

**MODELO PARA LA PREDICCIÓN DE CAUDALES MEDIANTE TÉCNICAS DE
INTELIGENCIA ARTIFICIAL**

Autor

NÉSTOR RUIZ BAUTISTA

**PROGRAMA DE INGENIERIA DE SISTEMAS
DEPARTAMENTO DE ELECTRICA, ELECTRONICA SISTEMAS Y
TELECOMUNICACIONES
FACULTAD DE INGENIERIAS Y ARQUITECTURA**



UNIVERSIDAD DE PAMPLONA

PAMPLONA, Julio 01 de 2019

**MODELO PARA LA PREDICCIÓN DE CAUDALES MEDIANTE TÉCNICAS DE
INTELIGENCIA ARTIFICIAL**

Autor

NÉSTOR RUIZ BAUTISTA

**Trabajo de grado presentado como requisito para optar al título de
INGENIERO DE SISTEMAS**

Director

JOSÉ ORLANDO MALDONADO BAUTISTA

**PROGRAMA DE INGENIERIA DE SISTEMAS
DEPARTAMENTO DE ELECTRICA, ELECTRONICA SISTEMAS Y
TELECOMUNICACIONES
FACULTAD DE INGENIERIAS Y ARQUITECTURA
UNIVERSIDAD DE PAMPLONA**

Pamplona, Julio 01 de 2019

Dedicatoria

Este trabajo está dedicado a mi familia por su apoyo incondicional para poder concluir mi carrera, a mis formadores y a todas aquellas personas que estuvieron acompañándome en este proceso de formación, brindando un poco de su tiempo para mi formación profesional.

Agradecimientos

Doy gracias a mi centro de estudios en el cual pude tener la oportunidad de formarme, de igual manera a todos los docentes con los cuales pude obtener conocimientos necesarios los cuales son soporte para mi vida profesional, agradezco también al profesor José Orlando Maldonado Bautista por ser mi tutor en mi proyecto de grado, sin sus conocimientos como base del proyecto no hubiese sido posible la culminación de este trabajo, la experiencia en el proyecto me deja claro que hay mucho conocimiento para aplicar a problemas de la vida cotidiana y que a partir de ahora inicia el proceso donde pongo a prueba todos estos conocimientos adquiridos.

Gracias a todos los que acompañaron mi proceso profesional, y a quienes dieron su apoyo incondicional.

Tabla de contenido

1. Introducción.....	9
2. Justificación.....	11
3. Objetivos.....	12
3.1 Objetivo general	12
3.2 Objetivos específicos.....	12
4. Estado del Arte y Marco Teórico	13
4.1 Estado del Arte	13
4.2 Marco Teórico	16
4.2.1 Ciclo hidrológico	16
4.2.2 Cuenca Hidrológica	17
4.2.2.1 Características de la Cuenca	18
4.2.3 Modelamiento Hidrológico.....	20
4.2.3.1 Clasificación de los modelos hidrológicos	21
4.2.4 Series de tiempo.....	22
4.2.4.1 Movimiento o Variación de una serie de tiempo.....	22
4.2.4.2 Predicción	24
4.2.5 Redes Neuronales Artificiales.....	24
4.2.5.1 La Neurona	24
4.2.5.2 La Neurona Artificial.....	26
4.2.5.3 Estructura de la Red Neuronal.....	28
4.2.5.4 Redes feedforward.....	30
4.2.5.5 Redes de Neuronas de base radial	32
4.2.5.6 Redes recurrentes.....	33
4.2.6 Validación Cruzada.....	36
4.2.6.1 Validación Cruzada con Una Observación Fuera (LOOCV)	36
4.2.6.2 Validación cruzada K-Fold Cross Validación.....	36
5. Metodología.....	37
5.1 Selección De Datos	37
5.2 Selección De Los Modelos.....	37
5.3 Selección De Las Herramientas	38
5.4 Aplicación De Los Modelos.....	38

5.5	Evaluación De Resultados.....	38
6.	Desarrollo del modelo	38
6.1	Metodología para objetivo 1	39
6.2	Metodología para objetivo 2	41
6.2.1	Toma de datos	41
6.2.2	Herramienta de modelado	43
6.2.2.1	Parámetros para utilizar la función newff para el modelo con Backpropagation 44	
6.2.2.2	Parámetros para utilizar la función newrb para el modelo con red de base radial 45	
6.2.3	Modelo de predicción hidrológico	45
6.2.3.1	Tipo de modelos estacionarios a implementar.	46
6.2.4	Implementación con función Backpropagation y validación cruzada	48
6.2.4.1	Experimento n° 1	49
6.2.4.2	Experimento n° 2.....	51
6.2.4.3	Experimento n° 3.....	56
6.2.5	Implementación con función base radial y validación cruzada	63
6.2.5.1	Experimento n° 4.....	63
6.2.6	Implementación con redes Feed Forward en Cascada y validación cruzada.....	65
6.2.6.1	Experimento n° 5.....	65
6.3	Objetivo 3.....	66
7.	Análisis de Resultados.....	67
8.	Conclusiones.....	70
9.	Recomendaciones Y Trabajos Futuros	72
10.	Bibliografía	73
11.	Anexos	76
11.1	Anexo 1:.....	76
11.2	Anexo 2	77

Lista de tablas

Tabla 1 Error medio para predicciones tiempo futuro con una red de retropropagación	62
Tabla 2 Error medio para predicciones tiempo futuro con una red de base radial	63
Tabla 3 Error medio para predicciones tiempo futuro con una red de base radial	65
Tabla 4. Error promedio para la predicción del caudal en el tiempo t , con diferentes retrasos en la dirección de entrada	67

Lista de figuras

figura 1 Ciclo Hidrológico	17
figura 2 Tipos de Cuenca	18
figura 3 Movimiento secular con recta de tendencia	23
figura 4 La Célula Nerviosa o Neurona	25
figura 5 Modelo Estándar de Neurona Artificial	26
figura 6 Funciones de Activación	27
figura 7 Arquitectura de Redes Neuronales	28
figura 8 Ejemplo del perceptron monocapa figura 9 Ejemplo del perceptron multicapa	31
figura 10 Arquitectura de la red de base radial	33
figura 11 Estructura básica de una red recurrente.....	35
figura 12 Ejemplo Red de Jordan figura 13 Ejemplo Red de Elman	35
figura 14 Ramas del machine Learning	40
figura 15 Estructura de Estación Hidrométrica.....	42
figura 16 Encabezado datos de la estación Glenamoy datos trimestral	42
figura 17 Encabezado datos de la estación Glenamoy datos anual.....	43
figura 18 Entrada de datos según el modelo autorregresivo	47
figura 19 Error de predicción del Caudal(t+1).....	50
figura 20 Predicción del caudal (t+1)	51
figura 21 Error de predicción del caudal (t+1).....	52
figura 22 Predicción del caudal (t+1)	53
figura 23 Predicción del caudal (t+1)	54
figura 24 Predicción del caudal (t+1)	54
figura 25 Error de predicción del caudal (t+1) y Predicción del caudal (t+1).....	55
figura 26 Error de predicción del caudal (t+1).....	56
figura 27 Error de predicción del caudal (t+2) y Predicción del caudal (t+2).....	57
figura 28 Error de predicción del caudal (t+5) y Predicción del caudal (t+5)	57
figura 29 Error de predicción del caudal (t+10) y Predicción del caudal (t+10)	58
figura 30 Error de predicción del caudal (t+15) y Predicción del caudal (t+15)	59
figura 31 Error de predicción del caudal (t+20) y predicción del caudal (t+20)	59
figura 32 Error de predicción del caudal (t+30) y predicción del caudal (t+30)	60
figura 33 Error de predicción del caudal (t+40) y predicción del caudal (t+40)	61
figura 34 Error de predicción del caudal (t+50) y Predicción del caudal (t+50)	61
figura 35 Comportamiento error medio para predicción de caudal $t + i$	62
figura 36 Comportamiento del error medio para $(t + i)$ para $i = 1,2,5,10,15,30$ y 50	64
figura 37 Predicción del caudal (t+50) mediante red neuronal RBF	64
Figura 38 Comportamiento del error promedio en la predicción del Caudal de agua en el tiempo $t+i$, para $i = 1, 2, 5, 10, 15, 20, 30, 40$ y 50 , utilizando una red neuronal FeedForward en cascada	65
Figura 39. Error de predicción del Caudal(t+50), para ventana de predicción con el Caudal en $t, t-1, \dots, t-5$ y Predicción del Caudal(t+50), dado el Caudal en $t, t-1, \dots, t-5$ vs la serie original de Caudal, al utilizar una red FeedForward.....	66

1. Introducción

Diferentes disciplinas que van desde la economía, ciencias e ingeniería, desde hace muchos años acogen a la predicción de series de tiempo como respuesta a problemas que requieren de toma de decisiones de corto mediano y largo plazo, las cuales a medida que surgen avances tecnológicos paralelamente avanzan las capacidades en procesamiento de los equipos utilizados, realizando cálculos complejos en minutos e incluso segundos según el modelo que se emplea para tales fines, siendo la predicción de series de tiempo muy importante como campo de investigación el cual se mantendrá activo y con muchas más aplicaciones en el futuro.

Una serie de tiempo es el registro de una secuencia de datos históricos ordenados de un determinado proceso con un intervalo de tiempo común, cuyos valores futuros son dependientes de los valores pasados. Con estos valores es posible predecir un comportamiento futuro utilizando modelos estadísticos y matemáticos que tienen gran importancia porque con estos modelos se puede llevar un control en la precisión de los valores a predecir.

Años atrás, antes de los avances tecnológicos actuales se utilizaban métodos estadísticos como los de Holt-Winters y Box-Jenkins (Gázquez Abad, Jiménez Guerrero, & Sánchez Fernández, 2006) para predecir datos en series de tiempo los cuales son adecuados para trabajarlos en series de tiempo lineales, pero con limitaciones en series de tiempo no lineales.

Para dar solución a series de tiempo que tienen una naturaleza no lineal se hace necesario implementar técnicas no estadísticas, obteniendo modelos más eficientes utilizando como alternativa a las redes neuronales artificiales (RNA), usadas en el campo de la investigación como

una herramienta eficaz para el modelado y la predicción de series de tiempo, pues las RNA tienen la característica de capturar relaciones lineales y no lineales demostrando ser más robustas, además las redes neuronales poseen la capacidad de aprender y procesar gran cantidad de información ya sea que este completa o incompleta y con alto grado de tolerancia al ruido presente en los datos.

En el presente documento se presentarán las redes neuronales como modelo útil para el pronóstico de series temporales en el campo de la hidrología, donde se analizarán datos reales de prueba como caso de estudio, para obtener como propuesta un modelo que sea óptimo en su implementación con un margen de error mínimo al ser comparado con los demás modelos de estudio de redes neuronales utilizados en el presente proyecto.

2. Justificación

Los modelos matemáticos para la predicción de series temporales suelen necesitar de un complicado ajuste de parámetros y un conocimiento profundo del fenómeno a predecir para realizar dicho ajuste. GIAASS, EDUMATEST y CICOM, están desarrollando un proyecto de investigación que requiere de la formulación de modelos de predicción hidrodinámicos desde diferentes enfoques, incluyendo modelos de ecuaciones lineales, modelos paramétricos, modelos no lineales, entre otros. El objetivo final del proyecto es "... realizar una descripción del comportamiento de los caudales del río pamplonita y el río Zulia, conforme a la variación de los fenómenos meteorológicos más relevantes durante un periodo de tiempo...". debido a que el mencionado proyecto está en una fase de recopilación y almacenamiento de datos de las estaciones disponibles, no se cuenta con una base de datos suficiente para realizar pruebas, sin embargo, se requiere avanzar con el estudio y diseño de los posibles modelos que puedan llegar a adoptarse para los fines últimos del proyecto. En el presente trabajo métodos no lineales, y en particular basados en redes neuronales serán analizados para determinar las ventajas de su adopción en el proyecto de investigación mencionado.

3. Objetivos

3.1 Objetivo general

Desarrollar un modelo para la predicción de caudales medios mediante redes neuronales y evaluar su desempeño para diferentes ventanas temporales.

3.2 Objetivos específicos

- Seleccionar y evaluar tres modelos de redes neuronales con capacidad para la predicción de series de tiempo.
- Construir un modelo hidrodinámico para la predicción de caudales medio en ríos utilizando las variables hidroclimáticas disponibles en la base de datos de prueba.
- Evaluar la capacidad de predicción para cada tipo de red, variando la ventana de predicción y los periodos de tiempo a predecir

4. Estado del Arte y Marco Teórico

4.1 Estado del Arte

(Campozano Parra, 2011) realizó en su tesis *‘Análisis de la robustez de un modelo de redes neuronales para la predicción de caudales en la cuenca alta del río Paute’* un análisis sobre la robustez del método de redes neuronales para predecir caudales, con la cual se evaluaron las variaciones respuesta del modelo teniendo en cuenta las variables de entrada y la capacidad de predicción del modelo, el análisis se realizó a la sensibilidad y la incertidumbre como cumplimiento de sus objetivos, la sensibilidad ayudó a estudiar cada variable del modelo mientras que la incertidumbre aportó al estudio de la predicción del caudal, análisis posteriores advierten sobre la calidad de los datos que se utilizan como entradas de datos deben ser de buena calidad, datos erróneos pueden ser causantes de ruido generando alteración en el modelo.

(Gomez Vargas, Obregón Neira, & Socarras Quintero, 2010) realizaron su trabajo *‘Aplicación del modelo neurodifuso ANFIS vs redes neuronales, al problema predictivo de caudales medios mensuales del río Bogotá en Villapinzón’* comparando seis tipos de modelos para el sistema neurodifuso ANFIS en el cual observan el tipo de comportamiento que estos modelos tendrán, utilizando datos de caudales medios, los resultados serán comparados con un modelo de red neuronal de tipo perceptron, en este trabajo observaran si los resultados son dependientes de la cantidad de entradas a los modelos, de estos resultados se evalúa si los modelos son factibles para la implementación en el campo de la hidrología.

(Gómez Gallego, 2012) propone en su trabajo '*Redes neuronales para la simulación, estimación y control de caudales impulsados por estaciones de bombeo en zonas regables*' la demanda en un sistema de distribución de agua, donde aplica inicialmente métodos estadísticos tradicionales como regresión lineal y modelo de redes neuronales, utilizando datos históricos de caudales diarios de los cuales se tiene que definir una relación de datos pasados y futuros, de igual manera tiene en cuenta factores externos como lo son las temperaturas , el viento , radiación solar, precipitación y evapotranspiración potencial, el objetivo del trabajo será definir el modelo más viable que sea capaz de predecir según los objetivos, la demanda en un tiempo definido de 24 horas, para ello utiliza datos históricos de todas las variables que influyen en sistema del caudal.

(Matinez Gomez & Pedraza Caballero, 2012) proponen en su trabajo '*Implementación web de redes neuronales artificiales aplicadas de la predicción de series de tiempo*' una plataforma web donde implementaran dos tipos de red neuronales artificiales, el perceptron multicapa y la red neuronal basada en función de base radial para el análisis de series de tiempo, este trabajo se realiza con el fin de establecer resultados para la simulación de los resultados , teniendo en cuenta los datos recolectados , estos resultados se analizaran con el fin de saber si los valores de las variables utilizadas crecen o decrecen y bajo que costo computacional se lograran los resultados.

(Moscoso Barrera, 2015) presenta la comparación entre modelos físicos (Hec-Ras) y modelos inteligentes (redes neuronales – redes ANFIS) , utilizadas para la predicción, se realiza la comparación de diferentes modelos tanto físicos como inteligentes para encontrar el modelo que arroje buenos resultados , tomando datos históricos de la cuenca del rio Bogotá comprendidos desde 2009 a 2013, con estos modelos se desea determinar el modelo que más se aproximan a los

datos reales, estos resultados son analizados y con ellos se determina según los objetivos que se plantean, generar datos que ayuden a pronosticar zonas inundadas.

(Pierini, Gomez, & Telesca, 2012) realizan en su trabajo '*Predicción de caudales en río Colorado, Argentina*' el análisis de resultados a partir de pruebas realizadas con redes neuronales y modelo autorregresivo, los datos a utilizar en estas pruebas corresponden a datos históricos de 18 años del caudal estudiado, en el trabajo se realizará el modelado y la predicción de las fluctuaciones diarias del caudal, los dos modelos a trabajar se compararan con el fin de determinar cual está sujeta a trabajar de manera eficiente y que los resultados obtenidos sean fiables, ya que dependiendo de los resultados se puede determinar por ejemplo la regularización de los embalses y centrales hidroeléctricas, y el abastecimiento de agua en ciudades como Bahía Blanca.

(Ocampo & Vélez, 2014) plantean en su trabajo '*Análisis comparativo de modelos hidrológicos de simulación continua en cuencas de alta montaña: caso del río Chinchiná*' la implementación de un modelo capaz de gestionar el recurso hídrico en alta montaña, en el trabajo se plantea un análisis multimodelo para capturar incertidumbre asociada mediante la aplicación de siete modelos hidrológicos con diferentes niveles de complejidad, el análisis realizado a los resultados se hicieron mediante métodos gráficos y estadísticos, los datos estadísticos, los hidrógramas y curvas de duración permiten analizar la escorrentía en cuencas de alta montaña.

4.2 Marco Teórico

4.2.1 Ciclo hidrológico

El ciclo hidrológico (Aparicio Mijares, 1992) son una serie de estados del agua presentes en la naturaleza, los cuales se ejecutan simultáneamente de forma continua y no tienen un principio o un fin, como foco central de la hidrología, este ciclo puede estudiarse de manera adecuada mediante su representación esquemática, la cual describe todos los procesos presentes en él. Partiendo de cualquier punto del esquema del ciclo, puede iniciarse por la evaporación del agua de los océanos o desde la superficie terrestre debido a la radiación solar y el viento para llegar a la atmosfera, este vapor de agua forma nubes y se condensa, luego se precipita sobre la superficie terrestre o los océanos. Durante su trayecto de agua precipitada esta puede volver a evaporarse o ser interceptada por plantas, convertirse en flujo superficial sobre el suelo, infiltrarse en él, volverse flujo por debajo del suelo y luego poder descargarse en ríos. El agua que se intercepta y la que corre por la superficie regresa a la atmosfera mediante la evaporación. De la que se precipita y llega a las corrientes parte de esta se infiltra profundamente para recargar aguas subterráneas de donde pueden emerger en forma de manantiales o se pueden deslizar hacia los ríos para que circulen libremente sobre la superficie, fluyendo hacia el mar o se evapora hacia la atmosfera. Del agua infiltrada, otra parte es absorbida por las plantas y posteriormente es transpirada, casi en su totalidad y a medida que continua el ciclo hidrológico.

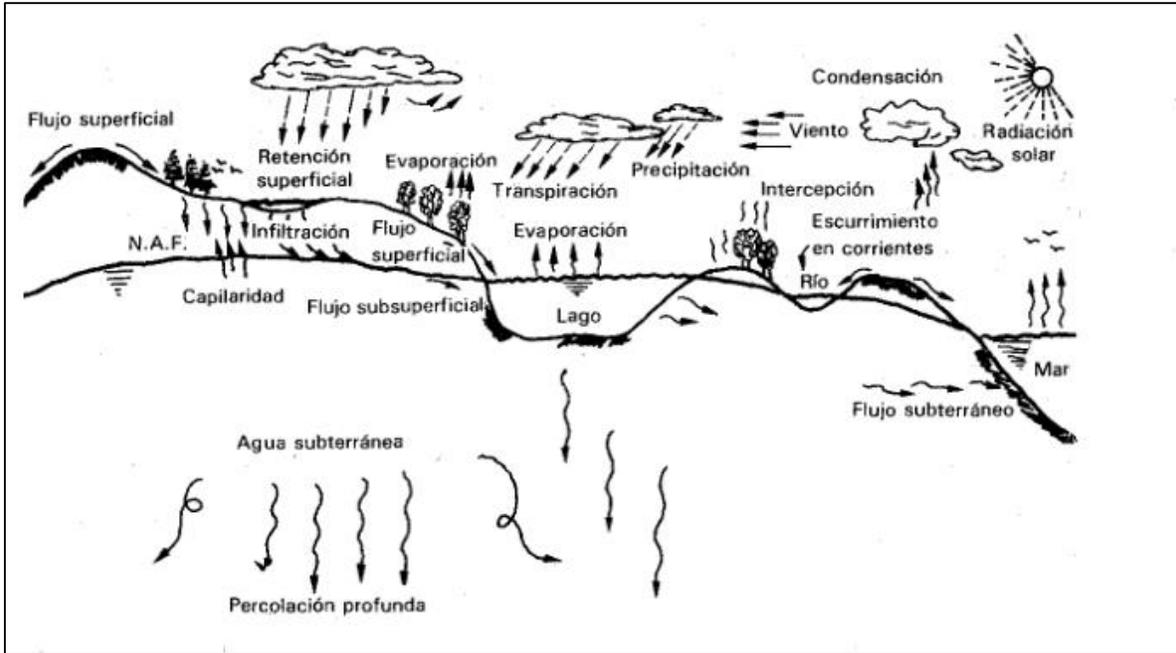


figura 1 Ciclo Hidrológico

4.2.2 Cuenca Hidrológica

La cuenca es la unidad básica de estudio del ciclo hidrológico, es un terreno ancho y profundo cuyas aguas llegan a un punto en común, es decir una cuenca hidrológica define a una red hídrica. Las cuencas pueden ser superficiales o subterráneas, las superficiales tienen a drenar las gotas de lluvia por un sistema de corrientes hacia un mismo punto de salida, y la cuenca subterránea, es semejante a la superficial ya que poseen la misma definición, diferenciándose en el punto de vista de su salida, ya que pueden haber dos tipos de cuencas: las cuencas endorreicas tienen un punto de salida que está dentro de los límites de las misma cuenca y generalmente es un lago, y las cuencas exorreicas cuyo punto de salida se encuentra en los límites de la cuenca y está en otra corriente o en el mar.

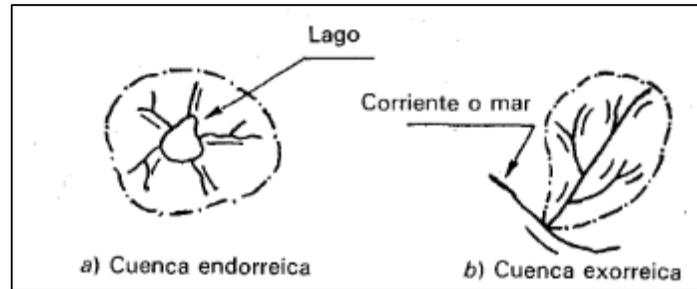


figura 2 Tipos de Cuenca

4.2.2.1 Características de la Cuenca

Las características de las cuencas; son de dos tipos los cuales se condicionan según los fenómenos como la precipitación y el escurrimiento, según la manera en que controlan los fenómenos mencionados:

- Las que condicionan el volumen de escurrimiento, como el área de la cuenca y el tipo de suelo.
- Las que condicionan la velocidad de respuesta, como son el orden de corrientes, pendiente de la cuenca y los cauces, etc.

Las características que a continuación se presentan son las de mayor importancia según la relación escurrimiento-precipitación.

- 1) **Parteaguas:** línea imaginaria formada por los puntos de mayor nivel topográfico la cual separa una cuenca de sus vecinas.
- 2) **Área de la cuenca:** superficie, en proyección horizontal, delimitada por el parteaguas.
- 3) **Corriente principal:** corriente que pasa por la salida de la misma, aplicada solamente a cuencas exorreicas, las demás corrientes son llamadas tributarias, todo punto en cualquier

corriente tiene una cuenca de aportación, y toda cuenca tiene solo una corriente principal. Las cuencas correspondientes a las corrientes tributarias o puntos de salida se llaman cuencas tributarias o sub-cuencas.

Mientras más corrientes tributarias tenga una cuenca, mayor será su grado de bifurcación en su sistema de drenaje, mayor será su respuesta a su precipitación.

Las corrientes se clasifican de manera por el tiempo en que transportan agua:

- a) **Corrientes perenne:** Es una corriente cuyo punto más bajo del cauce se encuentra abajo del nivel de aguas freáticas (aquellas que se acumulan bajo la tierra, almacenadas en los poros de arena y grava), son afluentes ya que transportan agua durante todo el año y siempre están alimentadas, totalmente o en parte por aguas subterráneas.
- b) **Corrientes intermitentes:** Transporta agua durante la época de lluvias de cada año, hay transporte de agua si el nivel freático asciende de un punto de referencia, y no hay transporte agua y desciende ese mismo punto de referencia.
- c) **Corrientes efímeras o influentes:** cuando el freático está bajo del punto de referencia, solo transporta agua después de una tormenta, alimentando solo a los almacenamientos de agua subterránea.

Y de la manera por su posición topográfica o edad geológica:

- a) **Montaña o juveniles:** los ríos de montaña, tienen la característica de tener cotas elevadas sobre el nivel del mar, tienen mucha pendiente y pocas curvas debido a las velocidades de sus aguas.

- b) **Transición o maduros:** se encuentran en medio de los ríos de montaña y los de planicie, presentan algunas curvas, con velocidades moderadas de agua.
- c) **Planicie o viejos:** presentan muchos meandros (curvas muy pronunciadas) debido a que poseen bajas velocidades de agua, estos ríos poseen cotas cercanas al nivel del mar.

4.2.3 Modelamiento Hidrológico

El modelamiento hidrológico (Te Chow, R. Maidment, & W. Mays, 1994) lo que busca es aproximarse al sistema real, ya que uno de los objetivos que busca el análisis de los sistemas hidrológicos es estudiar cómo funcionan dichos sistemas y predecir sus salidas, la modelación hidrológica es una herramienta utilizada para la gestión de cuencas urbanas y rurales (Mediero Orduña, 2007), puede involucrar la identificación de unidades de respuesta hidrológica denominada subcuentas.

Para cada una de estas unidades se caracterizan las fases del ciclo hidrológico: precipitación, infiltración, escorrentía y agua subterránea. De esta forma se pueden obtener los datos de entrada para la obtención de los modelos ya que sus entradas y salidas son variables que pueden ser medidas y la estructura es un conjunto de ecuaciones que conectan las entradas y las salidas. Con dichos modelos se puede simular el comportamiento de procesos hidrológicos (Ocampo & Vélez, 2014) debido a los cambios en las condiciones ambientales, tal como la variación de la precipitación o la escorrentía, así como la disminución de la infiltración a causa del crecimiento urbano. Los resultados obtenidos a través de simulaciones son utilizados en la gestión de recursos, así como la planificación y manejo de las obras hidráulicas, y la toma de decisiones para la prevención temprana de riesgos.

4.2.3.1 Clasificación de los modelos hidrológicos

1. **Modelos físicos:** poseen modelos a escala (ej: tamaño reducido de una presa mediante modelo hidráulico) y modelos análogos (ej: movimiento de fluido viscoso para modelar flujo a través de un acuífero).
2. **Modelos abstractos:** representación matemática del sistema, describiendo las operaciones mediante ecuaciones, relacionando las variables de entrada y las de salida, variables que pueden ser funciones de espacio y tiempo.

Trabajar con variables aleatorias es una tarea ardua, para ello se simplifican los modelos para que sea más práctico, despreciando algunas fuentes de variación, estos modelos hidrológicos pueden clasificarse como:

2.1) Modelo determinístico: no considera aleatoriedad, una entrada dada produce siempre una misma salida.

2.2) Modelo estocástico: tiene salidas que son por lo menos parcialmente aleatorias.

Los modelos determinísticos hacen pronósticos, mientras que los modelos estocásticos hacen predicciones.

4.2.4 Series de tiempo

Una serie de tiempo (Aparicio Mijares, 1992) es un conjunto de observaciones tomadas en momentos o tiempo específicos, generalmente en intervalos iguales como, por ejemplo: las temperaturas a cada hora anunciadas por el centro meteorológico de una ciudad.

Matemáticamente, una serie de tiempo se define por medio de los valores $y_1, y_2 \dots$ de una variable Y (temperatura) en los tiempos $t_1, t_2 \dots$, por lo tanto, Y es una función de t esto se denota por $y = f(t)$, una serie de tiempo que involucre la variable y se representa por la construcción de la gráfica de Y con respecto a t .

4.2.4.1 Movimiento o Variación de una serie de tiempo

El movimiento se describe en la gráfica como un punto moviéndose con el paso del tiempo, cuando se habla de movimiento se dice que es la trayectoria de una partícula influenciada por fuerzas físicas, pero en este caso este punto lo afecta fuerzas económicas, sociológicas, fisiológicas, etc.

Las variaciones o movimientos que se pueden presentar en una serie de tiempo pueden ser de acuerdo a (Levin & Rubin, 2004):

1. **Movimientos a largo plazo o seculares:** Estos se refieren a la dirección general en la que la gráfica de series de tiempo parece seguir sobre un intervalo grande de tiempo. Está indicado por una recta tendencia, que se muestra discontinua. Para algunas series de tiempo puede ser más adecuada una curva de tendencia.

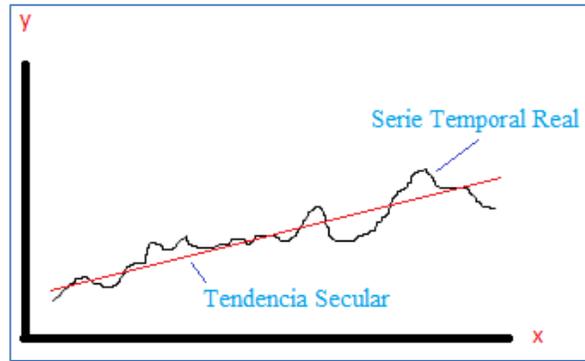


figura 3 Movimiento secular con recta de tendencia

2. **Movimientos cíclicos o variaciones cíclicas.** Estos tienen que ver con las oscilaciones o los movimientos respecto de una recta o curva de tendencia. Estos ciclos, como se les denomina en ocasiones, son o no periódicos; es decir, pueden o no seguir patrones exactamente similares, después de intervalos iguales de tiempo. En los negocios y las actividades económicas, los movimientos son considerados cíclicos sólo si se repiten después de intervalos mayores de un año.

3. **Movimientos o variaciones estacionales:** Estos se relacionan con los patrones idénticos o similares, que las series de tiempo parecen seguir durante los meses o trimestres correspondientes a los años sucesivos. Tales movimientos se deben a eventos recurrentes que suceden anualmente, como el repentino incremento en las ventas de las tiendas de departamento, previo a la navidad.

4. **Movimientos irregulares o aleatorios.** Estos se refieren a los movimientos esporádicos de las series de tiempo, debido a eventos aleatorios tales como: inundaciones, huelgas o elecciones. Aunque generalmente se considera que tales eventos producen variaciones que

duran poco tiempo, cabe la posibilidad de que sean tan intensos que resulten en nuevos movimientos cíclicos de otro tipo.

4.2.4.2 Predicción

Los métodos (porcentaje promedio, porcentaje de la tendencia o razón de la tendencia, promedio móvil o razón de promedio móvil) y principios (índice estacional, des-estacionalización de los datos, estimación de variaciones cíclicas, estimación de las variaciones irregulares, comparación de los datos) se utilizan en el importante trabajo de predecir series de tiempo, en donde se debe tener en cuenta que el tratamiento matemático de los datos no resuelve por sí mismo todos los problemas. sin embargo, el análisis matemático ha probado ser valioso para la predicción a corto y a largo plazo.

4.2.5 Redes Neuronales Artificiales

El objetivo de las redes neuronales es llegar a diseñar maquinas con elementos neuronales de procesamiento paralelo, de modo que se llegue a emular con la red, de forma similar a como lo hacen los sistemas neuronales de los animales.

4.2.5.1 La Neurona

La célula nerviosa o neurona es la parte esencial del sistema de comunicación neuronal, utilizan sus productos de secreción como señales químicas (transmisores), para la transmisión de la información, esta se envía entre las distintas neuronas a través de prolongaciones, formando redes donde se elabora y se almacena la información. Una parte de las neuronas está en relación con los

receptores, donde llegan comunicaciones procedentes del exterior o interior del organismo hasta las redes neuronales.

Las funciones parciales de la neurona son:

- Recoger información que llega a ella en forma de impulsos procedentes de otras neuronas o de los receptores.
- Integrar la información en un código de activación propio de la célula.
- Transmitir la información codificada en forma de frecuencia de impulsos a través del axón.
- Distribuir espacialmente los mensajes a través de las ramificaciones del axón.
- Transmitir los impulsos a las neuronas subsiguientes o a las células efectoras por sus terminales.

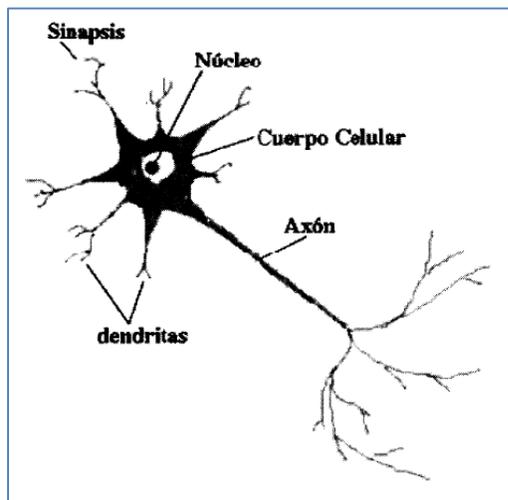


figura 4 La Célula Nerviosa o Neurona

4.2.5.2 La Neurona Artificial

Llamada también célula o autómatas, es un elemento que posee un estado interno, llamado nivel de activación, recibe señales que le permiten en su caso cambiar de estado.

Si se denomina $S = \{0,1\}$ siendo 0 el estado inactivo y 1 el activo. S también podrá tomar un mayor número de valores, $S = \{0,1, 2, \dots, n\}$ para representar.

Se puede representar gráficamente la unidad de proceso típica de la neurona artificial como se muestra a continuación:

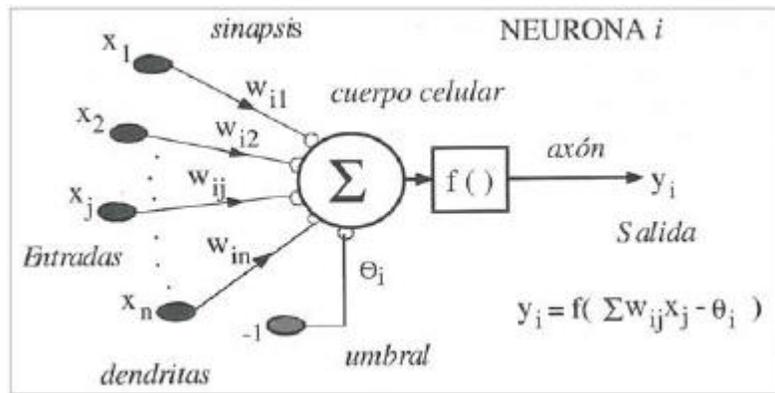


figura 5 Modelo Estándar de Neurona Artificial

partes de la neurona artificial:

- Conjunto de entradas $X_j(t)$ y pesos sinápticos W_{ij}
- Regla de propagación: $h_i(t) = \sigma(w_{ij}, x_j(t)); h_i(t), = \Sigma w_{i,j}x_j$
- Función de Activación $y_i(t) = f_i(h_i(t))$: salida y estado de activación.

El grupo de entradas x_1, x_2, \dots, x_n son introducidos en una neurona artificial. Estas entradas corresponden a las señales de sinapsis de una neurona biológica. Cada señal se multiplica por un peso asociado w_1, w_2, \dots, w_n antes de que sea aplicado al sumatorio etiquetado por Σ . cada peso

corresponde a la fuerza de conexión sináptica. El sumatorio, corresponde al cuerpo de la neurona, suma todas las entradas ponderadas algebraicamente produciendo una salida que se denomina E, así:

$$E = x_1w_1 + x_2w_2 + \dots + x_nw_n$$

Las señales E son procesadas por la función de activación o de salida F , que produce la señal de salida de la neurona S , dependiendo de la función F , habrán distintos modelos de autómatas.

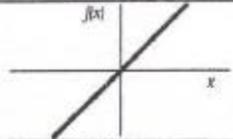
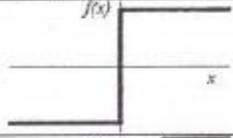
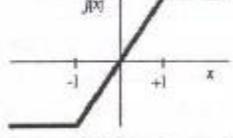
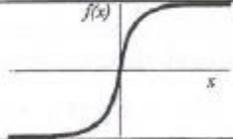
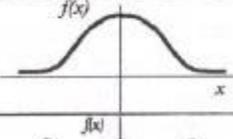
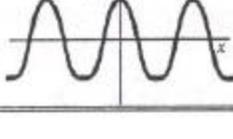
	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Línea a tramos	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } -l \leq x \leq +l \\ +1, & \text{si } x > +l \end{cases}$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1 + e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = Ae^{-Bx^2}$	$[0, +1]$	
Sinusoidal	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	

figura 6 Funciones de Activación

4.2.5.3 Estructura de la Red Neuronal

La red posee una serie de entradas a la neurona, cada una llega de la salida de otra neurona de la red. Una vez se calcula la salida de una neurona, esta se propaga, vía conexiones de salida a las células destino. Todas las conexiones de salida reciben el mismo valor de salida, a medida que las células se conectan entre sí, se denominan patrón de conectividad o arquitectura de la red.

A través de las interconexiones de red (rutas de comunicación) viajan valores numéricos de una célula a otra, estos valores son evaluados por los pesos de las conexiones. Los pesos de las conexiones se ajustan durante la fase de aprendizaje para producir una red de neuronas artificial final.

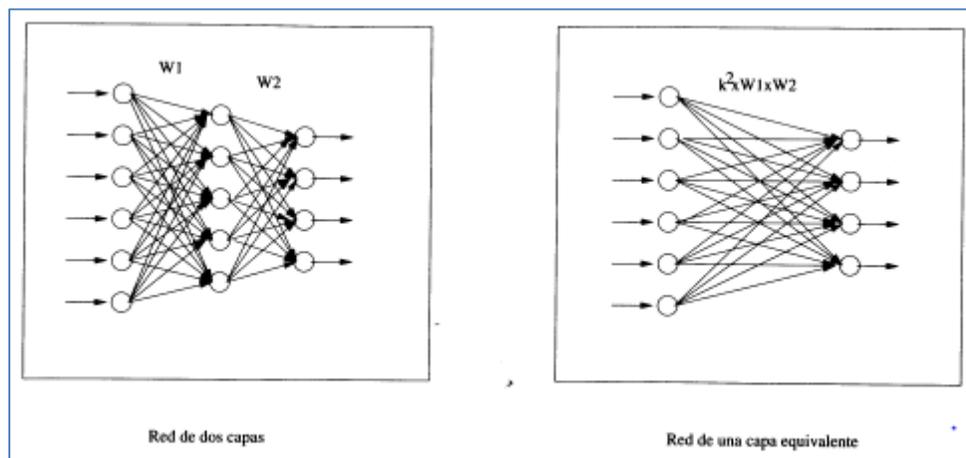


figura 7 Arquitectura de Redes Neuronales

4.2.5.3.1 Modos de operación

MATLAB define a la red neuronal como un modelo de computación cuya estructura se asemeja a una neurona del cerebro, con capas de nodos conectados entre sí, esta red posee la característica de aprender de la información que se le proporcione para que posteriormente se puedan reconocer patrones clasificar datos y pronosticar eventos.

Una de las características importantes de las redes neuronales de MATLAB es que descompone las entradas en capas, las entradas se propagan a través de la red con un algoritmo que se le asigne hasta el punto de obtener una salida, el algoritmo tendrá una característica especial para aprender, las redes sirven para modelar sistemas, así que al tener una entrada determinada la red iniciara con la fase de entrenamiento para dar una salida.

los dos modos de operación que tienen las redes neuronales son el de aprendizaje y recuerdo:

- La fase de aprendizaje convergencia o también llamado entrenamiento, es donde a partir de una estimulación del entorno que rodea la red se producen ajustes en los parámetros libres de la red en el desarrollo del algoritmo, el aprendizaje surge cuando se determinan los pesos sinápticos los cuales permiten que la red se transforme según lo deseado en función de los pesos y bias (umbral), dando como resultado una salida para las mismas entradas. El objetivo del entrenamiento es encontrar pesos ideales que sean óptimos para la aplicación que se le quiera dar a la red neuronal.

- La fase de recuerdo o ejecución le sigue a la fase de aprendizaje en la cual el aprendizaje se desconecta, es decir los pesos ya no se modifican y quedan fijos según el entrenamiento permitiendo tener una red neuronal lista para ser procesada y que pueda resolver un determinado problema.

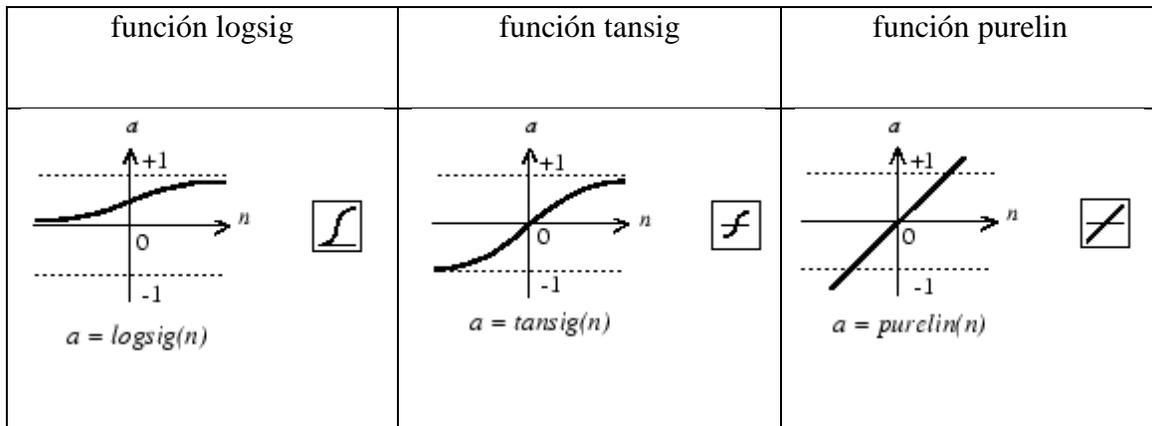


figura 8 Funciones de transferencia utilizadas por MATLAB
Fuente: documentación de Matlab

4.2.5.4 Redes feedforward

Son aquellas redes que se enfocan en, lo que se puede hacer mejor o distinto en un futuro y la forma en que la información fluye es unidireccional desde las neuronas de entrada a la(s) capa(s) de procesamiento, existiendo así dos tipos de redes las monocapa y multicapa hasta llegar a la capa de salida.

Existen dos modelos (Isasi Viñuela & Galván León, 2004) de redes feedforward el primero es el perceptron simple el cual es una red monocapa la cual es una red formada por una sola capa de neuronas desconectadas entre sí, y una o varias células de salida, las conexiones determinan las superficies de discriminación del sistema es decir utiliza formas lineales como funciones

discriminantes, su limitación está en que solo puede sintetizar clasificadores lineales, discriminando solo conjuntos linealmente separables.

El segundo modelo es el perceptron multicapa o red multicapa con conexiones hacia adelante, este es una generalización del perceptron simple, y que nace en consecuencia de la limitación que tiene la arquitectura del sistema simple referente al problema de separabilidad no lineal, las conexiones del perceptron multicapa siempre están dirigidas hacia adelante, es decir, las neuronas de una capa se conectan a una capa siguiente, de ahí surge el nombre de redes alimentadas hacia adelante o redes feedforward desde la capa de entrada existen conexiones a la primera capa oculta y de esta a otra capa, se dice entonces que la red está totalmente conectada.

Cuando se desea abordar un problema con el perceptron multicapa, uno de los pasos a seguir es realizar la arquitectura de la red, es decir, determinar la función de activación a emplear ya sea una función sigmoïdal o una función tangente hiperbólica, definir el número de neuronas y el número de capas de red, no importa la función de activación que se elija, esto no influye en la solución del problema sino en el recorrido que se desee utilizar.

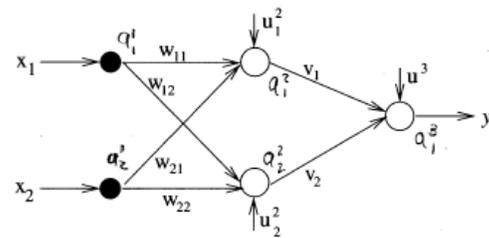
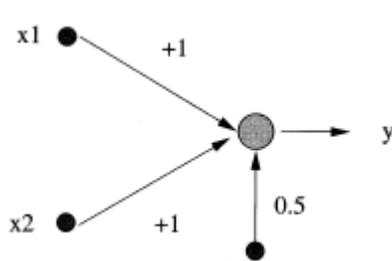


figura 8 Ejemplo del perceptron monocapa **figura 9** Ejemplo del perceptron multicapa

el proceso de aprendizaje o entrenamiento de la red, es necesario fijar los parámetros que intervienen en dicho proceso; la razón de aprendizaje y el número de ciclos de aprendizaje, ambos dependiendo del problema a resolver y se hace necesario realizar varias simulaciones previas para fijar ambos parámetros de acuerdo al problema.

4.2.5.5 Redes de Neuronas de base radial

Son redes multicapa de base radial (Isasi Viñuela & Galván León, 2004) con conexiones hacia adelante con similitud al perceptron multicapa al ser aproximadores universales ya que pueden aproximar cualquier función continua, caracterizado porque están formadas por una única capa oculta y cada neurona de la capa posee carácter local, haciendo que cada neurona oculta de la red se activa en una región diferente del espacio de patrones de entrada. El carácter local es dado por el uso de funciones de base radial como por ejemplo la función gaussiana como función de activación, estas redes realizan una combinación lineal de las activaciones de las neuronas ocultas, es decir, solo unas pocas neuronas ocultas pueden ser procesadas para nuevos patrones de entrada.

Cada neurona de la red de base radial construye una aproximación local y no lineal en una determinada región del espacio de entrada, ya que la salida de la red es una combinación lineal de las funciones de base radial, las aproximaciones que constituyen las redes de base radial son combinaciones lineales de múltiples funciones locales y no lineales, lo que hace que relaciones complejas se vuelvan colecciones locales menos complejas diferenciándose de las aproximaciones globales del perceptron, el cual hace que se pueda aplicar al campo del análisis de las series temporales.

Las redes de base radial están formadas por tres capas de neuronas, la capa de entrada la cual recibe señales de exterior, transmitiendo a la siguiente capa sin realizar procesos sobre las señales, en la capa oculta se reciben las señales de la capa de entrada y realizan la transformación local y no lineal de las señales, y la capa de salida quien es la que realiza la combinación lineal de las activaciones de las neuronas ocultas, actuando también como salida de la red.

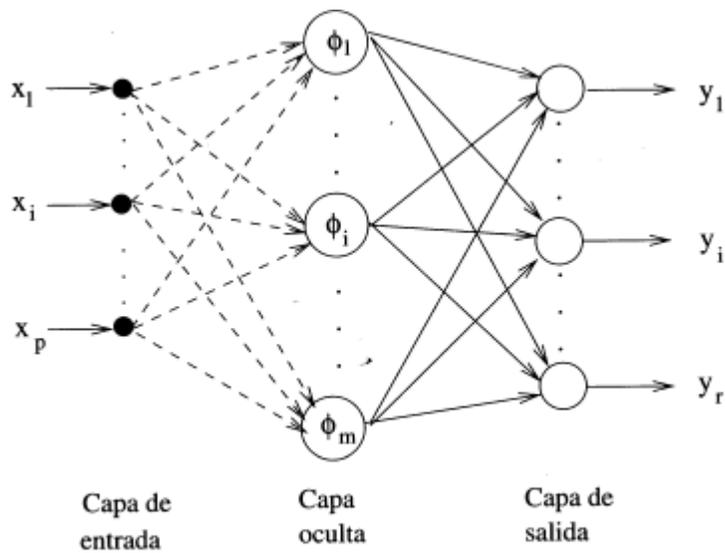


figura 10 Arquitectura de la red de base radial

el número de neuronas ocultas en la red se puede determinar a modo de prueba y error, el número de neuronas se puede variar hasta que se pueda conseguir una red que pueda ser capaz de resolver el problema, la diferencia de esta red con un perceptrón multicapa es que en la red de base radial al añadirse o eliminarse neuronas de la capa oculta influye significativamente en los resultados obtenidos para la red.

4.2.5.6 Redes recurrentes

Las redes recurrentes (Isasi Viñuela & Galván León, 2004) son aquellas que no están sometidas a la restricción de conexión entre neuronas que crean ciclos o bucles, este tipo de redes tienen la característica de crear bucles mediante el uso de las llamadas conexiones recurrentes, en donde conexiones de una neurona pueden llegar a ella misma, pueden aparecer de igual manera conexiones entre neuronas de una misma capa, conexiones de una capa a una anterior, esto sirve para aumentar pesos o parámetros ajustables en la red, aumentando así la capacidad de representación, sin embargo esto complica un poco el aprendizaje de las redes recurrentes.

La activación en una neurona con redes recurrentes no depende de la capa anterior, sino que también puede depender del estado o activación de cualquier otra neurona que tenga conexión con ella o incluso de su propia activación, es por eso que se incluye la variable del tiempo en la activación de una neurona.

Existen varios tipos de redes recurrentes que poseen características en cuanto a la actuación que poseen las redes recurrentes, la primera es conocida como redes parcialmente recurrentes las cuales se caracterizan porque solo aparecen pocas conexiones recurrentes introducidas en la red y la segunda es la red totalmente recurrente la cual no posee restricción en cuanto a la consideración de conexiones recurrentes en la red, ambos tipos de red poseen los mismos algoritmos de aprendizaje supervisado, característica fundamental ya que tiene como fin hacer predicciones.

Las redes recurrentes poseen un comportamiento dinámico lo cual facilita el tratamiento de información temporal o también llamados patrones dinámicos, siendo este tipo de redes el método indicado para abordar problemas que tengan patrones que dependen del tiempo en sentido de que

ese determinado patrón dependiese de un valor en instante anterior de tiempo, pudiéndose utilizar no solo a patrones dinámicos sino también a patrones estáticos es decir patrones en la que no interviene la variable del tiempo.

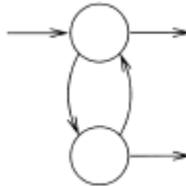


figura 11 Estructura básica de una red recurrente

cuando se habla de redes parcialmente recurrentes se pueden encontrar dos tipos conocidos de redes como los son la red de Elman y la segunda la red de Jordan, la diferencia radica en que las neuronas de contexto es decir neuronas especiales de entrada las cuales son las receptoras de las conexiones recurrentes y funcionan como memoria de red almacenando activaciones de una cierta capa de red reciben copias de diferentes capas de red, es decir, las red de Elman recibe copia de neuronas ocultas y la red de Jordan recibe una copia de las neuronas de la capa de salida y de sí mismas.

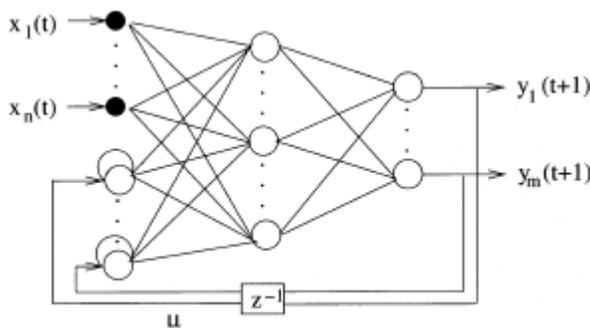


figura 12 Ejemplo Red de Jordan

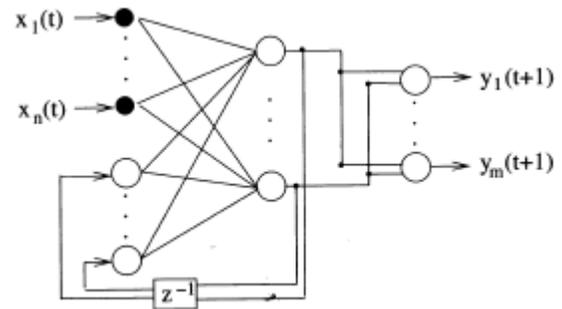


figura 13 Ejemplo Red de Elman

4.2.6 Validación Cruzada

Es una técnica utilizada principalmente en ambientes donde el objetivo principal es la predicción, donde se evalúa el comportamiento de los resultados de los análisis estadísticos en el modelo de división de datos, esta técnica obtiene un error de pronóstico para todas las observaciones disponibles con ajustes repetitivos en el modelo para las diferentes combinaciones en la selección de los conjuntos de entrenamiento y de prueba, todo esto para poder estimar la precisión con la que el modelo predictivo se desempeñara.

4.2.6.1 Validación Cruzada con Una Observación Fuera (LOOCV)

Como su nombre lo indica esta validación considera una sola observación (X_i, Y_i) como el conjunto de prueba (Ruiz Suarez, 2016) , dejándola fuera del ajuste del modelo, y prediciendo con el modelo ajustado. Este proceso es repetitivo en todo el conjunto de datos disponible, obteniendo así n valores distintos para la tasa de error, así se podrá estimar una tasa de error de clasificación promediando los errores en cada observación, esta técnica es ideal cuando no se trabajan datos masivos.

$$(V_n) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Tasa de error de clasificación LOOVC

4.2.6.2 Validación cruzada K-Fold Cross Validación

Esta técnica surge como alternativa cuando el conjunto de datos de observación es mucho más grande (Ruiz Suarez, 2016) , se deben considerar subconjuntos de prueba de mayor tamaño en comparación con LOOVC, el procedimiento para el tratamiento de los datos es de la misma forma para obtener la estimación del error de predicción para las observaciones disponibles, su

proceso inicia con la división de las observaciones en k sub-muestras aleatorias de un tamaño igual ajustando el modelo con $k-1$ de estas, prediciendo la restante, repitiendo el proceso k veces, de allí se estima la tasa de error de clasificación:

$$\hat{L}_{(k)} = \sum_{i=1}^k \frac{1}{k} L_i$$

Tasa de error de clasificación validación cruzada

5. Metodología

5.1 Selección De Datos

Se investigará el modo de recopilación de datos de las estaciones hidroclimáticas y se observaran diferentes fuentes de datos de cualquiera de las estaciones que tengan a disposición una serie de datos recopilados con variable del flujo y caudal que sean útiles para el desarrollo del modelo que se propone. Estos datos serán de prueba, pero se deberá observar que los datos entrenados concuerden con datos reales.

5.2 Selección De Los Modelos

Para la selección de los modelos que se trabajaran en el proyecto se deberá analizar como primera instancia los tipos algoritmos derivados del Machine Learning, y con esto entender el tipo de problema el cual se va a tratar, y así hacer uso de los modelos que estén estudiando que sean óptimos al desarrollo del modelo hidrodinámico.

5.3 Selección De Las Herramientas

Se investigarán frameworks que estén disponibles en el campo de la investigación, además de analizar que sea óptimo para la aplicación de modelos del Machine Learning y que puedan ser aplicados a series temporales, para ello se tendrá en cuenta que no tengan limitaciones en ejecución y consumo de recursos que puedan afectar el desarrollo del modelo.

5.4 Aplicación De Los Modelos

Una vez se tenga claro el framework con el cual se va a trabajar se deben analizar los datos fuente que se tengan a disposición, y se iniciara con el desarrollo del modelo con el cual se desea predecir el comportamiento de los datos, utilizando los modelos investigados y aplicando los procesos correspondientes para adquirir los datos resultantes.

5.5 Evaluación De Resultados

Se analizarán los resultados que se obtengan en las pruebas de cada modelo desarrollado, observando sus gráficas que se vayan obteniendo durante el proceso de pruebas y analizar el comportamiento de cada una de estas para poder describir la que mejor se ajusta a los parámetros que se le ingresen.

6. Desarrollo del modelo

En este trabajo se desea plantear la siguiente metodología a través de actividades para desarrollar un modelo de predicción de caudales teniendo en cuenta los objetivos propuestos, para alcanzar los objetivos del proyecto se plantea una metodología basado en etapas la cual se verá reflejada analizando cada uno de los objetivos planteados , iniciando desde la búsqueda de una base de datos que contengan variables hidrológicas y la elección de MATLAB como software adecuado para la ejecución de redes neuronales.

A continuación, se mostrarán ítems para los objetivos planteados en el proyecto y plantearán etapas que se verán involucradas en cada uno de estos.

6.1 Metodología para objetivo 1

Seleccionar y evaluar tres modelos de redes neuronales con capacidad para la predicción de series de tiempo.

- Se investigan redes neuronales que sirvan para la evaluación y predicción de series de tiempo que posean la característica de poder ser implementadas en MATLAB y que puedan trabajar específicamente con series de tiempo.

Inicialmente antes de buscar los tipos de redes neuronales que pueden ser útiles para aplicarlos en el proyecto se hace necesario investigar y tener claro el funcionamiento de una red neuronal y saber cómo es su modo de operación para que se conozcan los tipos de resultados que se van a obtener, para poder hacer un análisis adecuado al tipo de problema que se va a resolver.

- Evaluación y selección de diferentes arquitecturas de redes neuronales que no tengan como limitación el tiempo de procesamiento para que sean viables en las simulaciones y que cumplan las expectativas del modelamiento al trabajarlas con los datos de prueba.

Al tener un enfoque hacia las redes neuronales se hace necesario investigar a fondo que clase de algoritmos se pueden encontrar y que satisfagan la condición propuesta, y se encuentra que el aprendizaje automático o también llamado machine learning rama de la inteligencia artificial alberga una amplia gama de algoritmos, dependiendo del problema

se pueden agrupar en dos tipos, estos pueden ser: Problemas de regresión y Problemas de clasificación , según estos el que más se ajusta a los objetivos es el problema de regresión, ya que las variables a predecir son numéricas y los problemas de clasificación no lo son.

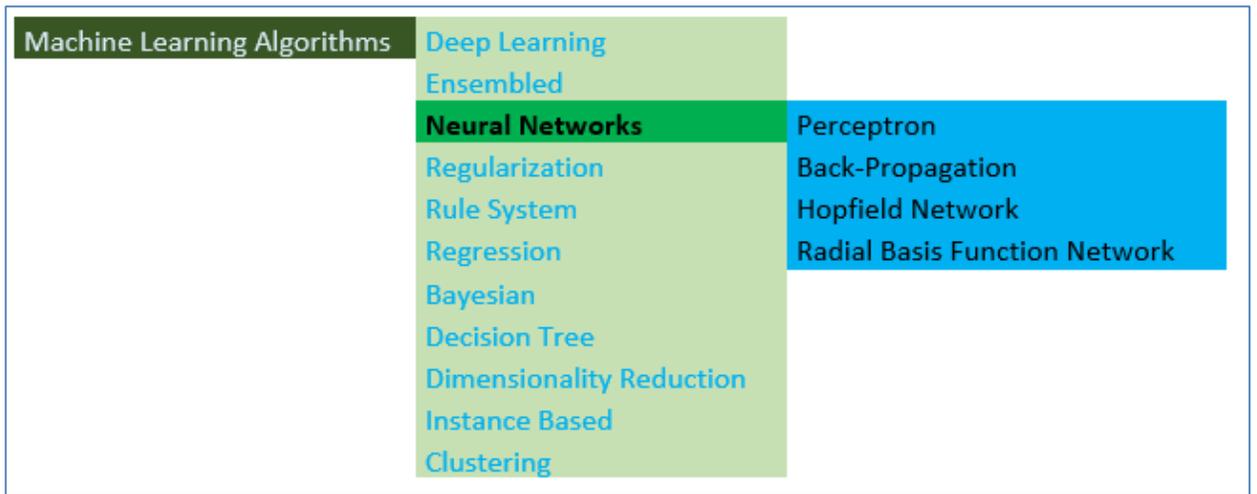


figura 14 Ramas del machine Learning

al tener definido el problema a afrontar, se define la solución utilizando uno de los modelos del machine Learning como, por ejemplo: Modelo Lineal, Modelo de Árbol, y Modelo de Redes Neuronales.

Se escoge el modelo de redes neuronales, y se observa que este modelo alberga varios algoritmos como los son: Radial Basis Function Network (RBFN), Perceptron, Back-Propagation y Hopfield Network.

6.2 Metodología para objetivo 2

Construir un modelo hidrodinámico para la predicción de caudales medio en ríos utilizando las variables hidroclimáticas disponibles en la base de datos de prueba.

- Análisis de la base de datos de prueba la cual posee parámetros tomados de estaciones hidrológicas, para iniciar posteriormente con la identificación de las variables hidroclimáticas adecuadas para el procesamiento de datos, utilizando los tipos de redes neuronales clasificadas y tener en cuenta la normalización de los datos para tener definido el rango a trabajar y analizar.

6.2.1 Toma de datos

Inicialmente se debe tener en cuenta como son recolectados los datos, para ello se utilizan estaciones hidrométricas (Marbello Pérez) las cuales son reglas que se ubican de forma escalonada en una de las márgenes del curso del agua y sea un río, arroyo, laguna o embalse, las cuales miden la cantidad de agua que está a disposición en los cuerpos de agua que dieron origen las lluvias y los escurrimientos, al no tener una estación hidrométrica al alcance, se recurre a la búsqueda de datos disponibles en la web los cuales se utilizaran como ejemplo en la aplicación del modelo.

Los datos históricos de prueba corresponden a una de muchas de las estaciones hidrométricas de recolección de datos en los cuales obtienen datos ya sean de medición del flujo o caudal.

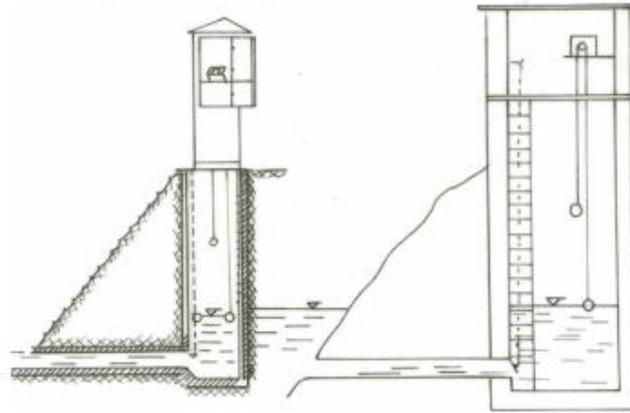


figura 15 Estructura de Estación Hidrométrica
Fuente: (Marbello Pérez)

Como datos de prueba se utilizaran los datos de recolección de la estación hidrométrica llamada GLENAMOY, estación la cual se encuentra en uno de los afluentes en Irlanda, estos datos están disponible en la web; Los datos que se tomaron corresponden a dos series de tiempo la primera tiene la característica de ser una serie de datos que fueron registrados cada 15 minutos durante 3 meses con una totalidad de 8531 filas con los datos de registro de cada medición, de las cuales cada una posee , la fecha y el momento en que fue tomada la medición además de tener el valor de la medición, la segunda serie corresponde a una serie que contiene 365 datos , cada fila tiene el valor de caudal tomado durante 1 año.

#Station Name	GLENAMOY									
#Station Number	33001									
#Station Paramete	River Discharge									
#Timeseries Name	Cmd.Base									
#Unit Symbol	m ³ /s									
#Rows	8531									
#Timestamp	Value	Quality Code	Name	Dia	Mes	Anio	Hora	Min	Caudal	
03/01/2018 0:00	23.5	Unchecked		03/01/2018 0:00	3	1	2018	0	0 23.5	

figura 16 Encabezado datos de la estación Glenamoy datos trimestral

#Station Name	GLENAMOY	
#Station Number	33001	
#Station Parameter Name	River Discharge	
#Timeseries Name	Day.Mean	
#Unit Symbol	m ³ /s	
#Rows	365	
#Timestamp	Value	Quality Code Name
	04/04/2017 1:00	1.78

figura 17 Encabezado datos de la estación Glenamoy datos anual

6.2.2 Herramienta de modelado

Se escoge para el desarrollo del modelo al Framework MATLAB (Demuth & Beale, 1998) ya que dentro de sus características este software resuelve problemas científicos y de ingeniería, está basado en matrices ya que es la forma natural de expresión de las matemáticas computacionales, además de tener características para la visualización de datos y obtener información a través de estos.

Matlab posee un conjunto de biblioteca de herramientas llamadas toolbox de forma integrada las cuales poseen características especiales para la resolución de problemas, se escoge una de estas toolbox la cual tiene la característica de trabajar con redes neuronales, herramienta útil en la investigación, los tipos de redes neuronales que permite el desarrollo y el entrenamiento con esta herramienta (N. Koivo, 2008) son por ejemplo redes de tipo: Perceptron, Adaline, Backpropagation, Redes de Base Radial, SOM, Elman, Hopfield, LVQ y que concuerdan con las redes que derivan del machine learning como lo indica la **figura 14**.

Los comandos importantes a utilizar en el desarrollo del modelo son los siguientes:

- **Newff:** Crea una red feed-forward backpropagation.
- **Newrb:** Diseña una red de base Radial.
- **Cvpartition:** Crea una partición de validación cruzada para los datos.

- **Sim:** Simula una red neuronal.
- **Train:** Entrena la red neuronal.

El algoritmo a utilizar en los entrenamientos es el de Levenberg-Marquardt (TRAINLM) MATLAB, ya que este algoritmo parece ser el método más rápido para entrenar redes neuronales de avance moderado (hasta varios cientos de pesos). También tiene una implementación en Matlab muy eficiente, ya que la solución de la ecuación matricial es una función integrada, por lo que sus atributos se vuelven aún más pronunciados en una configuración de Matlab (Demuth & Beale, 1998).

Las funciones que se utilizan para poder crear una red neuronal poseen algunas diferencias al momento de trabajar independientemente con cada algoritmo, por ejemplo, para trabajar con una red backpropagation necesita de los siguientes parámetros:

6.2.2.1 Parámetros para utilizar la función newff para el modelo con Backpropagation

Para la creación de la red neuronal se utiliza el comando newff (Neural Network Toolbox), con algunos parámetros unos obligatorios, y otros no se agregan ya se ejecutan por defecto, su estructura es la siguiente:

newff(PR,[S1 S2...SNI],{TF1 TF2...TFN1},BTF,BLF,PF)

- **PR:** Rx matriz de valores mínimos y máximos para R-elementos de entrada.
- **Si:** Tamaño de la capa i, por N-capas.
- **TFi:** Función de transferencia de la capa ith, por defecto = 'tansig'.
- **BTF:** Función de entrenamiento de la red de backpropagation, por defecto = 'traingdx'.
- **BLF:** Función de aprendizaje de peso / sesgo de contrapropagación, por defecto = 'learnngdm'.
- **PF:** Función de rendimiento, por defecto = 'mse'.

6.2.2.2 Parámetros para utilizar la función newrb para el modelo con red de base radial

Para la creación de la red neuronal se utiliza el comando newrb (Neural Network Toolbox), con algunos parámetros unos obligatorios, y otros no se agregan ya se ejecutan por defecto, su estructura es la siguiente:

newrb(P,T,goal,spread,MN,DF)

- **P:** Matriz R-por-Q de vectores de entrada Q.
- **T:** Matriz S-by-Q de vectores de clase objetivo Q.
- **Goal:** Objetivo de error cuadrático medio (por defecto = 0.0).
- **Spread:** Propagación de funciones de base radial (por defecto = 1.0).
- **MN:** Número máximo de neuronas (el valor predeterminado es Q).
- **DF:** Número de neuronas para agregar entre pantallas (predeterminado = 25).

6.2.3 Modelo de predicción hidrológico

Para el desarrollo del proyecto se opta por trabajar con el modelo autorregresivo (García Martos, 2012) y modelos autorregresivos con variable exógena, para ello se tiene que tener en cuenta que un modelo autorregresivo es un modelo de regresión en el que las variables explicativas, son la misma variable dependiente retardada.

Un ejemplo sencillo de un proceso autorregresivo es el siguiente:

- Modelo autorregresivo estacionario: $X_t = \theta X_{t-1} + a^t$
- Modelo autorregresivo no estacionario: $X_t = X_{t-1} + a^t$

Esto indica que X_{t-1} es conocido en t y a^t no es conocido. a^t constituye el valor de innovación.

Se debe tener en cuenta que tipo de modelo se debe implementar, para ello se observa los diferentes modelos estacionarios con el modelo no estacionario que puedan servir para el análisis posterior de la serie temporal, y haciendo un análisis de los modelos se puede definir que se utilizaran los modelos con condición de estacionariedad, esta condición significa que los shocks pasados afectan cada vez menos sobre los valores presentes de la serie temporal.

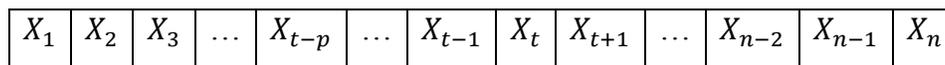
6.2.3.1 Tipo de modelos estacionarios a implementar.

- **Modelo Autorregresivo. AR.:** Se utilizará este modelo ya que la condición se ajusta a las necesidades de utilizar valores pasados y que este solo dependa de los valores que se tengan a disposición, y que ninguna otra variable externa esté presente en el modelo a comparación de los modelos autorregresivos de media móvil que dependen de valores actuales y variables externas estocásticas que afecten el comportamiento de la implementación del modelo

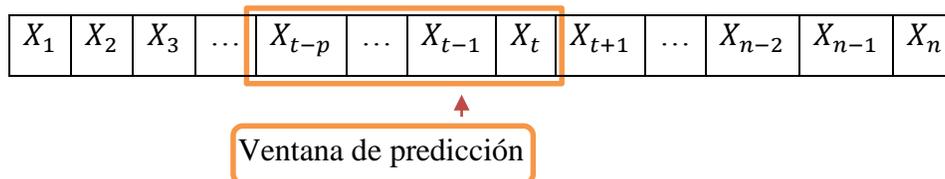
Modelo autorregresivo a implementar:

$$X_t = \theta_0 + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \dots + \theta_p X_{t-p} + \theta_t$$

De forma gráfica se puede observar cual será el comportamiento que tendrá el modelo de predicción como se muestra a continuación:



Lo que nos dice el modelo autorregresivo es que se tomaran datos pasados de una serie de tiempo y según estos datos pasados se podrá determinar el valor futuro a predecir para ello se tomara un conjunto de datos llamado Ventana de predicción que será un conjunto de datos que de desee tomar y el valor futuro será determinado según el modelo:



El valor X_{t+1} dependerá de los datos pasados según el siguiente comportamiento:

$$X_{t+1} = F(X(t), X(t - 1), X(t - 2), \dots, X(t - p))$$

El comportamiento que tendría este modelo para determinar el valor X_{t+1} en una red neuronal será el siguiente:

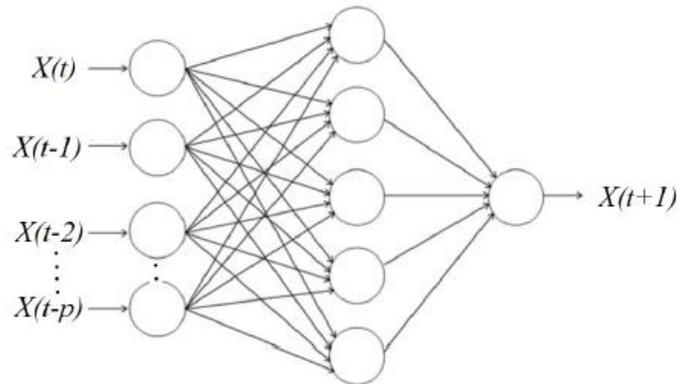


figura 18 Entrada de datos según el modelo autorregresivo

es decir que la red dependerá de ese conjunto de datos que se escojan para poder determinar la variable a predecir, estas entradas tendrán asignados sus pesos y seguido iniciara con el entrenamiento y el aprendizaje como característica esencial de la red.

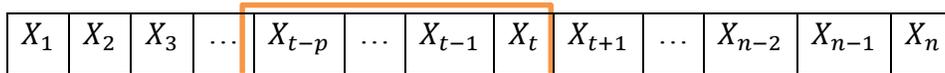
El comportamiento que tendrá la red para predecir dicho valor X_{t+1} esta determinado por:

$$X_{t+1} = F(X(t), X(t - 1), X(t - 2), \dots, X(t - p))$$

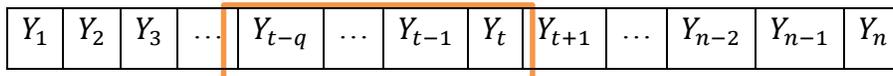
- **Modelos Autorregresivos con entrada exógena ARX:** este otro modelo es útil cuando la variable a predecir no solo dependerá de valores pasados sino también dependerá de otra ventana de predicción con variables exógenas cuyos valores tengan relación con los datos utilizados en la ventana de predicción principal.

Para entender este método se puede utilizar como ejemplo a (Agudelo, López Lezama, & Velilla, 2015) trabajo se basó en predicción se desea determinar el precio de energía de la bolsa, teniendo como entradas exógenas la demanda, la relación, generación hidráulica-térmica, fenómeno del niño y volumen diario de energía.

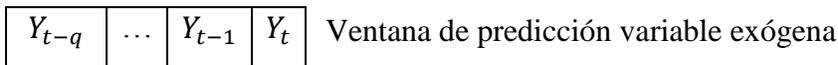
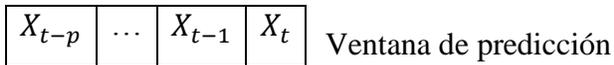
La representación del modelo de autorregresión con entrada exógena será de la siguiente manera:



Conjunto de datos principales con su ventana de predicción



Conjunto de datos secundarios con su ventana de predicción



Estos determinan el valor a predecir X_{t+1} como se muestra en el procedimiento:

$$X_{t+1} = F(X(t), X(t - 1), X(t - 2), \dots, X(t - p), Y(t), Y(t - 1), Y(t - 2), \dots, Y(t - q))$$

6.2.4 Implementación con función Backpropagation y validación cruzada

Para implementar redes con arquitectura de red feedforward backpropagation se utiliza la toolbox de redes neuronales de Matlab, y específicamente la estructura de red con función newff de la cual ya se conoce el tipo de parámetros que recibe dicha función para la creación de la red neuronal.

Antes de iniciar las pruebas se debe tener a disposición los datos hidroclimáticos que están disponibles en la web con los cuales se va a trabajar, inicialmente se guardan en un tipo de archivo .MAT para su posterior utilización en cualquiera de los datos, para la interpretación de los datos se hace necesario antes de aplicar cualquier función, la normalización de los datos para que estos puedan ser analizados de forma apropiada cuando se obtengan resultados de forma gráfica, de igual manera en la interpretación de los error que arrojen las pruebas, graficando los resultados obtenidos.

6.2.4.1 Experimento n° 1

Se implementa un modelo de red neuronal con una red de retropropagación tipo Perceptrón Multicapa. Se utilizan la serie de datos de Caudal, en un periodo trimestral, con intervalos de muestreo de 15 minutos.

En la **figura 19** se muestran los resultados de entrenar el modelo para determinar la capacidad de predecir el Caudal de agua en el tiempo $(t+1)$, utilizando como entrada el valor del Caudal en el tiempo t . Para determinar la estabilidad del modelo, se realiza la simulación para 20 entrenamientos. Para cada uno de ellos además se realiza validación cruzada, con 10 particiones. Se emplea una red neuronal con 2 capas ocultas y 10 neuronas en cada capa, experimentos preliminares muestran que, aumentando la complejidad de la red, no se mejoran considerablemente los resultados, pero si incrementa el tiempo de entrenamiento, por lo cual se ha decidido fijar dicho valor, los resultados de estas pruebas se pueden encontrar en la sección 6.3.1.1 del libro, para observar el comportamiento de las pruebas con los ajustes iniciales.

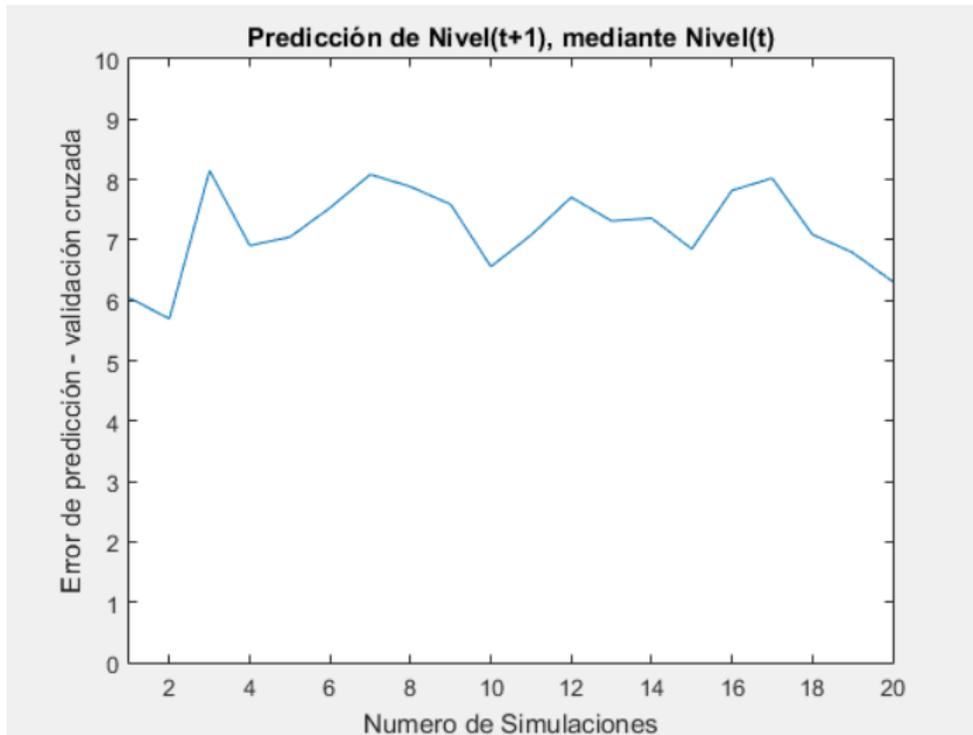


figura 19 Error de predicción del Caudal($t+1$)

Resultado del experimento 1 donde se grafica el error de predicción del caudal $t + 1$ dado el caudal t , para cada una de las 20 simulaciones que realizó la prueba, a continuación se muestra. En la **figura 20** se muestra en color azul la salida de una de las redes entrenadas, en color rojo se muestran los valores originales de la serie. Se puede observar que las dos gráficas se ajustan con un mínimo error.

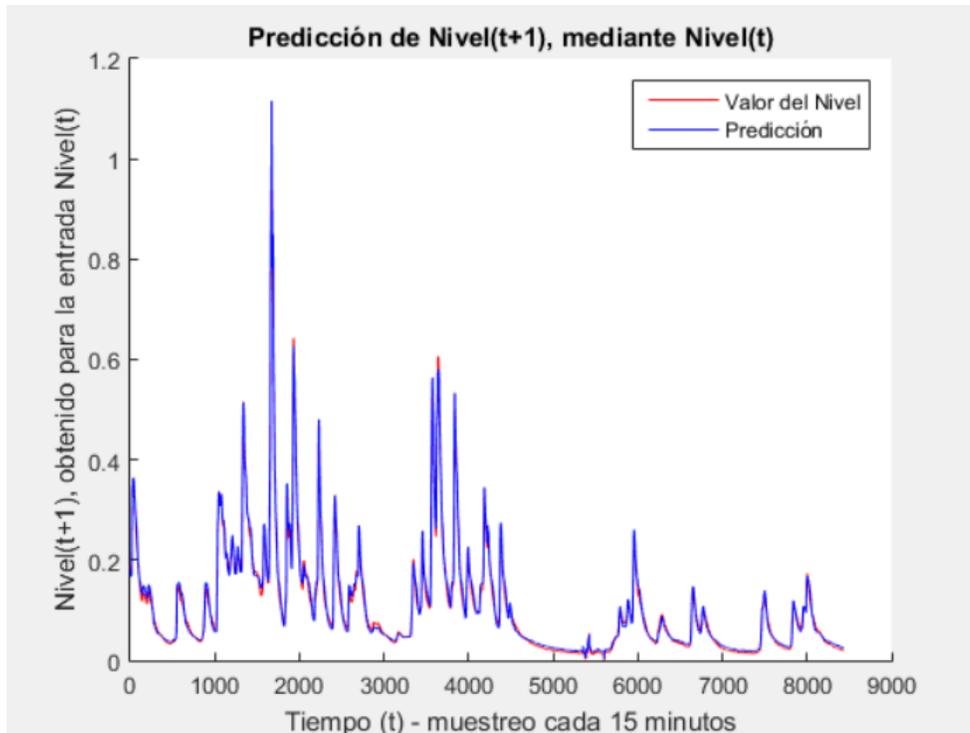


figura 20 Predicción del caudal (t+1)

Como resultado de esta prueba se puede observar la predicción del caudal (t+1) dado un caudal (t) comparada con la serie original del caudal.

6.2.4.2 Experimento n° 2

Para determinar la influencia que puedan tener los datos históricos de la serie para predecir el Caudal en el tiempo (t+1). Se realiza el entrenamiento del modelo con la red de retropropagación utilizando como conjunto de entrada, valores con diferentes ventanas de retraso de la siguiente manera.

- **Predicción para ventana 1:** Se trata de predecir Caudal en $t+1$, con los datos de Caudal observados en tiempo t , $t-1$ y $t-2$.

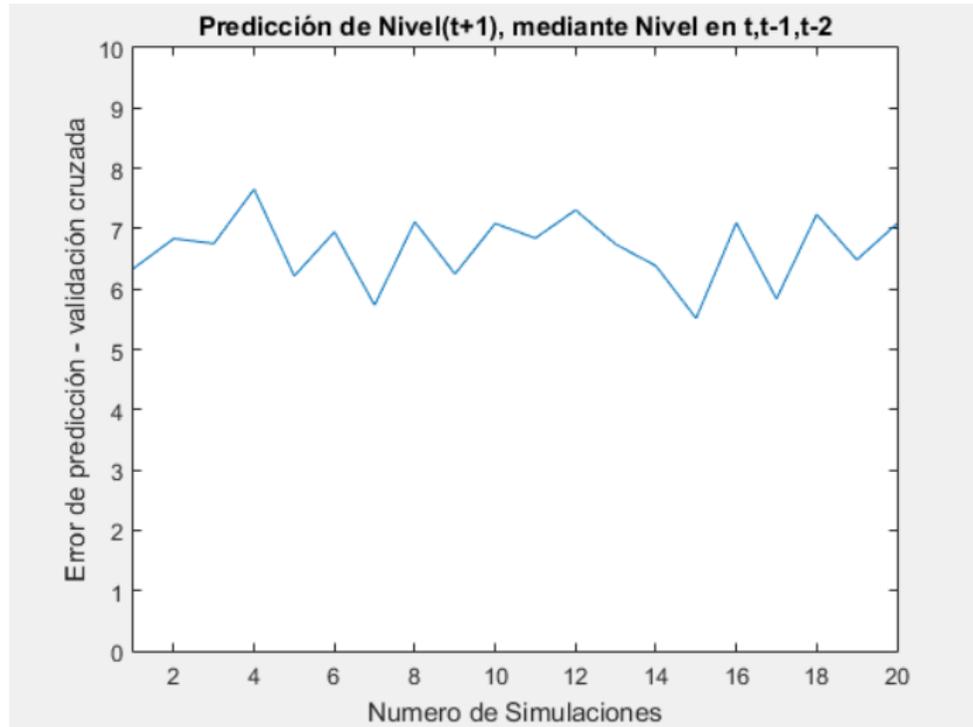


figura 21 Error de predicción del caudal ($t+1$)

como se observa se obtiene como resultado el error de predicción del caudal ($t+1$), para ventana de predicción con el caudal en los t , $t-1$ y $t-2$.

A continuación, se grafica la serie de tiempo, de azul se muestra la salida de la red entrenada y de rojo la serie original.

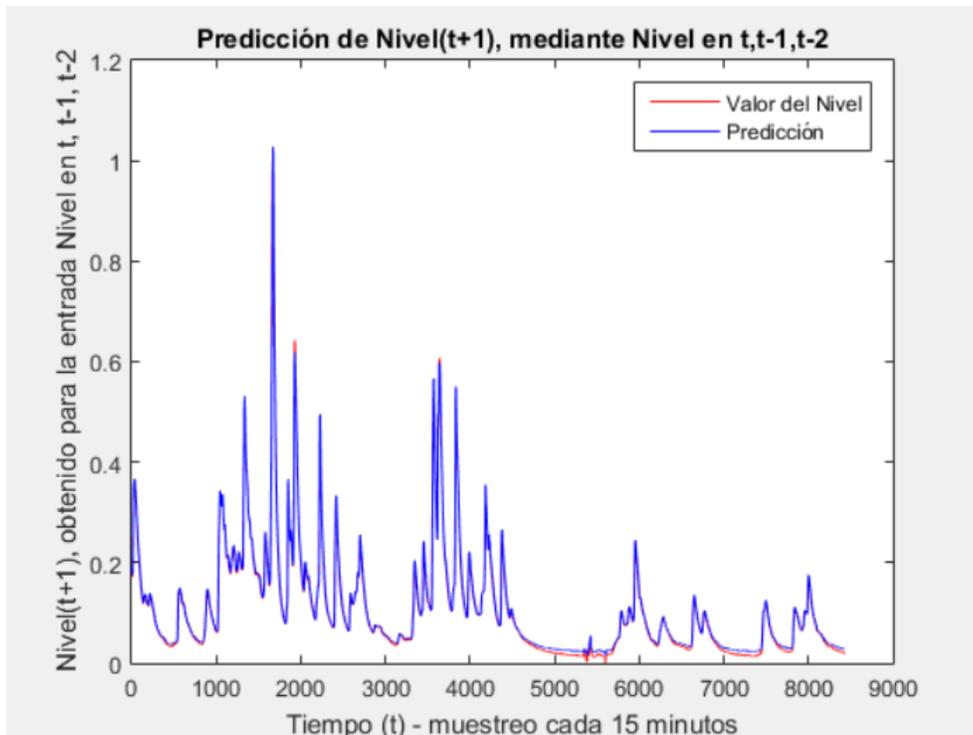


figura 22 Predicción del caudal (t+1)

Como resultado se puede observar la predicción del caudal ($t+1$) dado el caudal en t , $t-1$ y $t-2$ comparada con la serie original del caudal.

- **Predicción para ventana 2:** Se trata de predecir Caudal en $t+1$, con los datos de Caudal observados en tiempo t , $t-1$, $t-2$, ..., $t-5$. Los valores obtenidos para el error se muestran en la **figura 23**.

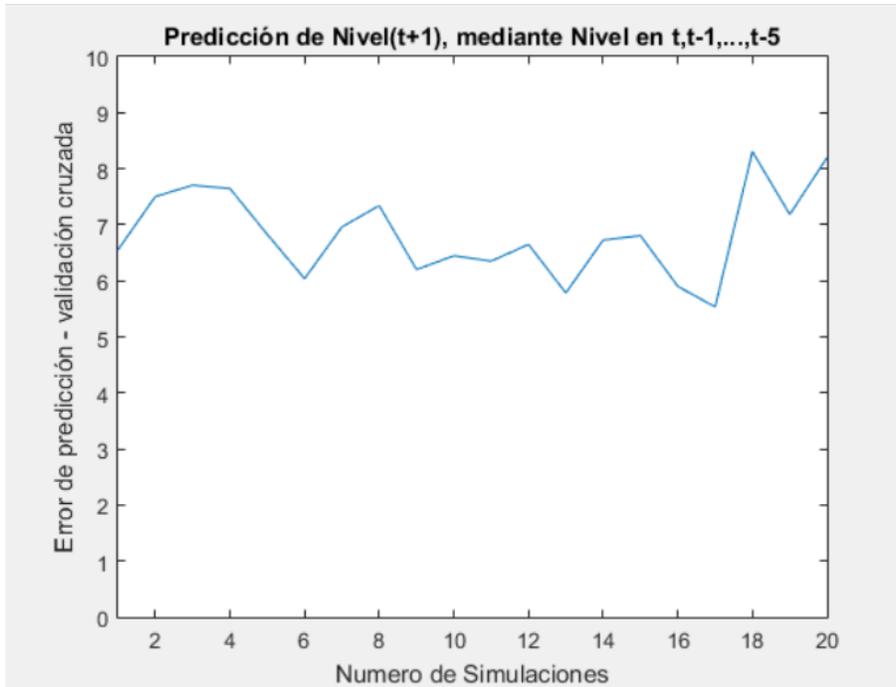


figura 23 Predicción del caudal ($t+1$)

como se observa se obtiene como resultado el error de predicción del caudal ($t+1$), para ventana de predicción con el caudal en los $t, t-1, \dots, t-5$, obteniendo un error medio del 6.8330 %.

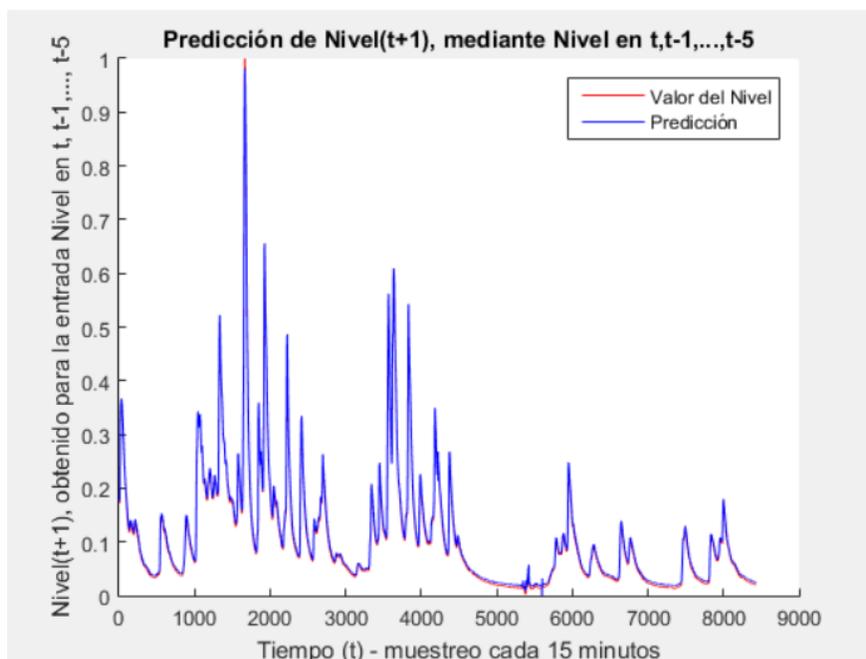


figura 24 Predicción del caudal ($t+1$)

En la **figura 24** se observa la predicción del caudal ($t+1$) dado el caudal en $t, t-1, \dots, t-5$, comparada con la serie original del caudal.

- **Predicción para ventana 3:** Se trata de predecir Caudal en $t+1$, con los datos de Caudal observados en tiempo $t, t-1, t-2, \dots, t-10$. La **figura 25** muestra el error de predicción y la serie obtenida de la última red entrenada.

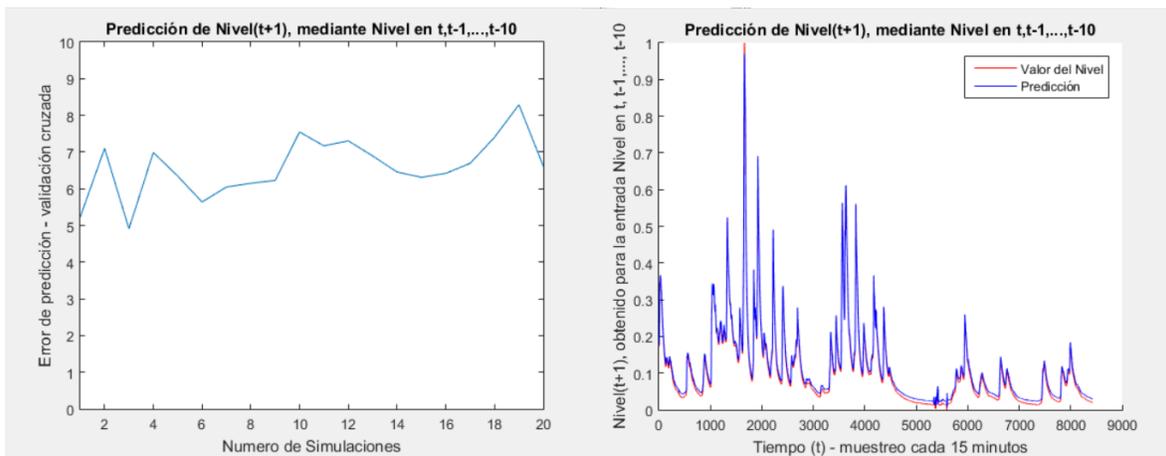


figura 25 Error de predicción del caudal ($t+1$) y Predicción del caudal ($t+1$).

Se observa como resultado el error de predicción del caudal ($t+1$), para ventana de predicción con el caudal en $t, t-1, \dots, t-10$, y predicción del caudal ($t+1$), dado el caudal en $t, t-1, \dots, t-10$, comparada con la serie original.

- **Predicción para ventana 4:** Para este último caso del primer experimento se trata de predecir Caudal en $t+1$, con los datos de Caudal observados en tiempo $t, t-1, t-2, \dots, t-20$. La **figura 26** muestra el error para cada una de las 20 simulaciones y la serie obtenida con la última red entrenada.

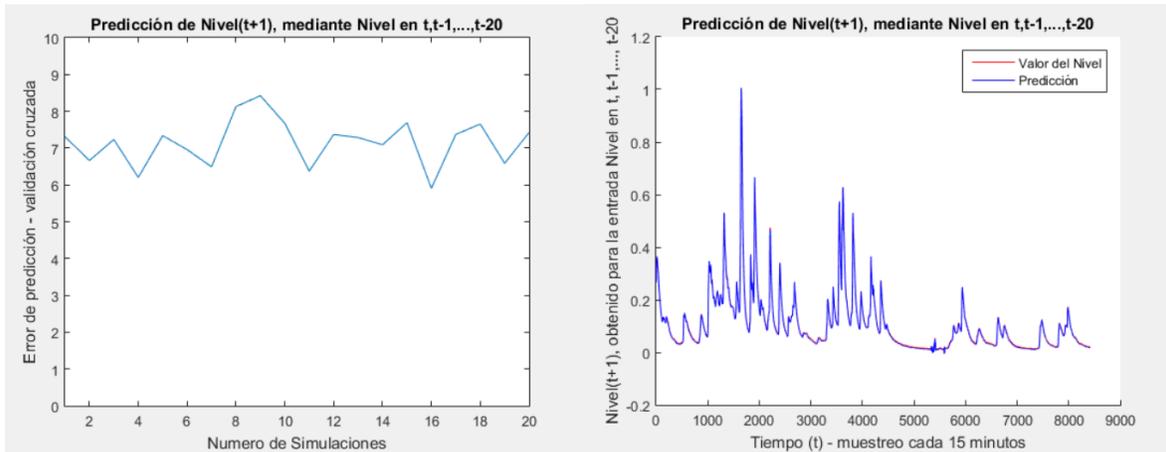


figura 26 Error de predicción del caudal ($t+1$)

Como resultado se puede observar el error de predicción del caudal ($t+1$) para la ventana de predicción del caudal en $t, t-1, \dots, t-20$, y predicción del caudal ($t+1$), dado el caudal en $t, t-1, \dots, t-20$, comprada con la serie original del caudal.

6.2.4.3 Experimento n° 3

En los experimentos anteriores se puede observar que el entrenamiento del modelo con entradas correspondientes a ventanas con grandes retrasos en el tiempo no tiene mayor influencia en la disminución del error de predicción, por el contrario, el error ha aumentado levemente lo que puede ser interpretado como que los datos históricos añaden ruido en la predicción. Por ello se fija una ventana de retrasos para tiempos $t, t-1, t-2, t-5$, y se trata de predecir el comportamiento para diferentes valores a futuro de la siguiente manera:

- **Predicción tiempo futuro t+2:**

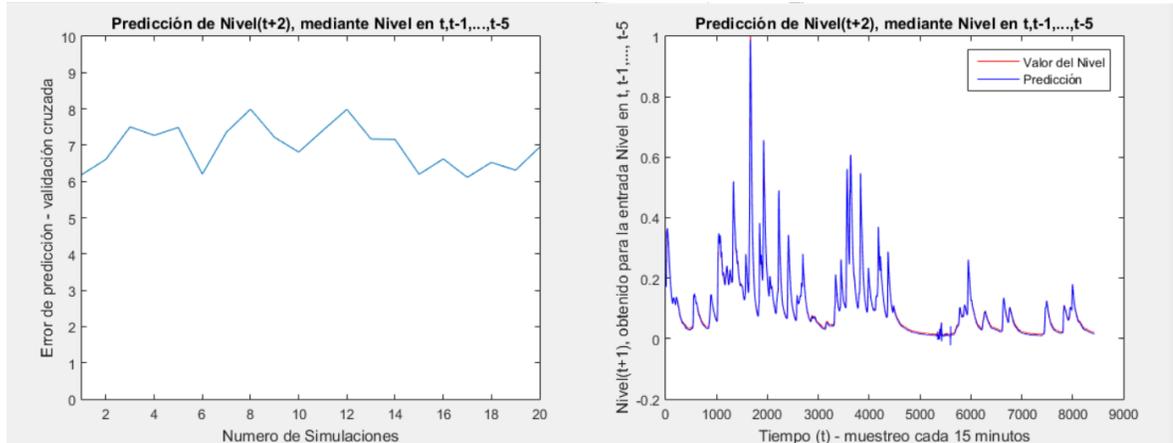


figura 27 Error de predicción del caudal (t+2) y Predicción del caudal (t+2).

La **figura 27** muestra los resultados obtenidos, primero muestra el error de predicción del caudal ($t+2$) para una ventana de predicción $t, t-1, \dots, t-5$, junto con la gráfica de predicción del caudal ($t+2$), dado el caudal de agua en los tiempos $t, t-1, \dots, t-5$ como entrada de la red, donde esta comparada con la serie original del caudal, arrojando resultados después de 20 simulaciones se obtiene un error del 6.9539 %.

- **Predicción tiempo futuro t+5:**

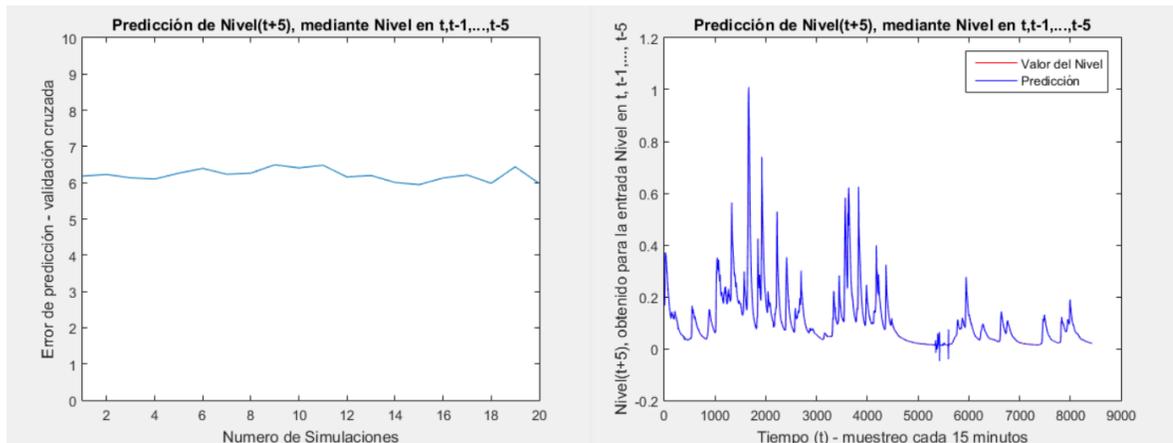


figura 28 Error de predicción del caudal (t+5) y Predicción del caudal (t+5)

La **figura 28** muestra los resultados obtenidos para un error de predicción del caudal ($t+5$), para la ventana de predicción con el caudal en $t, t-1, \dots, t-5$, junto con la predicción del caudal en ($t+5$), dado el caudal en $t, t-1, \dots, t-5$, comparándola con la serie original del caudal, en el cual se obtiene un error medio de 6.2131%.

El error vario levemente entre cada una de las simulaciones realizadas, en comparación con los casos anteriores, esta prueba se realizó varias veces para descartar cambios drásticos en el comportamiento del error, y se obtuvieron resultados similares.

- **Predicción tiempo futuro $t+10$:**

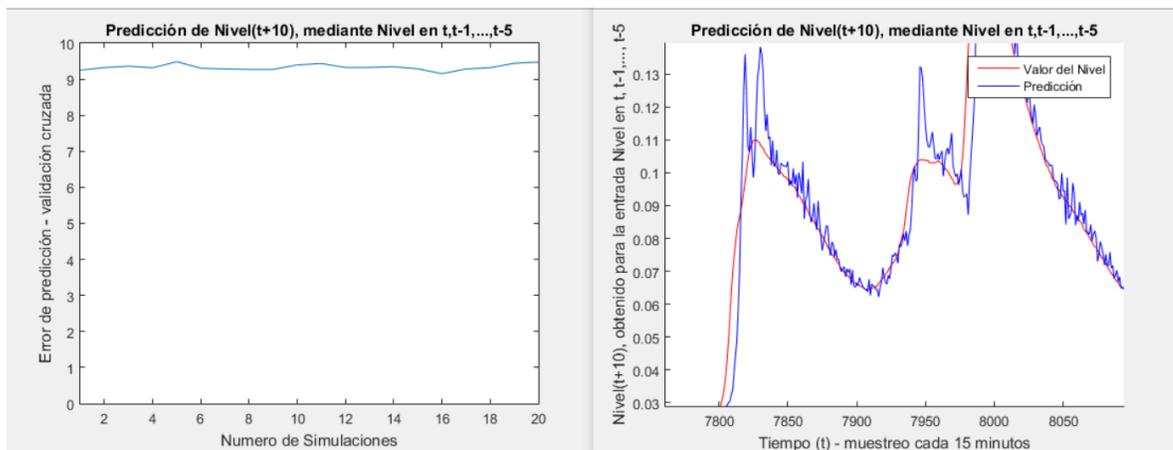


figura 29 Error de predicción del caudal ($t+10$) y Predicción del caudal ($t+10$)

La **figura 29** muestra los resultados obtenidos para un error de predicción del caudal ($t+10$), para la ventana de predicción con el caudal en $t, t-1, \dots, t-5$, junto con la predicción del caudal en ($t+10$), también se aprecia el error obtenido en cada una de las iteraciones, así como la gráfica obtenida por la simulación de la red, comparada con la serie original. Se ha realizado zoom sobre la imagen para observar en detalle las diferencias entre el Caudal que la red predice, y el Caudal en la serie original. La predicción tiene un cierto ruido añadido producto del error de predicción. Sin embargo, se observa que se mantienen la tendencia y estacionalidad de la serie, el error medio obtenido en esta prueba es 9.3329 %.

- **Predicción tiempo futuro $t+15$:**

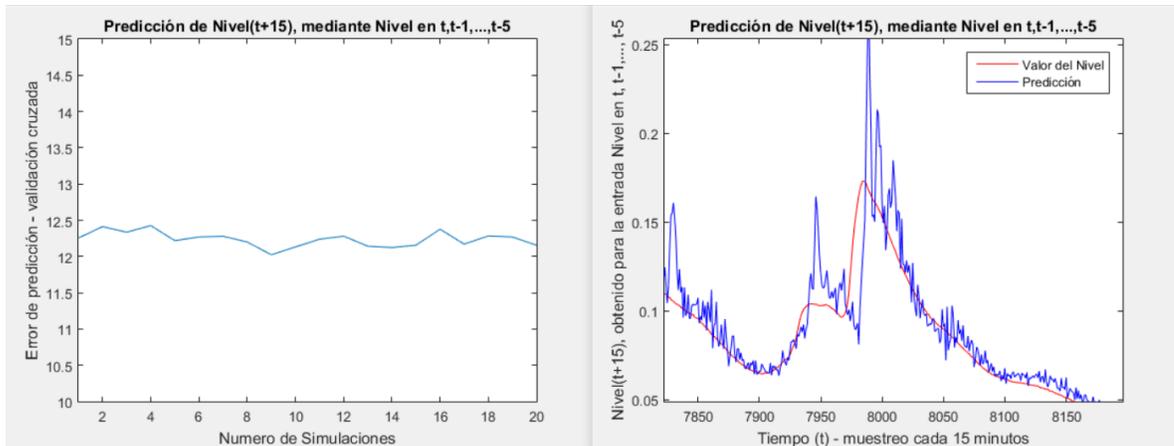


figura 30 Error de predicción del caudal ($t+15$) y Predicción del caudal ($t+15$)

La **figura 30** muestra los resultados obtenidos para un error de predicción del caudal ($t+15$), para la ventana de predicción con el caudal en $t, t-1, \dots, t-5$, junto con la predicción del caudal en ($t+15$), al igual que la gráfica anterior se hizo zoom para observar el comportamiento de la red resultante y se puede apreciar que existe mayor cantidad de ruido presente en la red por el error de predicción de la red, el error medio que se obtuvo en esta prueba es del 12.2376%.

- **Predicción tiempo futuro $t+20$:**

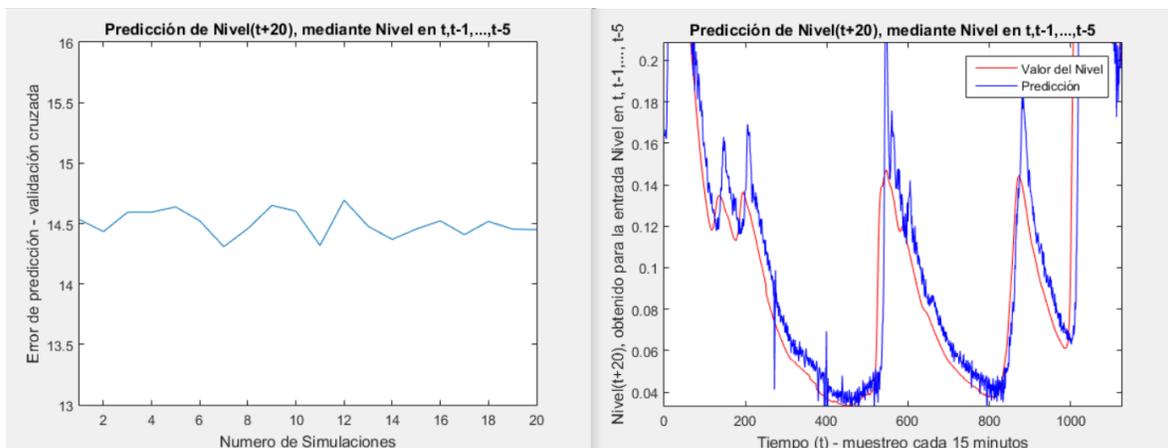


figura 31 Error de predicción del caudal ($t+20$) y predicción del caudal ($t+20$)

La **figura 31** muestra los resultados obtenidos para un error de predicción del caudal ($t+20$), para la ventana de predicción con el caudal en $t, t-1, \dots, t-5$, junto con la predicción del caudal en ($t+20$) comparándola con la serie original del caudal, se hace zoom para observar el comportamiento de la red resultante y se observa el aumento de ruido presente por el error de predicción, para esta prueba el error medio obtenido es del 14.4997%.

- **Predicción tiempo futuro $t+30$:**

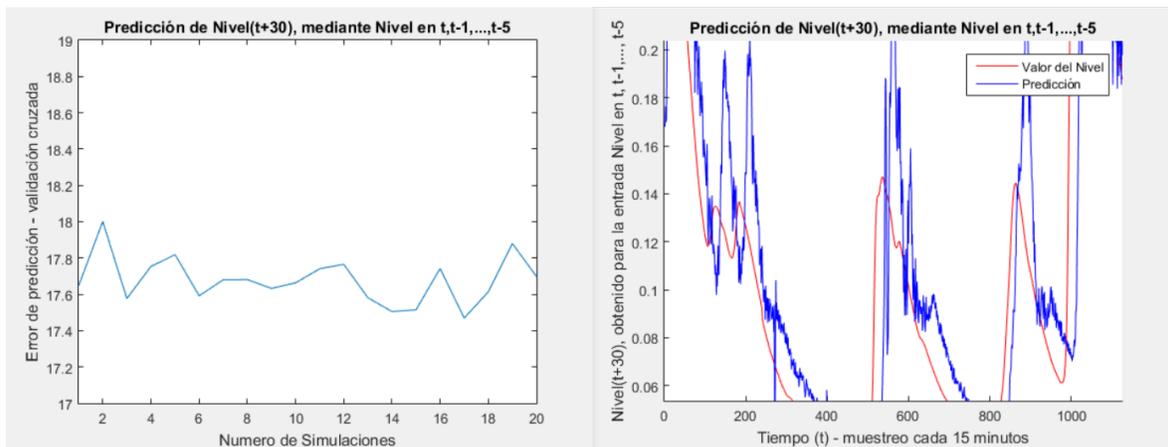


figura 32 Error de predicción del caudal ($t+30$) y predicción del caudal ($t+30$)

La **figura 32** muestra los resultados obtenidos para un error medio de predicción del caudal ($t + 30$), para la ventana de predicción con el caudal en $t, t-1, \dots, t-5$, junto con la predicción del caudal en ($t + 30$) comparándola con la serie original del caudal, donde se observa la distorsión de la red resultante por el ruido, la prueba arrojó un error de predicción del 17.6781%.

- **Predicción tiempo futuro $t+40$:**

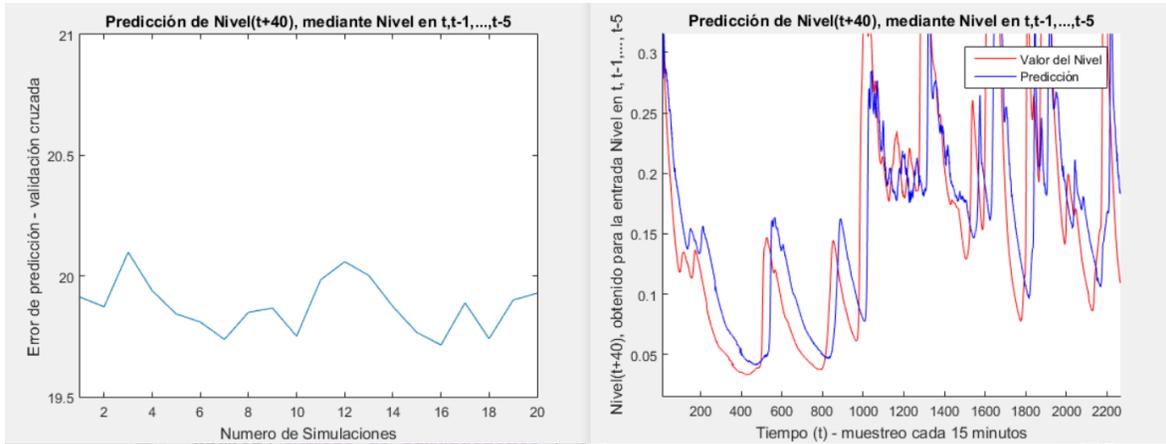


figura 33 Error de predicción del caudal ($t+40$) y predicción del caudal ($t+40$)

La **figura 33** muestra los resultados obtenidos para un error de predicción del caudal ($t+40$), para la ventana de predicción con el caudal en $t, t-1, \dots, t-5$, junto con la predicción del caudal en ($t+40$) comparándola con la serie original del caudal, aquí nuevamente se hace zoom sobre una parte de la red en la cual se puede detallar el desfase de la red resultante, de la cual se obtiene un error medio de 19.8771%.

- **Predicción tiempo futuro $t+50$:**

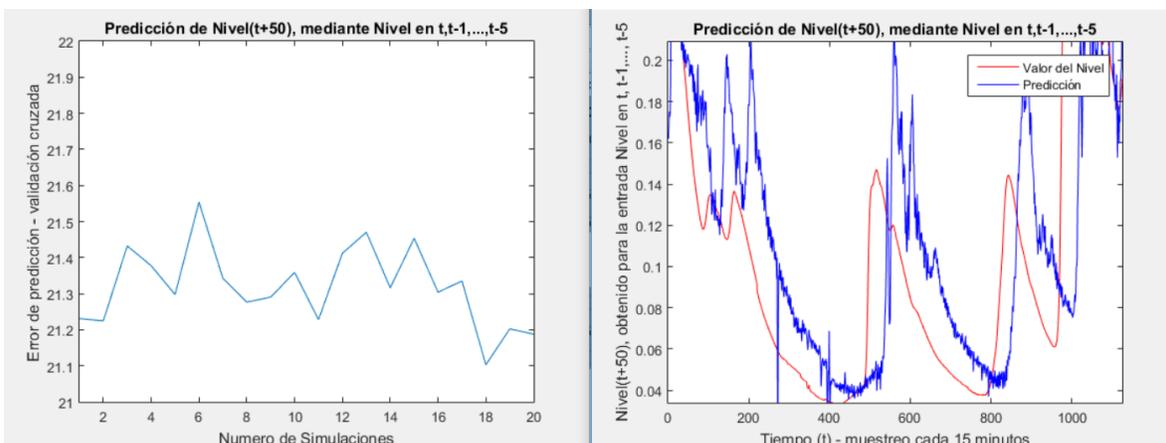


figura 34 Error de predicción del caudal ($t+50$) y Predicción del caudal ($t+50$)

La **figura 34** muestra los resultados obtenidos para un error de predicción del caudal ($t+50$), para la ventana de predicción con el caudal en t , $t-1$, ..., $t-5$, junto con la predicción del caudal en ($t+50$), La cual muestra como resultado después de 20 ejecuciones un error medio de predicción de 21.3199%.

La recopilación de los errores en las diferentes predicciones desde $t+1$ a $t+50$ con intervalos de tiempo definidos, para tiempo futuro son la siguientes:

Predicción	t+1	t+2	t+5	t+10	t+15	t+20	t+30	t+40	t+50
Error medio	6.5326%	6.9539%	6.2131%	9.3329%	12.2376%	14.4997%	17.6781%	19.8771%	21.3199%

Tabla 1 Error medio para predicciones tiempo futuro con una red de retropropagación

El comportamiento de los errores medio para la predicción del caudal en tiempos $t+i$ para $i = 1,2,4,10,15,20,30,40,50$ son como se muestran en la **figura 35**:

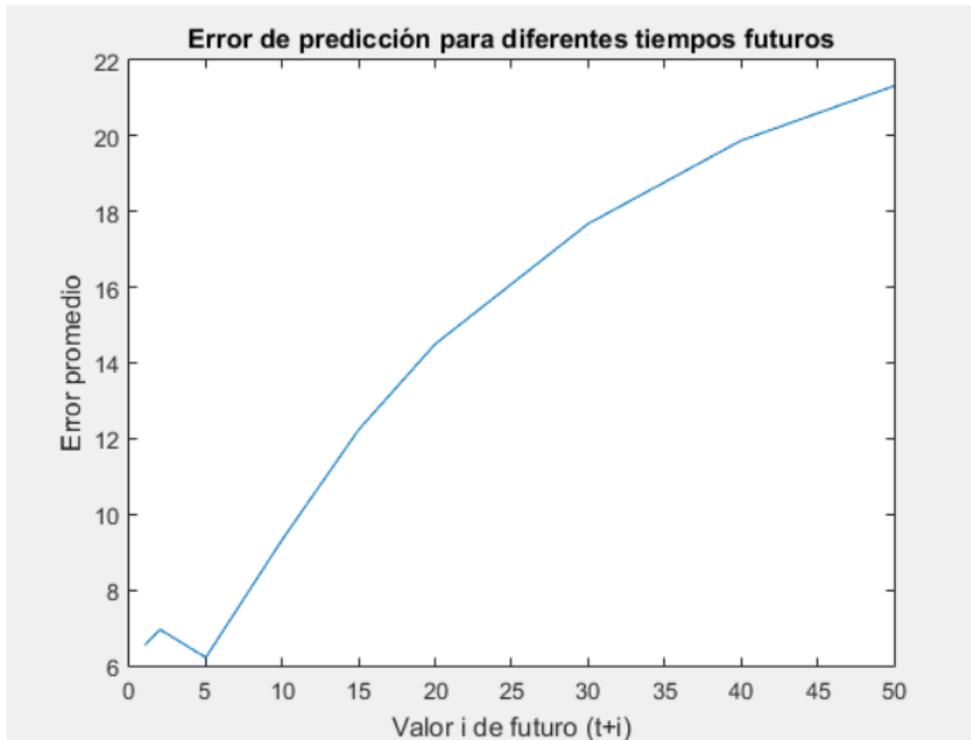


figura 35 Comportamiento error medio para predicción de caudal $t+i$

6.2.5 Implementación con función base radial y validación cruzada

Para implementar redes con arquitectura de red de base radial se utiliza la toolbox de redes neuronales de Matlab, y específicamente la estructura newrb de la cual ya se conoce el tipo de parámetros que recibe dicha función para la creación de redes neuronales, los siguientes experimentos se hacen con valores fijos después de realizadas las pruebas preliminares para observar el comportamiento del error con los parámetros por defecto que utiliza la red neuronal con función de base radial.

6.2.5.1 Experimento n° 4

Se realiza la predicción futura para los valores $t + i$, con $i = 1, 2, 5, 10, 15, 20, 30, 40$ y 50 , utilizando una red neuronal de función de base radial, la cual se caracteriza por un aproximador universal de funciones, cuyo entrenamiento suele darse en menor tiempo que el de un perceptrón multicapa, para este caso se utiliza un ancho de gaussiana de 0.15 . Este parámetro indica que se realiza una aproximación fina, lo cual requiere mayor número de iteraciones para reducir el error.

La siguiente tabla muestra los resultados obtenidos del error promedio para cada uno de los valores, teniendo en cuenta que se realizaron 20 simulaciones para cada valor de i .

Predicción	t+1	t+2	t+5	t+10	t+15	t+20	t+30	t+40	t+50
Error medio	18.564%	18.707%	19.056%	19.402%	20.045%	20.718%	22.093%	23.21%	24.001%

Tabla 2 Error medio para predicciones tiempo futuro con una red de base radial

En la tabla 2 se observan los resultados obtenidos en la simulación para diferentes tiempos futuros, donde se trata de predecir el valor del caudal en $(t + i)$ con un valor de $i = 1, 2, 5, 10, 15, 20, 30$ y 50 , de igual manera en la **figura 36** se grafica el comportamiento de los resultados de la simulación.

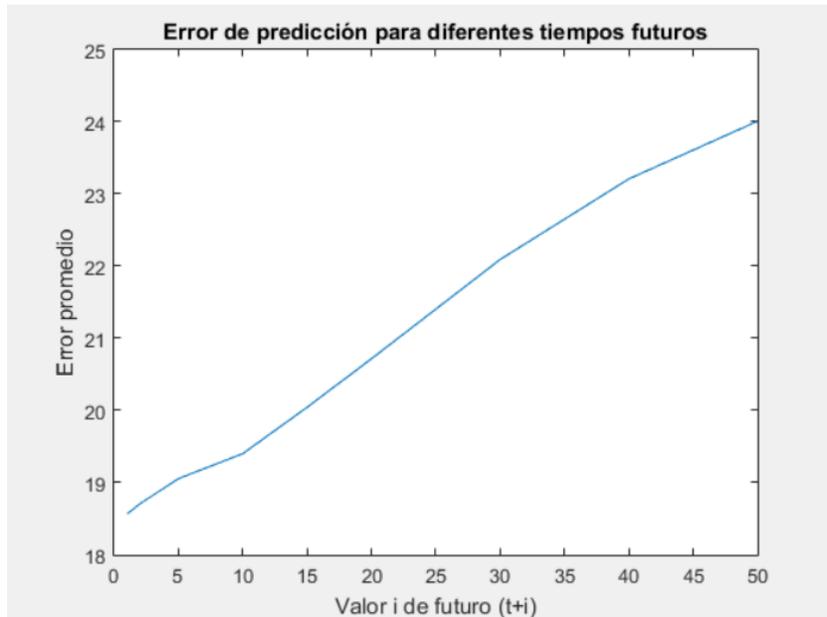


figura 36 Comportamiento del error medio para $(t + i)$ para $i = 1, 2, 5, 10, 15, 30$ y 50 .

A continuación, se observa en la **figura 37** el resultado de la predicción generada por la simulación comparada con la serie original, donde se predice el caudal para un tiempo futuro $(t + 50)$, dado el caudal en $t, t-1, \dots, t-5$.

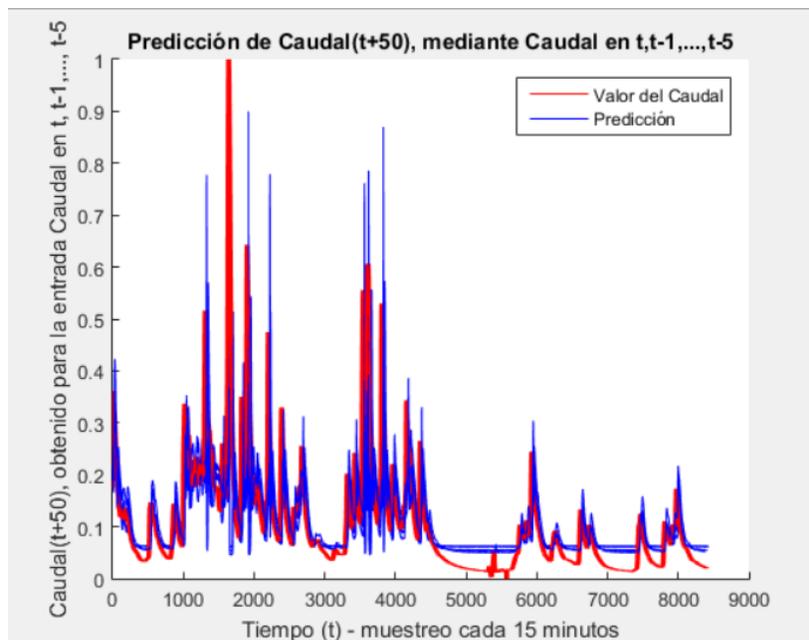


figura 37 Predicción del caudal $(t+50)$ mediante red neuronal RBF

6.2.6 Implementación con redes Feed Forward en Cascada y validación cruzada

6.2.6.1 Experimento n° 5

Se realiza el entrenamiento, y simulación utilizando una red FeedForward en Cascada. Esta es una red con una arquitectura similar a las redes FeedForward (Perceptrón multicapa, pero incluyen una conexión desde la entrada y cada capa anterior, a la siguiente capa. Nuevamente se realiza la predicción futura para los valores $t+i$, con $i = 1, 2, 5, 10, 15, 20, 30, 40$ y 50 , Los mejores resultados se obtienen al utilizar una red de dos capas ocultas con 10 neuronas en cada una de ellas.

La siguiente tabla 3, muestra los resultados obtenidos del error promedio para cada uno de los valores, teniendo en cuenta que se realizaron 20 simulaciones para cada valor de i .

Predicción	t+1	t+2	t+5	t+10	t+15	t+20	t+30	t+40	t+50
Error promedio	2.327%	3.242%	5.815%	9.404%	12.215%	14.480%	17.657%	19.841%	21.364%

Tabla 3 Error medio para predicciones tiempo futuro con una red feedforward en cascada

La **figura 38** muestra el error promedio obtenido a realizar la simulación para diferentes tiempos futuros, donde nuevamente se trata de predecir el valor del Caudal en $t+i$, con $i = 1, 2, 5, 10, 15, 20, 30, 40$ y 50 .

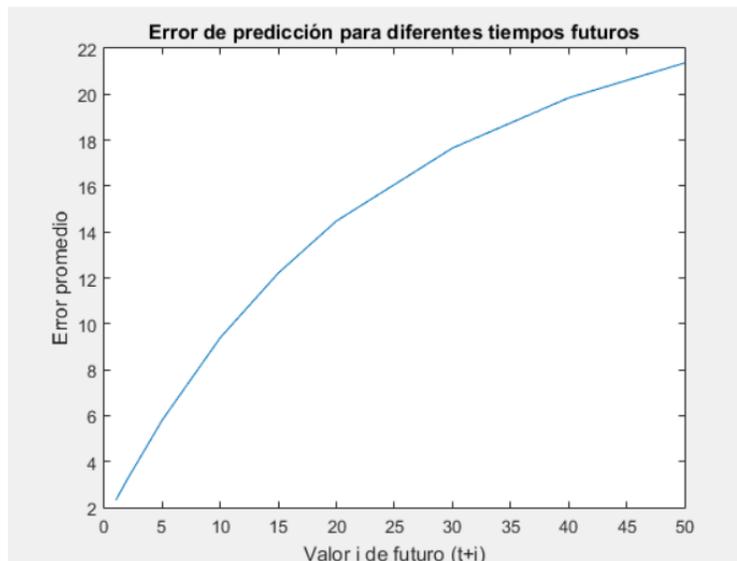


Figura 38 Comportamiento del error promedio en la predicción del Caudal de agua en el tiempo $t+i$, para $i = 1, 2, 5, 10, 15, 20, 30, 40$ y 50 , utilizando una red neuronal FeedForward en cascada

La **figura 39** muestra, en primer lugar, el error obtenido en cada una de las 20 iteraciones de prueba para predecir $t+20$. En segundo lugar, se muestran los resultados obtenidos con la simulación sobre la última red entrenada vs la serie original. Se puede apreciar que la simulación sigue la tendencia y estacionalidad de la red. Sin embargo, al hacer un zoom sobre ella se va a evidencias el incremento del error de predicción que para el caso de predecir $t+50$ es de 21.364%

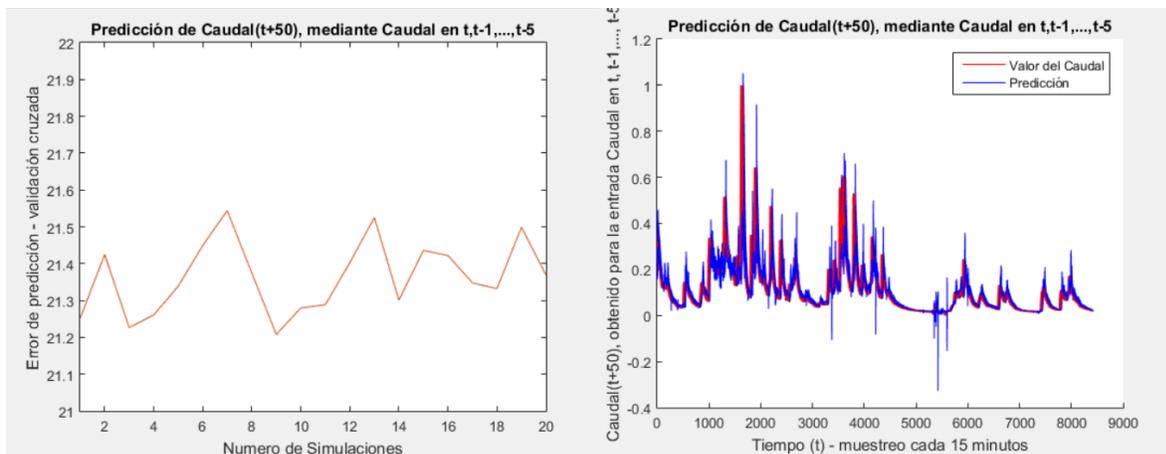


Figura 39. Error de predicción del Caudal($t+50$), para ventana de predicción con el Caudal en $t, t-1, \dots, t-5$ y Predicción del Caudal($t+50$), dado el Caudal en $t, t-1, \dots, t-5$ vs la serie original de Caudal, al utilizar una red FeedForward

6.3 Objetivo 3

Evaluar la capacidad de predicción para cada tipo de red, variando la ventana de predicción y los periodos de tiempo a predecir

- Analizar inicialmente las diferentes ventanas de predicción que se utilizaran como variables de entrada en los tipos de red seleccionados, variando la cantidad de datos de entrada, creando una tabla donde se refleje el error cuadrático medio que se genere al variar la ventana de predicción, esto para cada tipo de red, los resultados de los análisis para las ventanas de predicción aplicadas se ven reflejadas en el desarrollo del segundo objetivo del proyecto.

7. Análisis de Resultados

En el experimento 1 se trata de determinar la capacidad de una red feedforward, es decir un perceptrón multicapa, para predecir el caudal en el tiempo $t+1$, utilizando como entrada el valor del nivel de caudal en el tiempo t . Es decir, el caso más sencillo de predicción. En este caso se esperan los mejores resultados. Se realizaron 20 simulaciones y la figura 19 muestra el error obtenido para cada simulación. El error promedio está en 7%. Se realizaron simulaciones con diferentes arquitecturas de red (aumentando diferentes combinaciones y números de neuronas), y se optó por una arquitectura de 2 capas con 10 neuronas cada 1, con buen rendimiento en términos de coste computacional / Error de predicción.

En el experimento 2 se realizó el entrenamiento y simulación de la red para predecir el caudal en el tiempo $(t+1)$, utilizando como entradas los valores de la serie en varios tiempos anteriores, para lo cual se definieron diferentes ventanas de predicción:

- Ventana 1. Se trata, de predecir el caudal $t+1$, utilizando como entradas el caudal en t , $t-1$, $t-2$
- Ventana 2. Se trata de predecir el caudal $t+1$, utilizando como entradas el caudal en t , $t-1$, $t-2$, $t-3$, $t-4$, $t-5$.
- Ventana 3. Se trata de predecir el caudal $t+1$, utilizando como entradas el caudal en t , $t-1$, $t-2$, ..., $t-10$.
- Ventana 4. Se trata de predecir los caudales $t+1$, utilizando como entradas el caudal en t , $t-1$, $t-2$, ..., $t-20$.

En la tabla 4. se muestra los resultados de error promedio en la predicción del caudal en t , con las diferentes entradas.

Retrasos	Ventana1	Ventana 2	Ventana 3	Ventana 4
Error promedio	6.963%	6.833 %	6.588%	7.1590%

Tabla 4. Error promedio para la predicción del caudal en el tiempo t , con diferentes retrasos en la dirección de entrada

Los resultados de la tabla indican que, si se toma una ventana amplia como información de entrada en la red neuronal, los retrasos más viejos generan cierto ruido en la entrada, que empeora el error de predicción.

Por lo anterior, en el experimento 3 se decide utilizar como entradas para la predicción de datos futuros, solo los valores del caudal en $t, t-1, \dots, t-5$. Luego se desea predecir valores de caudal para tiempos de futuro más distantes, y no solo la $t+1$. En concreto se entrena el modelo para predecir los valores de caudal en los tiempos $t+2, t+5, t+10, t+15, t+20, t+30, t+40$ y $t+50$. Para cada valor futuro a predecir, se realizan 20 entrenamientos y simulaciones, para determinar la estabilidad del error obtenido. En las figuras 27 a 34 se puede apreciar el comportamiento del error en cada caso. Se observa estabilidad en su variación, por lo cual se toma el valor promedio como referencia. La tabla 3 resume los resultados del error para cada tiempo futuro, donde el valor mínimo de error es de 6.2% en $t+5$ y el valor máximo es de 21.3 % en $t+50$. La figura muestra el comportamiento del error, donde se observa que este es creciente a medida que el valor futuro a predecir es mayor. Teniendo en cuenta que el muestreo en la serie es de cada 15 minutos, se puede interpretar que el error de predicción del caudal en 12.5 horas, tiene un error de 21.3%.

En el experimento 4, se repite el modelo de predicción anterior, utilizando otro tipo de red neuronal. Es decir, predecir el valor del caudal en los tiempos futuros $t+2, t+5, t+10, t+15, t+20, t+30, t+40$ y $t+50$, utilizando la información del pasado en $t, t-1, \dots, t-5$. Para ellos se utiliza una red neuronal de función de Base Radial, la cual se caracteriza por ser un aproximador universal de funciones. Los resultados obtenidos se muestran en la tabla 4. Se observa que el error de predicción aumenta, de manera que la predicción empeora con respecto a los resultados obtenidos en el caso

de la red feedforward de tipo perceptrón multicapa. El error más bajo obtenido se encuentra para $t+2$, siendo del 18.56% y el valor más alto se tiene en $t+50$ siendo este del 24%. La figura 36 muestra la gráfica del error, cuyo comportamiento es estrictamente creciente y empeora de forma casi lineal, con lo cual la predicción para valores futuros más altos se incrementa indeseadamente.

En el experimento 5 se prueba el mismo modelo, pero con una red de tipo feedforward escalonada. En este modelo de red neuronal, cada capa de la red se conecta con todas las subsiguientes directamente, generando un modelo altamente conectado. Los resultados obtenidos se muestran en la tabla 5. En este caso el error de predicción mejora con respecto a las redes neuronales previamente utilizadas. En el mejor de los casos el error es de 2.32%, en $t+2$, mejorando notablemente los resultados anteriores. El peor valor del error está en $t+5$, siendo de 21.36%, mejor que en el caso de la red de Función de Base Radial, y casi idéntico a la obtenida por el perceptrón multicapa convencional.

8. Conclusiones

- Se ha desarrollado un modelo para la predicción de caudales medios utilizando redes neuronales y se ha evaluado su desempeño para diferentes ventanas temporales.
- El modelo utiliza la técnica de validación cruzada para mostrar la independencia de los resultados respecto a la selección de los datos de entrenamiento y datos de prueba de la serie de datos disponibles.
- La serie de datos utilizada muestra la captura del caudal de agua realizada en una estación particular, durante tres meses, en intervalos de 15 minutos. El modelo sin embargo se puede aplicar a series con diferentes periodos de captura y frecuencias de muestreo, pero los resultados aquí mostrados son particulares de los datos utilizados
- Los modelos de redes neuronales seleccionados son la red neuronal *feedforward* de tipo perceptrón multicapa, la red neuronal de Función de Base Radial y la red neuronal *feedforward* en cascada.
- El modelo diseñado permite entrenar la red utilizando diferentes retrasos para la predicción, es decir permite utilizar como entrada diferentes intervalos de tiempo pasado, así como predecir el caudal para diferentes ventanas de tiempo futuro.
- Las pruebas preliminares determinan que al utilizar valores de tiempo pasado mayores a $t-5$, no aporta mejores resultados en cuanto al error de predicción, por lo cual se fijó el intervalo $t-1 \dots, t-5$ como dato de entrada para predecir caudales futuros.
- Se realizaron las predicciones en el tiempo futuro $t+2, t+5, t+10, t+15, t+20, t+30, t+40$ y $t+50$. Es decir, las predicciones de caudal obtenidas fueron para los siguientes 30 minutos, 1 hora 15 minutos, 2 horas 30 minutos, 3 horas 45 minutos, 5 horas, 7 horas 30 minutos, 10 horas y 12 horas 30 minutos.

- Los mejores resultados fueron obtenidos por la red neuronal feedforward en cascada, que obtuvo un error de predicción del 2.32% para el caudal en los siguientes 30 minutos y un error del 21.3% para predecir el caudal en las siguientes 12 horas 30 minutos.
- Por el tipo de datos utilizados y con los resultados obtenidos se puede apreciar que un modelo basado en la predicción a través de redes neuronales puede ser implantado dentro de un sistema que permita proveer el comportamiento en periodos cortos de tiempo (12 horas), con un error aceptable. El modelo puede aplicar en casos en los que las estaciones de monitoreo han sido recientemente instaladas y con datos en un intervalo de al menos 3 meses.
- La capacidad de predicción puede mejorar si se cuenta con información adicional, por ejemplo, la obtenida en estaciones meteorológicas, pero en este trabajo no se tuvo acceso a dicha información. También es posible mejorar los resultados incluyendo variables exógenas en los datos de entrada como el nivel de agua, pero en este caso se ha utilizado un modelo exclusivamente AR, autorregresivo.

9. Recomendaciones Y Trabajos Futuros

- Probar el modelo con variables exógenas obtenidas por ejemplo de estaciones meteorológicas, lo cual depende de la disponibilidad de los datos
- Verificar los resultados para valores promedios diarios, en cuyo caso el comportamiento puede ser diferente.
- Replicar el modelo en estaciones diferentes, para el mismo intervalo (3 meses) y utilizando el mismo tipo de muestreo (15 minutos), para evaluar la dependencia que tienen los resultados de las características físicas del lugar de muestreo.

10. BIBLIOGRAFÍA

- Marbello Pérez, R. (s.f.). Manual de practicas de laboratorio de hidráulica. En *HIDROMETRÍA Y AFORO DE CORRIENTES NATURALES* (pág. 242). Medellín.
- Agudelo, A., López Lezama, J., & Velilla, E. (2015). Predicción del precio de la electricidad en la bolsa mediante un modelo neuronal no lineal autorregresivo con entradas exógenas. *Información Tecnológica*, 1.
- Aparicio Mijares, F. J. (1992). *fundamentos de hidrología de superficie*. México D.F: Grupo Noriega Editores.
- Campozano Parra, L. V. (2011). *Análisis de la robustez de un modelo de redes neuronales para la predicción de caudales en la cuenca alta del río Paute*. Cuenca, Ecuador.
- Demuth, H., & Beale, M. (10 de 06 de 1998). *Neural network toolbox for use with matlab* (3 ed.). The mathWorks. Obtenido de producto MATLAB.
- García Martos, D. (2012). Procesos autorregresivos. En *Econometría II, grado en finanzas y contabilidad*. Madrid, España.
- Gázquez Abad, J. C., Jiménez Guerrero, J. F., & Sánchez Fernández, R. (2006). La capacidad predictiva en los métodos Box-Jenkins y Holt-Winters: una aplicación al sector turístico. *Revista europea de dirección y economía de la empresa*, 15(3), 185-198.
- Gómez Gallego, M. (2012). *Redes neuronales para la simulación, estimación y control de caudales impulsados por estaciones de bombeo en zonas regables*. Huelva, España.

- Gomez Vargas, E., Obregón Neira, N., & Socarras Quintero, V. (2010). Aplicación del modelo neurodifuso ANFIS vs redes neuronales, al problema predictivo de caudales medios mensuales del río Bogotá en Villapinzón. *Tecnura*, 14(27), 18 - 29.
- Isasi Viñuela, P., & Galván León, I. (2004). *Redes de Neuronas Artificiales un enfoque práctico*. Madrid, España: Pearson.
- Levin, R., & Rubin, D. (2004). *Estadística para administración y economía* (Vol. 7). Mexico: Pearson Educación.
- Matinez Gomez, E. J., & Pedraza Caballero, L. E. (2012). *Implementacion web de redes neuronales artificiales aplicadas de la predicción de series de tiempo*. Barranquilla, Colombia.
- Mediero Orduña, L. (2007). *Pronostico probabilístico de caudales de avenida mediante redes bayesianas aplicadas sobre un modelo hidrológico distribuido*. Madrid, España.
- Moscoso Barrera, W. D. (2015). *Análisis de la relación entre variables hidrometeorológicas y zonas inundables por medio de sistemas inteligentes artificiales en un tramo del río Bogotá comprendido entre estaciones de puente florencia (Gachancipá) y Tocancipá*. Chia, Colombia.
- N. Koivo, H. (2008). *Neural Networks: Basics usign Matlab Neural Network Toolbox*. Espoo, Finland.
- Neural Network Toolbox*. (s.f.). Recuperado el 01 de 04 de 2019, de <https://edoras.sdsu.edu>

- Ocampo, O. L., & Vélez, J. J. (Enero-Junio de 2014). Análisis comparativo de modelos hidrológicos de simulación continua en cuencas de alta montaña: caso del río Chinchiná. *Ingenierías Universidad de Medellín*, 13(24), 214.
- Pierini, J., Gomez, E., & Telesca, L. (2012). *Predicción de caudales en río Colorado , Argentina*. Bahia Blanca, Argentina.
- Ruiz Suarez, S. (2016). *Métodos de clasificación supervisada aplicados a datos de radares metereologicos*. Buenos Aires, Argentina.
- Spiegel, M., & Stephens, L. (s.f.). *Estadística* (4 ed.). Mexico: McGraw-Hill.
- Te Chow, V., R. Maidment, D., & W. Mays, L. (1994). *Hidrología Aplicada*. Santafé de Bogotá, Colombia: McGraw-Hill.

11. Anexos

A continuación, se anexan los modelos utilizados para la ejecución de las pruebas de predicción, así como tablas de resultados desarrollados en el transcurso de la implementación de los modelos de prueba.

11.1 ANEXO 1:

```
%load DatosCaudalGlenamoyAnualDiario.mat
%E = Caudal;
%LonDatos = length(E);
%E = E/max(E);
%numMaxVentanas = 60;
%for vta = 1:numMaxVentanas
%   nPruebas = 20;
%for j = 1: nPruebas
%   LonES = LonDatos - vta;
%   paso = vta-1;
%   k = 10;
%   for i = 1:LonES
%       x(:,i) = E(i:i+paso)';
%       y(i) = E(i+vta);
%   end
%   c = cvpartition(LonES, 'Kfold', k);
%   for i = 1:k
%       tst = c.test(i);
%       tr = c.training(i);
%       red = newff(x(:,tr), y(:,tr), [10 10]);
%       red.trainparam.epochs = 200;
%       red.trainparam.goal = 0.0001;
%       red = train(red, x(:,tr), y(:,tr));
%       s = sim(red, x(:,tst));
%       err = sum(abs(s-y(:,tst)))/length(s);
%   end
%   disp('Error del ');
%   disp(err*100);
%   disp('%');
%   errV(vta, j) = err*100;
%   clear('x');
%   clear('y');
%end
%plot(mean(errV), 'blue');
%legend('Media');
%hold on
%plot(min(errV), 'green');
%legend('Minimo');
%plot(max(errV), 'red');
%legend('Maximo');
%plot(std(errV), 'black');
```

11.2 ANEXO 2

Tabla de resultados para red feedforward aplicándole número de intentos y pruebas individuales para ajuste de k particiones.

Pruebas del código: PrediccionesCredaICY_R2P1.m
 Última ejecución: 30 marzo 2019

VENTANA = 1		INTENTOS										20																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	PROMEDIO	VARIANZA	DESVIACION				
k = 5	0.0043	0.003	0.013	0.084	0.0247	0.0302	0.0248	0.0334	0.0401	0.0258	0.0371	0.0541	0.0639	0.0745	0.0805	0.0536	0.0533	0.0347	0.0396	0.0344	0.0396	0.0343	0.0383	0.0342	0.0547	0.0005707	0.01144526	0.028433286
k = 6	0.0092	0.0039	0.0142	0.0134	0.0236	0.0288	0.0336	0.0387	0.0434	0.0485	0.0537	0.0642	0.0656	0.0697	0.0735	0.051	0.0864	0.0921	0.0971	0.0829	0.0824	0.0833	0.0831	0.08245	0.0008301	0.00279655	0.002443207	
k = 7	0.0042	0.0052	0.0084	0.0044	0.0043	0.0049	0.0053	0.0062	0.0064	0.0068	0.0073	0.0083	0.0081	0.0084	0.0091	0.0094	0.0101	0.0105	0.0111	0.0105	0.0105	0.0105	0.0105	0.0105	0.000822	0.000822	0.000822	
k = 8	0.005	0.0013	0.0162	0.0213	0.0271	0.031	0.0359	0.042	0.0474	0.053	0.058	0.0636	0.0682	0.0738	0.0788	0.0943	0.0903	0.0951	0.1005	0.1064	0.095435	0.0905314	0.085521584	0.080515284	0.075515284	0.070515284	0.065515284	
k = 9	0.0047	0.01	0.0182	0.021	0.0266	0.031	0.0358	0.041	0.0464	0.0517	0.0574	0.0631	0.0685	0.0742	0.0803	0.0865	0.0931	0.0992	0.1053	0.094345	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	0.05934	
k = 10	0.0041	0.0053	0.017	0.0222	0.0269	0.0313	0.0354	0.0419	0.0473	0.052	0.0572	0.0621	0.0663	0.0707	0.0751	0.0796	0.0844	0.0894	0.0944	0.0994	0.09434	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	
VENTANA = 2		INTENTOS										20																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	PROMEDIO	VARIANZA	DESVIACION				
k = 5	0.0062	0.005	0.0151	0.0254	0.0307	0.0343	0.0381	0.0423	0.0464	0.0504	0.0542	0.0583	0.0624	0.0664	0.0704	0.0744	0.0784	0.0824	0.0864	0.0904	0.0944	0.0984	0.1024	0.1064	0.1104	0.1144	0.1184	0.1224
k = 6	0.0043	0.0059	0.0142	0.0181	0.0216	0.0246	0.0274	0.0301	0.0324	0.0343	0.0359	0.0371	0.0376	0.0376	0.0371	0.0361	0.0341	0.0311	0.0271	0.0221	0.0171	0.0121	0.0071	0.0021	0.0001	0.0001	0.0001	0.0001
k = 7	0.0043	0.0052	0.0084	0.0044	0.0043	0.0049	0.0053	0.0062	0.0064	0.0068	0.0073	0.0083	0.0081	0.0084	0.0091	0.0094	0.0101	0.0105	0.0111	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105
k = 8	0.005	0.0013	0.0162	0.0213	0.0271	0.031	0.0359	0.042	0.0474	0.053	0.058	0.0636	0.0682	0.0738	0.0788	0.0943	0.0903	0.0951	0.1005	0.1064	0.095435	0.0905314	0.085521584	0.080515284	0.075515284	0.070515284	0.065515284	
k = 9	0.0047	0.01	0.0182	0.021	0.0266	0.031	0.0358	0.041	0.0464	0.0517	0.0574	0.0631	0.0685	0.0742	0.0803	0.0865	0.0931	0.0992	0.1053	0.094345	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	0.05934	
k = 10	0.0041	0.0053	0.017	0.0222	0.0269	0.0313	0.0354	0.0419	0.0473	0.052	0.0572	0.0621	0.0663	0.0707	0.0751	0.0796	0.0844	0.0894	0.0944	0.0994	0.09434	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	
VENTANA = 3		INTENTOS										20																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	PROMEDIO	VARIANZA	DESVIACION				
k = 5	0.0055	0.0075	0.015	0.022	0.0284	0.0347	0.0409	0.047	0.0531	0.0592	0.0651	0.0709	0.0766	0.0822	0.0877	0.0931	0.0984	0.1036	0.1088	0.114	0.1192	0.1244	0.1296	0.1348	0.14	0.1452	0.1504	0.1556
k = 6	0.0051	0.0059	0.0142	0.0181	0.0216	0.0246	0.0274	0.0301	0.0324	0.0343	0.0359	0.0371	0.0376	0.0376	0.0371	0.0361	0.0341	0.0311	0.0271	0.0221	0.0171	0.0121	0.0071	0.0021	0.0001	0.0001	0.0001	0.0001
k = 7	0.0043	0.0052	0.0084	0.0044	0.0043	0.0049	0.0053	0.0062	0.0064	0.0068	0.0073	0.0083	0.0081	0.0084	0.0091	0.0094	0.0101	0.0105	0.0111	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105
k = 8	0.005	0.0013	0.0162	0.0213	0.0271	0.031	0.0359	0.042	0.0474	0.053	0.058	0.0636	0.0682	0.0738	0.0788	0.0943	0.0903	0.0951	0.1005	0.1064	0.095435	0.0905314	0.085521584	0.080515284	0.075515284	0.070515284	0.065515284	
k = 9	0.0047	0.01	0.0182	0.021	0.0266	0.031	0.0358	0.041	0.0464	0.0517	0.0574	0.0631	0.0685	0.0742	0.0803	0.0865	0.0931	0.0992	0.1053	0.094345	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	0.05934	
k = 10	0.0041	0.0053	0.017	0.0222	0.0269	0.0313	0.0354	0.0419	0.0473	0.052	0.0572	0.0621	0.0663	0.0707	0.0751	0.0796	0.0844	0.0894	0.0944	0.0994	0.09434	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	
VENTANA = 4		INTENTOS										20																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	PROMEDIO	VARIANZA	DESVIACION				
k = 5	0.0037	0.008	0.015	0.022	0.0284	0.0347	0.0409	0.047	0.0531	0.0592	0.0651	0.0709	0.0766	0.0822	0.0877	0.0931	0.0984	0.1036	0.1088	0.114	0.1192	0.1244	0.1296	0.1348	0.14	0.1452	0.1504	0.1556
k = 6	0.0033	0.0041	0.0084	0.0123	0.0162	0.0191	0.0216	0.0236	0.0251	0.0261	0.0266	0.0266	0.0261	0.0251	0.0236	0.0216	0.0191	0.0162	0.0123	0.0084	0.0041	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
k = 7	0.0043	0.0052	0.0084	0.0044	0.0043	0.0049	0.0053	0.0062	0.0064	0.0068	0.0073	0.0083	0.0081	0.0084	0.0091	0.0094	0.0101	0.0105	0.0111	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105
k = 8	0.005	0.0013	0.0162	0.0213	0.0271	0.031	0.0359	0.042	0.0474	0.053	0.058	0.0636	0.0682	0.0738	0.0788	0.0943	0.0903	0.0951	0.1005	0.1064	0.095435	0.0905314	0.085521584	0.080515284	0.075515284	0.070515284	0.065515284	
k = 9	0.0047	0.01	0.0182	0.021	0.0266	0.031	0.0358	0.041	0.0464	0.0517	0.0574	0.0631	0.0685	0.0742	0.0803	0.0865	0.0931	0.0992	0.1053	0.094345	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	0.05934	
k = 10	0.0041	0.0053	0.017	0.0222	0.0269	0.0313	0.0354	0.0419	0.0473	0.052	0.0572	0.0621	0.0663	0.0707	0.0751	0.0796	0.0844	0.0894	0.0944	0.0994	0.09434	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	
VENTANA = 5		INTENTOS										20																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	PROMEDIO	VARIANZA	DESVIACION				
k = 5	0.0051	0.0059	0.0142	0.0181	0.0216	0.0246	0.0274	0.0301	0.0324	0.0343	0.0359	0.0371	0.0376	0.0376	0.0371	0.0361	0.0341	0.0311	0.0271	0.0221	0.0171	0.0121	0.0071	0.0021	0.0001	0.0001	0.0001	0.0001
k = 6	0.0043	0.0052	0.0084	0.0044	0.0043	0.0049	0.0053	0.0062	0.0064	0.0068	0.0073	0.0083	0.0081	0.0084	0.0091	0.0094	0.0101	0.0105	0.0111	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105	0.0105
k = 7	0.005	0.0013	0.0162	0.0213	0.0271	0.031	0.0359	0.042	0.0474	0.053	0.058	0.0636	0.0682	0.0738	0.0788	0.0943	0.0903	0.0951	0.1005	0.1064	0.095435	0.0905314	0.085521584	0.080515284	0.075515284	0.070515284	0.065515284	
k = 8	0.0047	0.01	0.0182	0.021	0.0266	0.031	0.0358	0.041	0.0464	0.0517	0.0574	0.0631	0.0685	0.0742	0.0803	0.0865	0.0931	0.0992	0.1053	0.094345	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	0.05934	
k = 9	0.0041	0.0053	0.017	0.0222	0.0269	0.0313	0.0354	0.0419	0.0473	0.052	0.0572	0.0621	0.0663	0.0707	0.0751	0.0796	0.0844	0.0894	0.0944	0.0994	0.09434	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	
k = 10	0.0041	0.0053	0.017	0.0222	0.0269	0.0313	0.0354	0.0419	0.0473	0.052	0.0572	0.0621	0.0663	0.0707	0.0751	0.0796	0.0844	0.0894	0.0944	0.0994	0.09434	0.08934	0.08434	0.07934	0.07434	0.06934	0.06434	