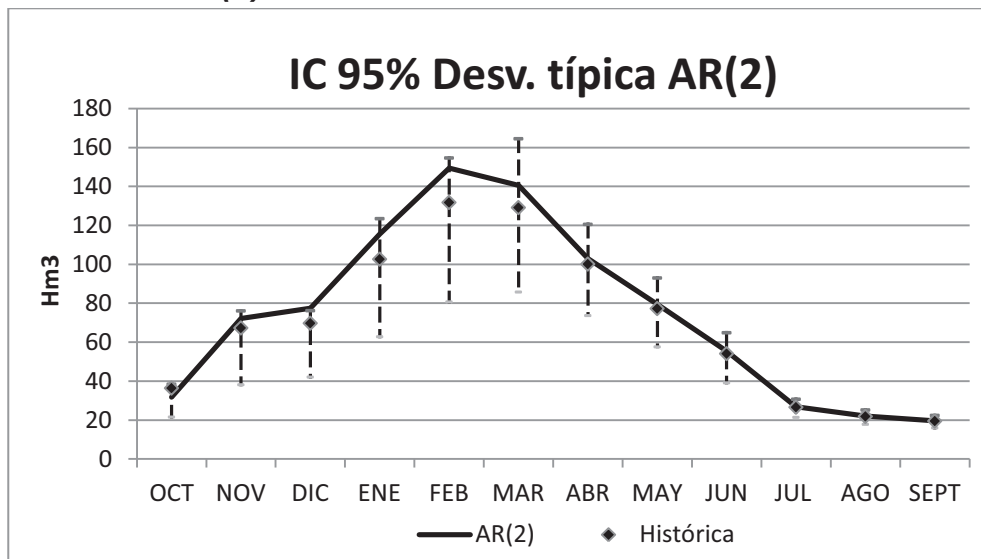
Wurbs, R. A., 1993. Reservoir system simulation and optimization models. *Journal of Water Resources Planning and Management*, 119 (4). pp 455-472,

Yevjevich, V., 1963. Fluctuations of wet and dry years. Part 1. Research data assembly and mathematical models, *Hydrology Paper 1*, Colorado State University, Fort Collins, Colorado.

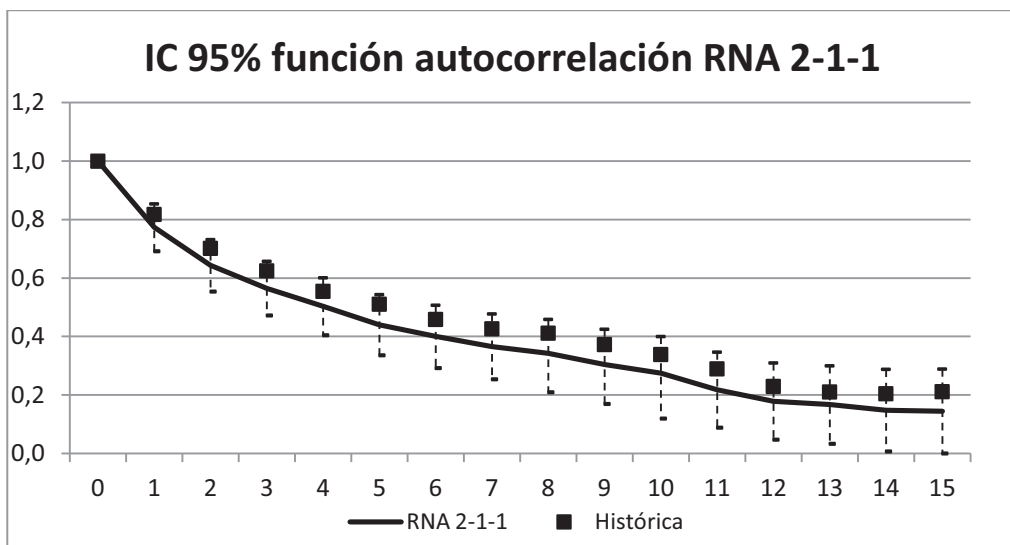
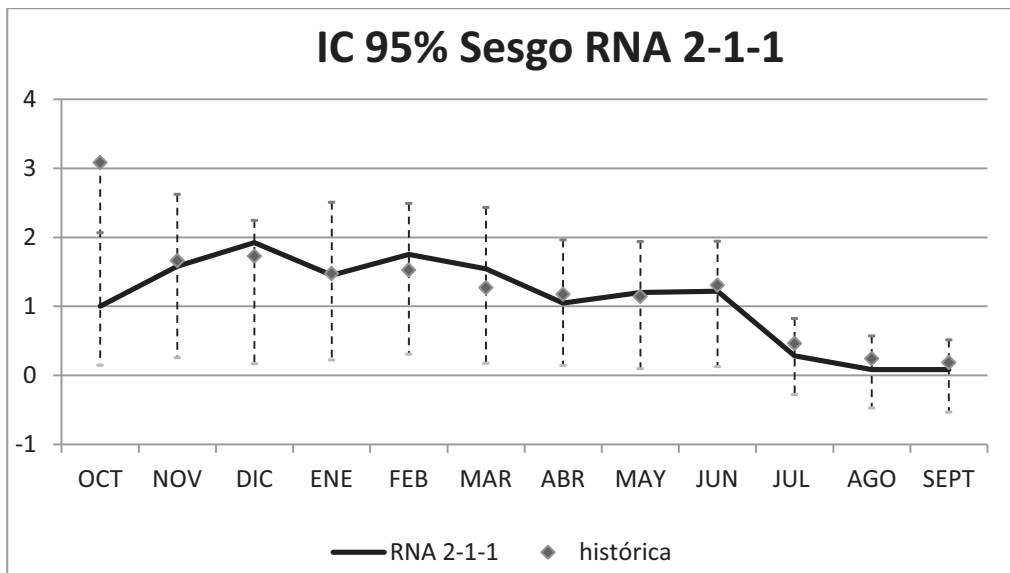
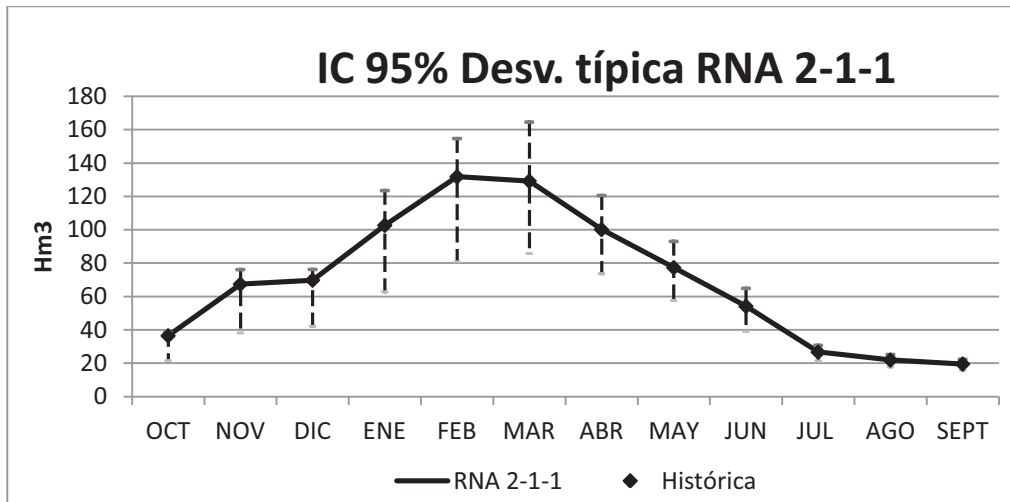
Anexo

Intervalos de confianza 95% de los estadísticos básicos para el caso univariado.

- **Modelo AR(2):**

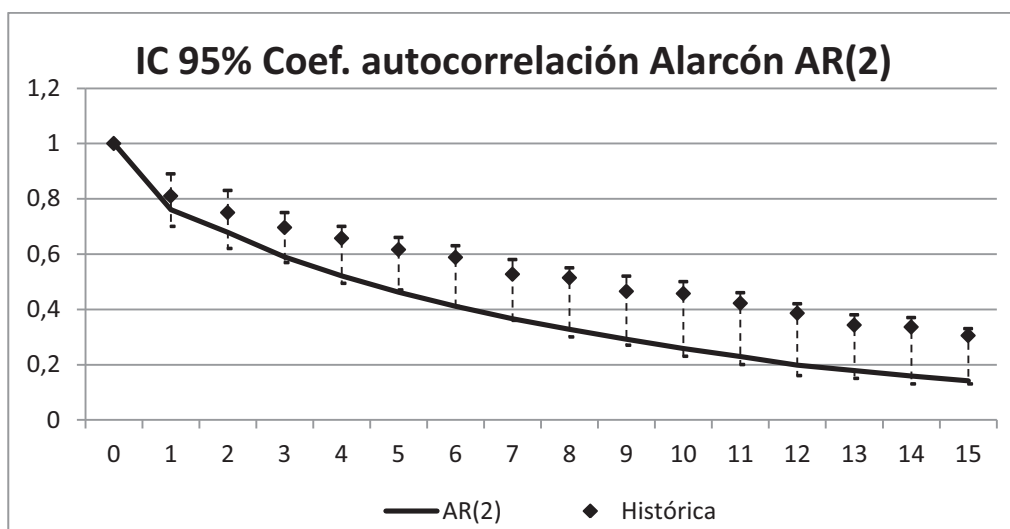
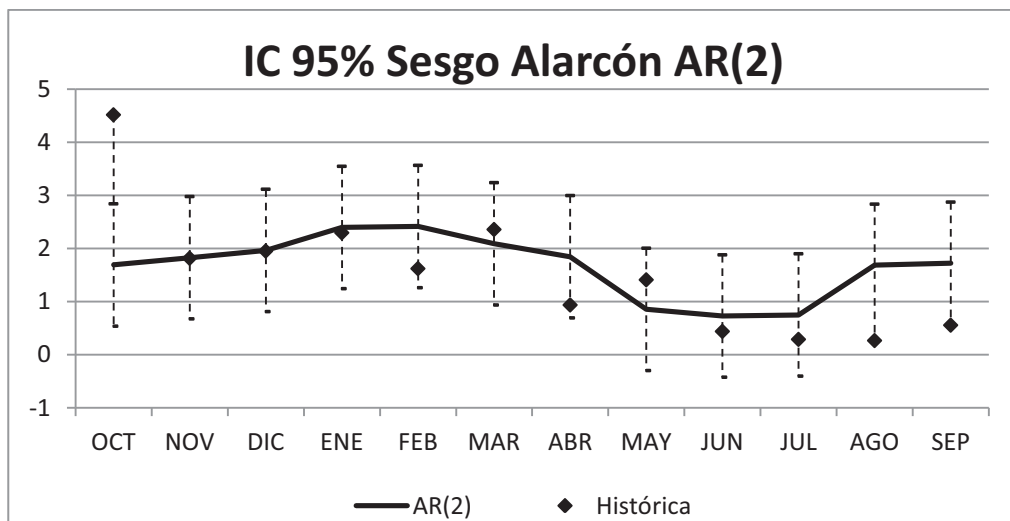
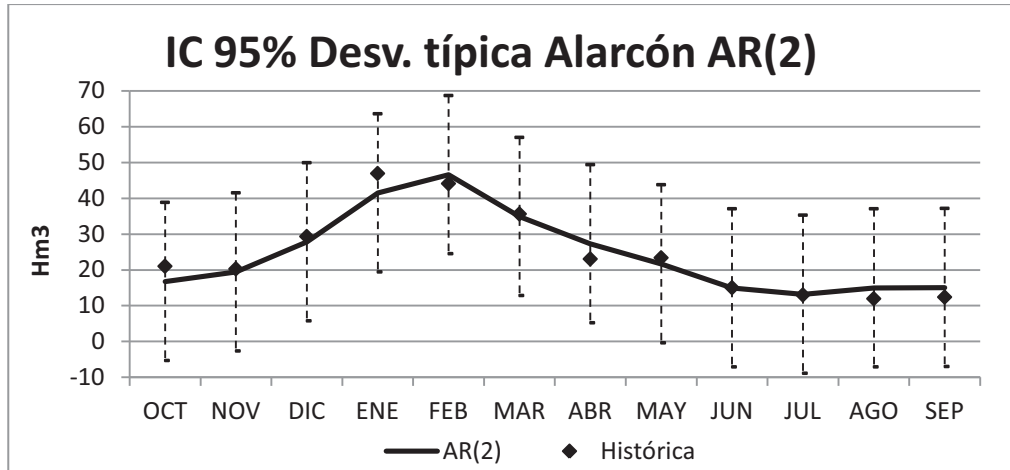


- **Modelo RNA 2-1-1:**

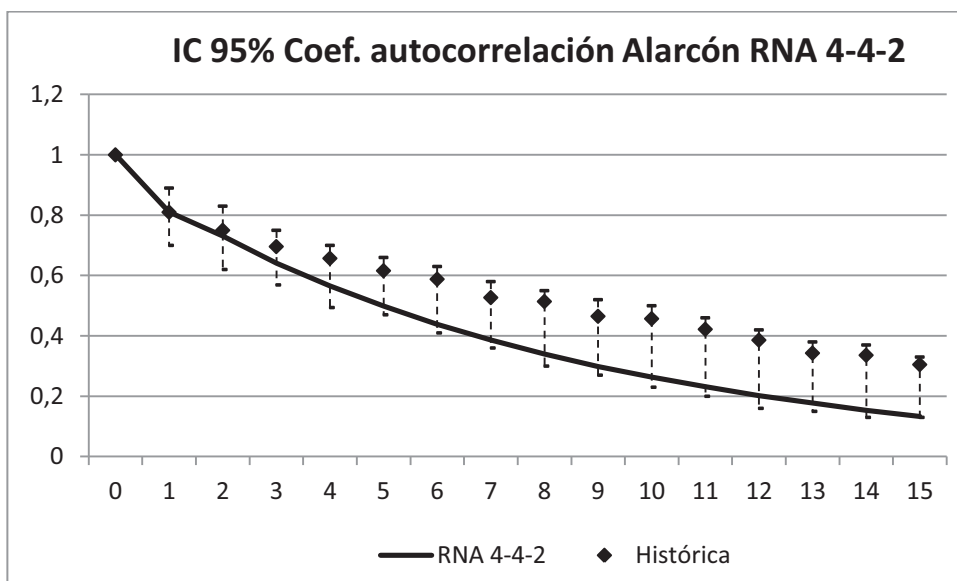
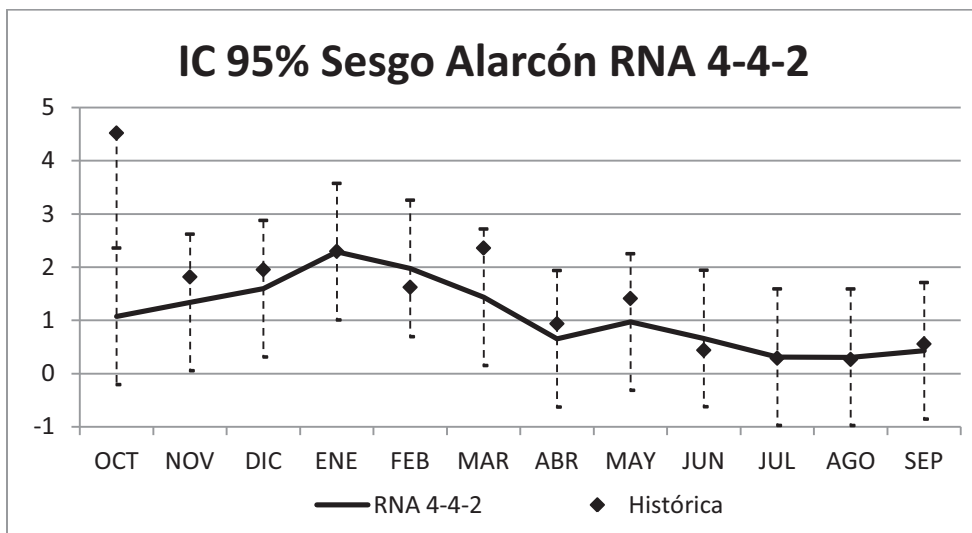
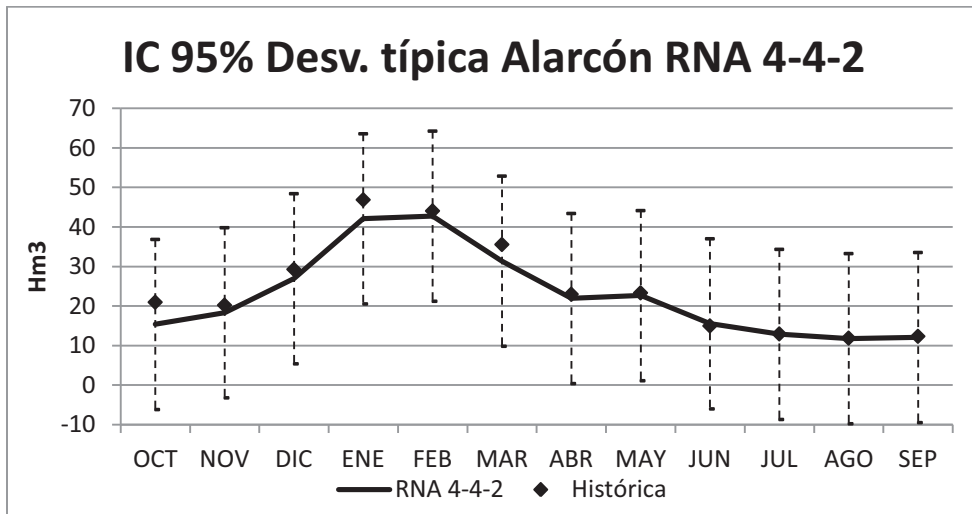


Intervalos de confianza 95% de los estadísticos básicos para el caso bivariado.

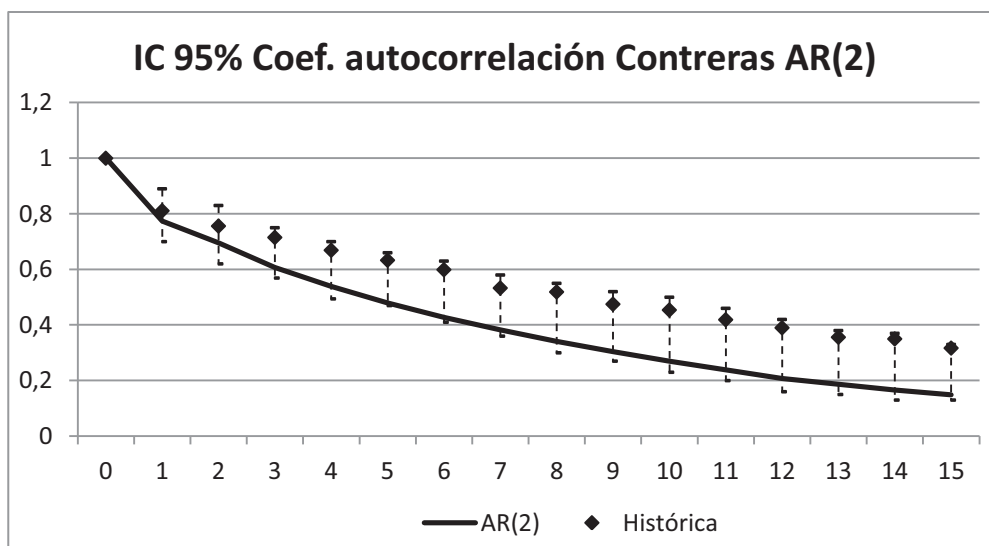
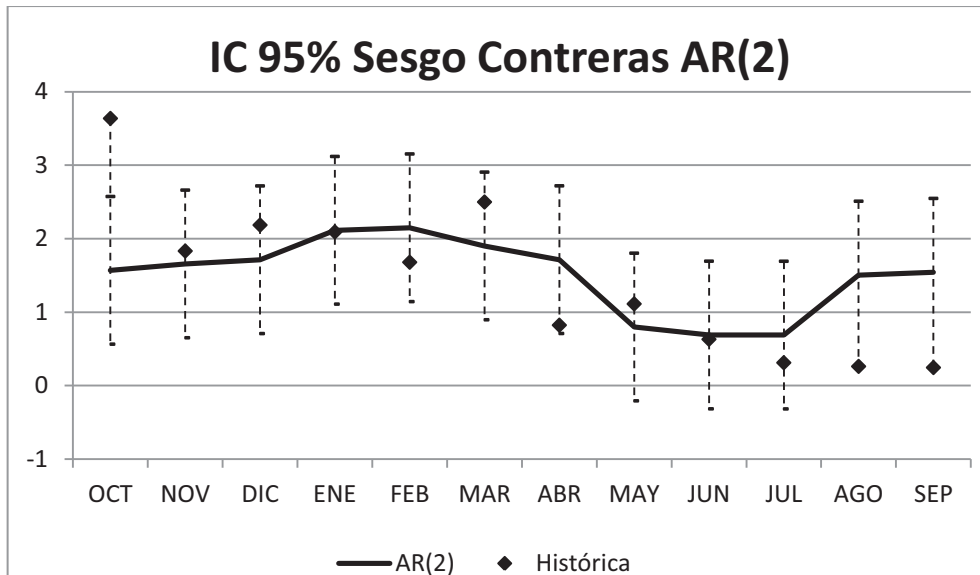
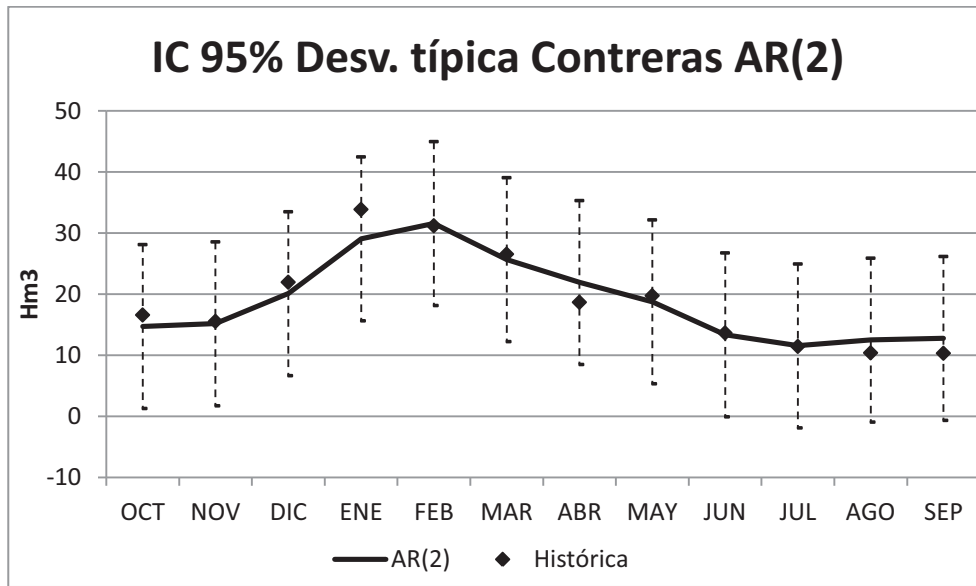
- **Alarcón- Modelo AR(2):**



- **Alarcón- Modelo RNA 4-4-2:**



• Contreras- Modelo AR(2):



- **Contreras- Modelo RNA 4-4-2:**

