



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
INGENIERÍA CIVIL - HIDRÁULICA

APLICACIÓN DE REDES NEURONALES ARTIFICIALES PARA LA PREDICCIÓN
DE ESCURRIMIENTOS EN EL SISTEMA CUTZAMALA

TESIS PARA OPTAR POR EL GRADO DE:
MAESTRO EN INGENIERÍA

PRESENTA:

Ing. José Antonio Romero Gil

TUTOR (ES) PRINCIPAL(ES)

Dr. Fernando Jorge González Villarreal
Instituto de Ingeniería

COMITÉ TUTOR

DR. FELIPE IGNACIO ARREGUÍN CORTÉS	FACULTAD DE INGENIERÍA
DR. RAMÓN DOMÍNGUEZ MORA	INSTITUTO DE INGENIERÍA
DR. RAFAEL BERNARDO CARMONA PAREDES	INSTITUTO DE INGENIERÍA
DR. ERICK MOLINO MINERO RE, IIMAS	IIMAS

CIUDAD DE MÉXICO, OCTUBRE, 2021



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



Dedicatoria:

A la memoria de mi padre, espero desde el cielo puedas ver que sigo creciendo.

Agradecimientos

A mi familia, por todo el apoyo y cariño que me han brindado para alcanzar mis metas.

Al Dr. Fernando J. González Villarreal, por compartir conmigo su amplio conocimiento y experiencia. Le agradezco su guía e instrucción, la confianza, y el apoyo que me ha brindado para crecer tanto académica como profesionalmente.

A mis profesores, tanto del Instituto e Ingeniería como de la Facultad, por todo su esfuerzo y dedicación para compartir su conocimiento.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por la beca brindada que me permitió continuar con mis estudios.

Tabla de contenido

<u>CAPÍTULO I: INTRODUCCIÓN</u>	1
JUSTIFICACIÓN Y ALCANCES	2
ESTRUCTURA DEL TEXTO	3
<u>CAPITULO II ANTECEDENTES</u>	5
MODELOS LLUVIA - ESCURRIMIENTO	5
MODELOS FÍSICOS	8
MODELOS CONCEPTUALES	9
MODELOS EMPÍRICOS	10
MODELOS BASADOS EN REDES NEURONALES ARTIFICIALES	11
MODELOS DE PRONÓSTICO HIDROLÓGICO	19
REDES NEURONALES ARTIFICIALES APLICADAS A MODELOS DE PRONÓSTICO	21
CONTEXTO DEL SISTEMA CUTZAMALA:	23
<u>CAPITULO III MARCO TEÓRICO</u>	27
REDES NEURONALES ARTIFICIALES	27
APRENDIZAJE EN UNA RNA	30
PROBLEMA DE SUB-AJUSTE Y SOBRE-AJUSTE	36
REDES NEURONALES RECURRENTE (RNR):	37
REDES LSTM (LONG SHORT-TERM MEMORY)	40
<u>CAPITULO IV: METODOLOGÍA</u>	43
ANÁLISIS DE LA ZONA DE ESTUDIO	43
INSTRUMENTACIÓN EN LA ZONA DE ESTUDIO	45
PLANTEAMIENTO DEL MODELO:	48
SELECCIÓN DE DATOS Y VARIABLES	49
PREPROCESAMIENTO DE DATOS	51
ANÁLISIS DE DATOS:	51
INTEGRACIÓN DE CONJUNTOS DE DATOS	56
ESCALADO DE ATRIBUTOS	59

DISEÑO DE LA RED	61
ENTRENAMIENTO	63
VALIDACIÓN DEL MODELO	65
<u>CAPÍTULO V: ANÁLISIS DE RESULTADOS</u>	69
PREDICCIÓN HORIZONTE 1 DÍA.	69
PREDICCIÓN HORIZONTE 2 DÍAS.	71
PREDICCIÓN HORIZONTE 3 DÍAS.	72
PREDICCIÓN HORIZONTE 4 DÍAS.	73
PREDICCIÓN HORIZONTE 5 DÍAS.	74
<u>CAPÍTULO VI CONCLUSIONES</u>	75
<u>REFERENCIAS:</u>	79
<u>ANEXO: CÓDIGO DEL MODELO</u>	88

Resumen

El Ciclo Hidrológico es un fenómeno complejo que interconecta diversos fenómenos físicos; no obstante, su entendimiento y modelación es de vital importancia para fines de diseño de infraestructura, planeación de los recursos hídricos, etc. Existen varios enfoques para esto, los cuales van desde modelos basados en ecuaciones que describen los fenómenos físicos (modelos físicos), los cuales tienen la desventaja de requerir una gran cantidad de información que en muchos casos no se encuentra disponible; hasta modelos basados totalmente en la inferencia de datos (empíricos) como los basados en redes neuronales. En este trabajo se propone el uso de este último tipo de modelo para pronosticar los esguimientos en la cuenca “El Bosque” del Sistema Cutzamala, lo anterior con el objetivo de evaluar su desempeño y la factibilidad de aplicarse en un sistema de monitoreo y pronóstico de disponibilidad para el Sistema Cutzamala.

Capítulo I: Introducción

El Ciclo Hidrológico es un fenómeno complejo que interconecta diversos fenómenos físicos; uno de los cuales es el esguerrimiento. Este es de vital importancia para la humanidad ya que además de fungir como fuente de abastecimiento de agua; también provee de otros servicios ecosistémicos y puede afectar gravemente a centros de población a través de inundaciones.

Por ello, este fenómeno se ha tratado de describir a través de modelos. Esto con el objetivo de poder anticipar su comportamiento y gestionar este recurso de manera eficaz. Existen diversos enfoques para esto, los cuales van desde modelos basados en ecuaciones matemáticas que describen a dichos procesos (modelos físicos), que tienen la desventaja de requerir una gran cantidad de información que en muchos sitios, no existe o se encuentra sin actualizar ; hasta modelos basados totalmente en la inferencia de la relación de datos (empíricos), los cuales tienen la ventaja de requerir de un menor número de variables para funcionar, por lo que son más fáciles y económicos de implementar.

Dentro de este trabajo se analiza la implementación un modelo perteneciente a este último grupo que se basa en el uso de redes neuronales artificiales para inferir el volumen de agua que esguerra en una cuenca a partir de la cantidad que se precipita dentro de ella. Para ello, se utiliza un tipo especial de red neuronal artificial denominada “recurrente”; la cual permite procesar series de tiempo y analizar las tendencias y patrones de comportamiento de estas. Esto permite que este tipo de modelos adicionalmente pueda utilizarse para generar pronósticos sobre el comportamiento del esguerrimiento.

De manera que, el objetivo de este trabajo es evaluar el desempeño de este tipo de modelos y su factibilidad para aplicarse en un sistema de monitoreo y pronóstico de disponibilidad para el Sistema Cutzamala. Para ello, se realizó un modelo piloto para pronosticar parte los esguerrimientos que llegan al vaso de la presa “El Bosque” por parte del río Zitácuaro. Se utilizó información de fácil acceso y el menor número de variables posibles, así como software de código abierto; de manera que sea fácilmente escalable en las demás cuencas que conforman al Sistema Cutzamala (S.C). Así mismo, para realizar el cómputo de la red se utilizó una herramienta basada en *cloud computing* gratuita; esto con el objetivo de que pueda ser replicado por cualquier persona interesada sin la necesidad de contar con equipo especializado.

Justificación y Alcances

El sistema Cutzamala es una de las principales fuentes de abastecimiento de la Zona Metropolitana del Valle de México (ZMVM), por lo que es de suma importancia contar con herramientas que permitan realizar el monitoreo de las cuencas que lo conforman. No obstante, los modelos tradicionales requieren ser alimentados con información de la cuenca que muchas veces no está disponible o se encuentra desactualizada; aunado a esto, el cambio global de la cuenca (aquel que se debe a los efectos del cambio climáticos y modificaciones en los tipos y usos del suelo) propicia que estos modelos tengan que ser revaluados y actualizados constantemente, lo que incrementa su costo económico.

Por lo anterior, en este proyecto se propone el desarrollo de un modelo basado en una arquitectura específica de las redes neuronales artificiales denominada recurrente, para realizar el pronóstico de los volúmenes de escurrimiento que llegan a este. Para ello, se propone como cuenca piloto “El Bosque”, y se selecciona un horizonte de pronóstico de hasta 5 días.

Estructura del Texto

En la siguiente sección de este trabajo se presenta una breve descripción de los principales fenómenos físicos relacionados con el proceso de esguimiento en una cuenca; y la manera en la cual estos son abordados por diferentes enfoques de modelación. Posteriormente, se introduce a los modelos basados en redes neuronales artificiales y se presenta una breve síntesis del estado del arte de su aplicación dentro del desarrollo de modelos lluvia-esguimiento. Dado que para poder establecer políticas de operación y planificar la gestión de este recurso no solo es importante conocer la relación que existe entre la lluvia y el esguimiento, sino también lo es poder anticipar su comportamiento, se realiza también una breve introducción a los modelos pronóstico y de la aplicación que tienen las redes neuronales artificiales en este campo. Finalmente, se realiza una breve introducción al contexto y problemáticas actuales referentes a la gestión del agua dentro de las cuencas que comprenden al Sistema Cutzamala (S.C.) . De este último apartado se destaca la necesidad de contar con una herramienta que permita estimar los esguimientos en la cuenca; así como las características particulares del sitio que hacen que este tipo de modelos se adapte fácilmente a sus necesidades.

Posteriormente, dentro del marco teórico se realiza una descripción detallada de qué son y cómo funcionan las redes neuronales artificiales; para lo cual primero se presentan los paralelismos que existen entre las redes neuronales artificiales y las biológicas con el objetivo de brindar una intuición de su funcionamiento. Posteriormente, se presenta formalmente el funcionamiento matemático de estos modelos y se introduce a los diferentes tipos de aprendizajes existentes dentro del campo de la inteligencia artificial, haciendo énfasis en el aprendizaje supervisado para problemas de regresión, el cual será el eje central de modelo desarrollado. Al final de este capítulo, se presenta la arquitectura *Long-Short Term Memory* (LSTM), la cual es un tipo especial de arquitectura recurrente que permite trabajar con series de tiempo; y será la utilizada para desarrollar el modelo.

En el siguiente capítulo se muestra una descripción detallada de la metodología y consideraciones tomadas para el desarrollo del modelo. Se presenta una descripción del sitio de estudio así como la instrumentación y tipo de información disponible en este. A continuación, se realiza el planteamiento del modelo acorde a las necesidades específicas del sitio y de los objetivos anteriormente establecidos. Seguido de ello, se muestra el proceso de selección y preprocesamiento de las variables utilizadas, así como de la programación de la red. Finalmente, se describe el proceso de entrenamiento, prueba y validación del modelo.

Una vez que el modelo ha sido probado y validado, dentro de la siguiente sección se analizan los resultados obtenidos para los diferentes horizontes de pronósticos. Dentro de esta sección se discuten las bondades y debilidades del modelo; así como las aplicaciones que podría tener la información provista por este para la planeación de este recurso. Finalmente, dentro la siguiente sección se establecen las conclusiones, recomendaciones y líneas de investigación propuestas para continuar con este estudio.

Capitulo II Antecedentes

Modelos Lluvia - Esgurrimiento



Figura 2.1: Esquema del Ciclo Hidrológico (USGS, 2006)

El Ciclo Hidrológico, el cual describe los medios y dinámicas a través de las cuales el agua fluye y se distribuye en nuestro planeta, está compuesto por diversos fenómenos que interactúan entre sí (ver Fig.2.1); uno de los cuales es el esgurrimiento. Este comprende al proceso mediante el cual, el agua proveniente de precipitación y deshielo fluye de manera superficial y subterránea hasta su llegada hacia una corriente, para finalmente ser drenada hasta la salida de la cuenca (Aparicio, 1992).

El esgurrimiento se conforma por dos tipos de flujo, siendo el principal el denominado "esgurrimiento directo"; el cual es generado a partir de la precipitación mediante la combinación de dos fenómenos (Yang et al., 2015): El primero es la saturación del suelo, es decir al momento en el cual se excede la capacidad de retención del este y el agua que se precipita comienza a encharcarse y fluir de manera superficial (Johnson et al., 2003); el segundo es la capacidad de infiltración del suelo, ya que si la intensidad de la tormenta supera su capacidad de absorción, el excedente fluirá de manera superficial (Horton, R. E., 1933).

Así, este depende de variables fisiográficas de la cuenca tales como: el tipo y usos del suelo, la configuración del terreno, el tipo y cobertura de vegetación; así como de características de la tormenta tales como su duración, intensidad y

volumen; por lo que la relación que guarda el volumen de agua que se precipita y aquel que escurre a la salida de una cuenca es variable, tanto espacial como temporalmente en un mismo sitio.

Posteriormente, un porcentaje del agua que se infiltra en el suelo fluye de manera subsuperficial y subterránea hasta desembocar en un cuerpo de agua; a este flujo se le conoce cómo “escurrimiento base” y es el que predomina durante los periodos de estiaje (Aparicio, 1992). A diferencia del escurrimiento directo, este no está relacionado con una tormenta específica; sino que es originado por el intercambio de agua del subsuelo hacia la superficie; por lo que su magnitud depende del volumen de agua almacenada en este; así como de características físicas tales como su estratigrafía y conductividad hidráulica (Singh, K. P. ,1968).

Finalmente, durante el proceso de transporte hacia la salida de la cuenca, aproximadamente dos terceras partes del agua que escurre regresa a la atmosfera en forma de vapor (Perlman, 2016). Esto, debido a la combinación de dos fenómenos: el primero, la evaporación del agua superficial, el cual depende de variables meteorológicas tales como la radiación solar, temperatura, humedad relativa del aire y la velocidad del viento (Rijtema, P. E.;1965); el segundo, la transpiración de la vegetación, el cual es un fenómeno más complejo ligado al proceso de producción de biomasa a partir de CO₂ y luz solar (fotosíntesis), y es por tanto dependiente del tipo de vegetación (particularmente de características morfológicas tales como el número y dimensiones de sus estomas), su cobertura, el número de horas con luz solar disponible y los niveles de CO₂ en el aire (de Wit C.T, 1958).

Al efecto combinado de estos dos fenómenos se le denomina evapotranspiración (Chow et al., 1988); y ha sido objeto de estudio de diversos autores (Allen R.G et al, 1998; Hargreaves, 2003; Santiago-Rodríguez et al, 2012), para su estimación y comprensión. Cabe mencionar que la dinámica de este fenómeno es sumamente compleja; ya que, por ejemplo, en zonas tropicales este fenómeno genera “mini ciclos del agua” donde el agua evapotranspirada vuelve a condensarse en forma de precipitación en un lapso de tan solo 1-2 días, dentro de la misma cuenca o en las cuencas contiguas (Peters T, 2016).

De este modo, es posible describir el escurrimiento en una cuenca para un periodo de tiempo determinado mediante la *ecuación 2.1* (Sitterson Jan, et al,2017); al realizar el balance de materia de la iteración de los fenómenos antes descritos. Sin embargo, tal como se ha presentado, cuantificar a cada uno de los elementos que conforman a esta ecuación puede resultar en una labor sumamente compleja debido a la cantidad de procesos involucrados y a la variación espacial y temporal.

$$Q_s = P - ET - \Delta SM - \Delta GW \quad Ec 2.1$$

Ecuación 2.1. Donde Q_s es el escurrimiento superficial, P es la precipitación, ET es la evapotranspiración, ΔSM es el cambio en la humedad del suelo, y ΔGW es el cambio en el almacenamiento del subsuelo.

No obstante, el estudio y análisis del escurrimiento es de vital importancia para la sociedad; ya que este, además de proveer de una serie de servicios ecosistémicos tales como el transporte de nutrientes, sedimentos y fungir como fuente de abastecimiento de agua para el desarrollo de actividades humanas (Alan Yeakley, J et al, 2016); también puede causar afectaciones a través de inundaciones o la degradación de suelos (Sitterson Jan, et al, 2017). Por lo que es una necesidad contar con herramientas para describir a este flujo y poder anticipar su comportamiento para el diseño de infraestructura y la elaboración de planes de gestión sobre este.

Al respecto, diversos autores han intentado representar a este fenómeno a través de modelos, los cuales son un conjunto de ecuaciones que ayudan en la estimación de la cantidad de lluvia que se convierte en escurrimiento en función de varios parámetros utilizados para describir el comportamiento de una cuenca (Devi et al 2015). Esto tiene dos retos fundamentales: el primero es establecer las relaciones que permitan cuantificar a cada uno de los elementos antes mencionados; el segundo es el de contar con la información necesaria (en cantidad y calidad) para poder alimentar a estos modelos.

En la actualidad, gracias al desarrollo tecnológico y al mejor entendimiento de los procesos hidrológicos, es posible desarrollar modelos complejos que emulen casi a la perfección las dinámicas intrínsecas dentro del ciclo hidrológico. Sin embargo, estos requieren de una gran cantidad de datos sobre la cuenca; los cuales no siempre están disponibles o se encuentran actualizados. De manera que, algunos autores han optado por idealizar algunos de los procesos con el objetivo de simplificar los cálculos y reducir el número de variables e información necesaria para alimentar al modelo.

Esto lleva a un amplio abanico de posibilidades para elegir a la hora de buscar una herramienta para modelar el escurrimiento dentro de una cuenca; cada una con sus respectivos beneficios y limitaciones. La elección de uno u otro depende esencialmente del propósito del modelo, la cantidad y calidad de datos, tiempo y recursos disponibles (Devi et al 2015); de manera que ninguno es superior a otro y cada uno tiene su contexto específico de aplicación. Todos ellos comparten elementos tales como entradas, ecuaciones de gobierno, condiciones o parámetros de contorno, procesos del modelo y resultados (Singh, 1995).

De este modo, los modelos hidrológicos se pueden clasificar de acuerdo con la manera en la que describen el ciclo hidrológico en tres grupos: físicos, conceptuales y empíricos (EPA 2017); a continuación, se presentan una breve descripción de cada uno de estos grupos (ver *tabla 2.1*) así como algunos ejemplos de estos encontrados en la bibliografía consultada:

Tabla 2.1. Comparación de la estructura básica para modelos Lluvia Esguimiento

	Físico	Conceptual	Empírico
Método	Leyes físicas y ecuaciones basadas en respuestas hidrológicas reales.	Ecuaciones simplificadas que representan procesos físicos	Relación no lineal entre entradas y salidas, concepto de caja negra
Fortaleza	Es versátil y permite la simulación de diferentes escenarios con modificaciones en los datos iniciales.	Fácil de calibrar, requiere de poca información.	Requiere de pocos datos de entrada. Preciso.
Debilidad	Requiere de una gran cantidad de información y su certeza depende de la calidad de esta.	No se considera variabilidad espacial	No tienen conexión física.
Uso Recomendado	Cuencas estudiadas con gran cantidad de información disponible.	Se cuentan con pocos recursos computacionales o los datos son limitados.	Cuencas con poca información.

Modelos Físicos

Estos son una representación matemática idealizada del fenómeno real; también se les denominan modelos “mecanicistas” o “basados en procesos” ya que se basan en los principios de los procesos físicos. Utilizan las ecuaciones físicas que rigen los procesos hidrológicos para describir el flujo del agua en una cuenca (Devi et al., 2015). Para lo cual, requieren contar con datos físicos de esta tales como: el contenido de humedad del suelo, la profundidad inicial del agua, la topografía, las dimensiones de la red fluvial, etc.

Tienen como beneficio el que no es necesario contar con extensos registros de datos hidrológicos y meteorológicos para su calibración. No obstante, requieren del estudio minucioso de las características de la cuenca a pequeña escala. A continuación, se presentan algunos ejemplos de este tipo de modelos encontrados en la bibliografía consultada:

Modelo SWAT (Soil and Water Assessment Tool)

Este es un modelo físico desarrollado por el departamento de agricultura de Estados Unidos. Requiere para su funcionamiento datos de precipitación diaria, la temperatura máxima y mínima del aire, radiación solar, humedad relativa del aire, velocidad del viento, topografía, uso de la tierra, vegetación y características del suelo. Permite describir eficientemente la circulación de agua y sedimentos, el crecimiento de la vegetación, la circulación de nutrientes, así como la tasa de nevadas. (White et al., 2011). Este modelo ha sido ampliamente utilizado en diversos campos de aplicación que van desde el pronóstico de inundaciones para las ciudades de Etiopía y Nueva York (White et al 2011), análisis de escenarios de cambio climático (Aana E.K. et al 2019) y de cambios en el uso de suelo (P.P. Saha et al 2019).

MIKE SHE model (Systeme Hydrologique European)

Este modelo fue desarrollado en colaboración entre el Instituto Británico de Hidrología, Instituto Danés de Hidráulica (DHI) y la empresa francesa de consultoría SOGREAH, formando la “*Association pour le SHE*” (ASHE). Este modelo considera varios procesos del ciclo hidrológico como la precipitación, evapotranspiración, interceptación, escurrimientos, flujo del agua subterránea saturada, flujo de agua subterránea no saturada, etc. Puede simular el movimiento de las aguas superficiales y subterráneas, sus interacciones, el transporte de sedimentos, nutrientes y plaguicidas en el área del modelo y varios problemas de calidad del agua y se puede aplicar para grandes cuencas hidrográficas (Refsgaard et al 2010); de manera similar al modelo anterior, este requiere de datos físicos de la cuenca para establecer las condiciones iniciales. Este ha sido ampliamente utilizado para el monitoreo de los cuerpos de agua en Dinamarca (Henriksen et al. 2008), establecer políticas de operación de presas (Refsgaard et al. 1998), entre otras aplicaciones.

Modelos Conceptuales

Los modelos conceptuales interpretan los procesos de escurrimiento conectando componentes simplificados del proceso hidrológico general. Se basan en ecuaciones que idealizan los procesos físicos, brindando una idea conceptual de los comportamientos de la cuenca (Devi et al., 2015). Requieren como información registros hidrometeorológicos, así como datos generales de la cuenca, y utilizan una serie de parámetros para ajustarse a la realidad.

Tienen como ventaja su simplicidad y el que, en comparación con los modelos físicos, requieren de información más simple y menos precisa. Por otro lado, tienen como inconveniente el que no consideran la variabilidad espacial y que, debido a la falta de significado físico de los parámetros, presentan limitaciones para evaluar escenarios donde se modifican las condiciones de la cuenca (EPA,2017). Algunos ejemplos de este tipo de modelo encontrados en la bibliografía se muestran a continuación:

Modelo para Pronóstico de Esguimiento (MPE)

Este modelo fue desarrollado por el Instituto de Ingeniería de la UNAM, el cual permite obtener hidrogramas de esguimiento de una cuenca; este está basado en el método de número de curva o SCS (*Soil Conservation Service*), con adaptaciones que permite considerar aspectos tales como el secado del suelo después de la lluvia (Ramón et al, 2008). Cabe mencionar que este modelo ha sido ampliamente utilizado en diversas cuencas de México.

TOPMODEL (Topographical Model)

Este es un modelo conceptual que se basa en la topografía de la cuenca y las características del suelo para describir las dinámicas dentro del proceso hidrológico de esguimiento a partir del análisis de áreas de contribución variables (Beven et al 2020). Debido a su simplicidad y fácil aplicación, este modelo ha sido aplicado ampliamente para diversos estudios en todo el mundo.

Modelos Empíricos

Los modelos empíricos, a veces llamados modelos “basados en datos” , utilizan relaciones estadísticas no lineales entre entradas y salidas. Están orientados a la observación y dependen en gran medida de la precisión de los datos de entrada (Kokkonen et al., 2001). En general, estos modelos se basan en la asociación de la lluvia con el esguimiento, por lo que estas suelen ser sus únicas variables de entrada. De manera que, este tipo de modelos puede ser descrito de la siguiente forma:

$$Q = f(X, Y)_{Eq 2.2}$$

Ecuación 2.2. Donde Q es el esguimiento; X y Y son las bases de datos de precipitación y registro histórico de esguimiento; y f la función del modelo que relaciona ambas variables.

La mayoría de los modelos empíricos son modelos de “caja negra”, lo que significa que se sabe muy poco sobre los procesos internos que controlan cómo se determinan los resultados (Beven, 2012). De manera que la función (f) utilizada para transformar la lluvia en esguimiento es un procedimiento desconocido (como en el aprendizaje automático) o sin ninguna referencia a los procesos físicos (como en el método del número de curva).

Este tipo de modelos ofrecen la ventaja de requerir de una menor cantidad de información y de estudios para su funcionamiento, brindando resultados similares a los anteriores enfoques. Lo que hace que los modelos basados en datos sean fáciles y económicos de aplicar; además de que disminuyen la incertidumbre asociada a errores de distorsión producidos por los datos de entrada (Dawson y Wilby, 2001).

Algunos ejemplos de modelos empíricos son el método del *Soil Conservation Service* (SCS) basado en el número de curva o las ecuaciones de regresión, las cuales encuentran la relación funcional entre entradas y salidas (Devi et al., 2015); asimismo en los últimos años una técnica que se ha desarrollado considerablemente gracias al desarrollo tecnológico es la implementación de redes neuronales artificiales (RNA), las cuales tienen la habilidad de reconocer patrones y de las que se hablará más a detalle a continuación.

Modelos basados en Redes Neuronales Artificiales

Las redes neuronales artificiales son una técnica, dentro del campo de la inteligencia artificial y el aprendizaje automatizado, que emula la manera en la cual el cerebro aprende. Tienen la capacidad de identificar relaciones y patrones, por lo que son eficientes resolviendo problemas complejos y a gran escala relacionados con: modelado no lineal, clasificación, asociación y control (ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000).

Debido a la habilidad que tienen las redes neuronales para mapear relaciones multidimensionales entre distintas variables, estas han sido aplicadas ampliamente dentro del campo de la hidrología en diferentes ámbitos (ASCE ²,2000). Dentro de este trabajo únicamente se analizará su aplicación para el desarrollo de modelos lluvia-esguimiento y de pronóstico. Sin embargo, cabe mencionar que esta técnica también tiene aplicaciones en otros ámbitos tales como para el análisis y pronóstico del nivel en un acuífero (Roshni. T et al, 2020); predicción del flujo base (Gerald Corzo et al, 2007); y de la calidad de agua en cuerpos superficiales (Yirgalem Assegid et al, 2012).

Particularmente, en su aplicación para el desarrollo de modelos lluvia-escurrecimiento, se destaca que estos no requieren de información fisiográfica de la cuenca de estudio; lo que les hace especialmente útiles dentro del contexto de sitios aforados, pero con estudios limitados o sin actualizar. Adicionalmente, estos modelos tienen la habilidad de continuar aprendiendo (durante su operación estos se retroalimentan para ajustar sus parámetros) lo que les permite adaptarse ante el cambio global en una cuenca. Finalmente, este tipo de modelos puede vincularse con otro tipo de herramientas tales como sistemas de medición en tiempo real, sistemas de alerta temprana, etc ; así como otros enfoques de gestión relacionados con el concepto de “ciudad inteligente”.

Es necesario mencionar también que este tipo de enfoque tiene limitaciones; ya que su precisión depende de la cantidad y calidad de los datos con los cuales se alimenta a la red, por lo que requiere de base de datos confiables de mediciones (Banko & Brill, 2001). Adicionalmente, demandan de una alta capacidad computacional para su entrenamiento; y requieren de conocimientos avanzados de programación. Motivos por los cuales su desarrollo hasta ahora ha sido limitado.

La manera en la cual funcionan y son entrenados estos modelos basados en el uso de redes neuronales artificiales se discutirá a detalle dentro del capítulo III. Por lo que a continuación, únicamente se presenta una síntesis de la bibliografía consultada relacionada con la aplicación de este tipo de modelos con el objetivo de mostrar el estado del arte de su desarrollo, evolución y alcances:

Half et al. (1993) desarrolló uno de los primeros modelos lluvia-escurrecimiento basado en redes neuronales; utilizando una arquitectura feedforward con una red de tres capas y 5 nodos para representar los escurrecimientos medidos en una cuenca ubicada en Bellevue, Washington. El modelo fue alimentado con hietogramas de 4 tormentas y sus respectivos hidrogramas, obteniendo resultados aceptables y similares a los de un modelo lineal de un hidrograma unitario.

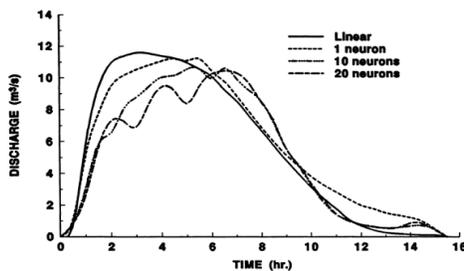


Fig. 5: Response of two-layer linear and three-layer nonlinear neural networks to a unit input

Table 3. Percent error in peak discharge prediction using hidden layers of various sizes.

Flood Date	Linear	Error in Peak Discharge (%)		
		1 Neurons in Hidden Layer	10	20
May 30, 1974	25.47	21.22	33.42	37.45
June 08, 1974	21.12	28.70	24.17	27.98
May 06, 1977	34.91	42.46	39.91	39.43
Aug. 01, 1978	5.29	13.75	0.58	10.27
Apr. 30, 1979	8.98	4.51	12.89	9.40
July 03, 1980	26.58	29.65	39.18	30.03
June 19, 1981	10.98	50.14	24.79	40.58
July 23, 1981	13.67	31.57	7.76	25.96
June 08, 1982	10.33	19.47	4.35	26.94
Sept. 01, 1982	6.86	42.06	5.30	7.81
Dec. 02, 1982	16.00	22.27	26.91	14.57
Apr. 30, 1983	20.17	14.78	25.75	16.25
June 08, 1983	6.59	6.09	3.80	1.15
Oct. 31, 1984	0.79	4.87	0.78	5.12
May 16, 1986	14.41	17.79	7.43	2.79
Sept. 23, 1986	42.24	49.77	50.31	53.53
Average values	16.52	21.82	16.13	17.28

Fig 2.2. Hidrograma Unitario RN (Half et al. 1993)

H.Y. Gupta, K. HSU, & S. Sorooshian (2000), desarrollaron un modelo similar utilizando una arquitectura *feedforward* con una capa intermedia de 3 nodos para la predicción de caudales en la cuenca del Río Leaf en Collins, Mississippi. Alimentaron al modelo con los registros de 7 años de precipitación y los caudales registrados 3 días antes para predecir el siguiente día; de manera que el modelo podía identificar el caudal base y ajustarse de manera satisfactoria a los periodos húmedos y secos. El modelo fue comparado con el utilizado por el “*National Weather Service River*” para la predicción de inundaciones, demostrando un desempeño similar, e incluso superior durante periodos de secos.

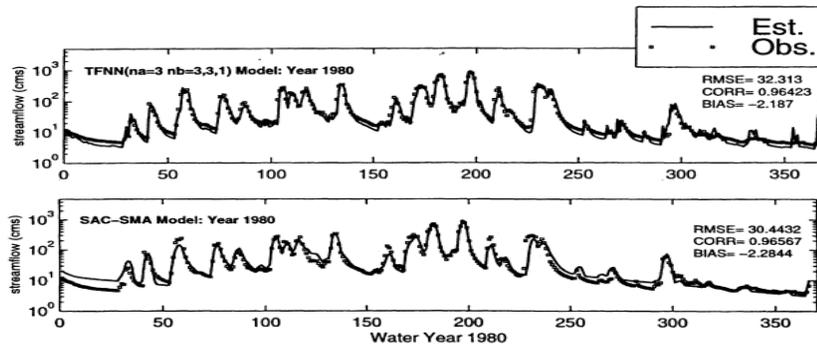


Fig 2.3. Resultados Modelo RN (H.Y. Gupta, K. HSU, & S. Sorooshian .2000)

J.D. Salas Et Al. (2000) realizaron pruebas con diferentes modelos para la predicción de caudales; uno de los modelos fue entrenado con datos de escurrimiento mensuales medidos en la cuenca del río Frazer en Colorado, de manera que a partir de los valores de junio se pudiese predecir los escurrimientos en julio; fue alimentado con información de 9 años de registro, obteniendo resultados favorables con un coeficiente de correlación de 0.928 entre los datos estimados y los reales. Posteriormente, en la cuenca del río Little Patuxent en Maryland; desarrollaron un modelo de predicción con datos diarios.

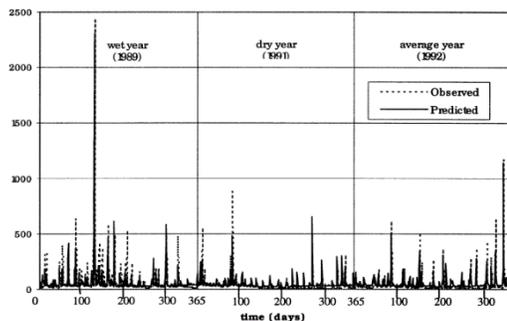


Table 2.4. Validation results of the ANN model (2.34) for forecasting July's flows in the period 1961-1970. Results are given for both the standardized and original flows.

Year	Results for the Standardized Flows				Results for the Original Flows		
	$q_{v,6}$	$q_{v,7}$	$\bar{q}_{v,7}$	e_t	$Q_{v,7}$	$\bar{Q}_{v,7}$	e_t
1961	-0.644	-0.775	-0.529	0.246	12,689.5	15,338.5	2,649.0
1962	0.789	0.865	0.477	-0.388	30,335.7	26,160.9	-4,174.8
1963	-1.451	-1.410	-0.713	0.696	5,858.7	13,353.6	7,494.8
1964	-0.712	-0.696	-0.551	0.144	13,543.4	15,094.3	1,550.9
1965	1.449	1.804	1.282	-0.521	40,439.6	34,828.8	-5,610.8
1966	-1.338	-1.053	-0.696	0.356	9,701.1	13,535.9	3,834.8
1967	0.039	0.323	-0.192	-0.514	24,501.1	18,968.1	-5,533.0
1968	0.357	-0.206	0.051	0.257	18,808.8	21,578.1	2,769.4
1969	0.448	0.653	0.132	-0.522	28,058.8	22,444.9	-5,613.9
1970	1.062	0.495	0.795	0.301	26,351.1	29,587.8	3,236.7
Sum	-0.001	0.000	0.056	0.055	210,287.8	210,890.9	603.0
Mean	0.000	0.000	0.006	0.005	21,028.8	21,089.1	60.3
Std.dev.	1.000	1.000	0.678	0.448	10,760.8	7,296.9	4,824.1
SkewCo	-0.168	0.301	0.726	-0.098	0.301	0.726	-0.097

Fig 2.4. Resultados Modelo (J.D. Salas Et Al. ,2000)

El entrenamiento se realizó con datos de un año considerado “húmedo” uno “seco” y uno “promedio”. Se realizaron pruebas con 9 variaciones distintas del modelo, alternando los siguientes datos de entrada: precipitación, caudales medidos en los días previos, deshielo y temperatura promedio. Se encontró que los mejores resultados para esta cuenca se obtenían al considerar únicamente la precipitación de los dos días anteriores y la temperatura promedio. Los resultados obtenidos para ambos modelos fueron favorables; en el modelo de predicción mensual, la predicción era más acertada que la brindada por modelos estocásticos ARMA(p); por otra parte, el modelo diario tenía un desempeño equiparable al de un modelo físico, el cual debía considerar, además, parámetros tales como la saturación del suelo, rugosidades, entre otras variables difíciles de encontrar en un gran número de cuencas. Cabe mencionar que el número de datos para alimentar el modelo fue considerablemente inferior a lo que hoy en día es recomendable, probablemente debido a la capacidad computacional de la época. Sin embargo, se destaca que los resultados fueron favorables.

Jy S. Wu Et. Al. (2005) entrenaron dos modelos basados en redes neuronales en la cuenca del río Cape Fear, en Carolina del Norte Estados Unidos, para la predicción de caudales con el objetivo de desarrollar una herramienta de alerta temprana para inundaciones. El primer modelo fue entrenado para representar la relación lluvia escurrimiento; el cual fue alimentado con la precipitación registrada dos horas antes, en intervalos de 15 minutos, y el caudal registrado una hora antes, con el objetivo de realizar predicción del escurrimiento a 15, 30, 45 minutos y una hora en el futuro.

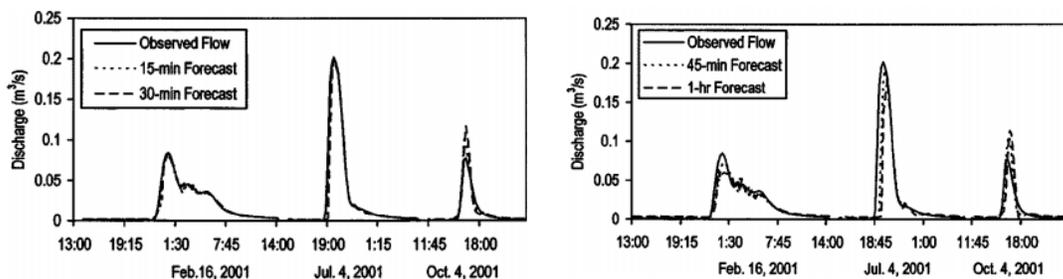


Fig 2.5. Hidrograma Estimado del Primer Modelo (J.D. Salas Et Al. ,2000)

El segundo modelo fue entrenado para predecir los caudales de la estación hidrométrica siguiente en la cuenca del río, utilizando como datos de entrada los valores medidos en la primera estación; con el objetivo de realizar el tránsito de avenidas en el río. En ambos modelos, únicamente se disponía con información registrada durante 2001; por lo cual, de las 12 tormentas registradas, 9 se utilizaron para realizar el entrenamiento y 3 para realizar la validación de los modelos. Pese

a la corta longitud de los registros, ambos modelos presentaron un buen desempeño. En el primer modelo, los resultados de la predicción fueron aceptables hasta un horizonte de 45 min en el futuro y se concluyó que utilizar registros de lluvia que excedieran el tiempo de concentración únicamente entorpecía el proceso de entrenamiento sin presentar una mejora significativa (J.D. Salas Et Al. ,2000).

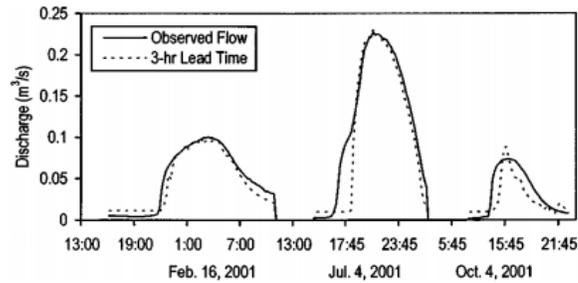


Fig 2.6. Hidrograma Tránsito de Avenida Estimado vs Real (J.D. Salas Et Al. ,2000)

Lucas Pujol Reig (2009) desarrolló una metodología para la aplicación de redes neuronales en la predicción de caudales en tiempo real; para ello utilizó dos tipos diferentes de modelos. El primero fue un modelo convencional con arquitectura feedforward; el segundo modelo consistía en una arquitectura modular, el cual primero clasificaba los datos de entrada automáticamente en grupos (dependiendo su comportamiento), y posteriormente utilizaba un modelo convencional para cada grupo. Posteriormente se comparó los resultados obtenidos contra los de un modelo estocástico autorregresivo de media móvil con variable exógena (ARMAX); siendo los modelos basados en redes neuronales los que mejor desempeño mostraron en una prueba piloto en la cuenca del río Neuquén, ubicado al sur de argentina. De la metodología propuesta destaca el desarrollo de un método para evaluar la incertidumbre en este tipo de modelos.

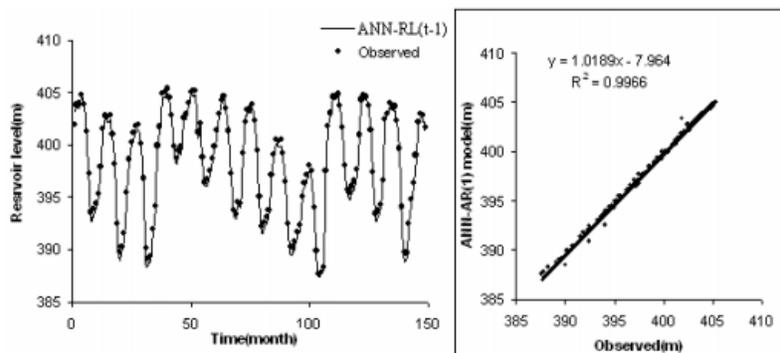


Fig 2.7. Comparación entre niveles estimados y reales (Fatih Üneş, 2010)

Fatih Üneş (2010) desarrolló un modelo basado en redes neuronales para la predicción de volúmenes de almacenamiento en la presa Tahtakopru, al sureste de Turquía. La red utilizaba una arquitectura feedforward, con una capa intermedia de 4 nodos. Como datos de entrada la red utilizó el nivel del reservorio para el tiempo t y $t-1$, así como la precipitación mensual en el tiempo t y la demanda de agua estimada, para predecir el nivel del reservorio en $t+1$; utilizando el registro de 1976 a 2004 para todas las variables de entrada, se generaron 200 pares de datos de entrenamiento y 176 pares para validación. Los resultados de este modelo fueron favorables, con un coeficiente de determinación de 0.97, presentando un mejor desempeño que los obtenidos por un modelo estocástico y de regresión lineal múltiple.

Yu. J . Et Al. (2014) desarrollaron dos modelos basados en una arquitectura feedforward para predecir escurrimientos en dos cuencas diferentes, la cuenca del río Heshui, en China; y la cuenca del río Coquitlam, en Canada. El primer modelo fue alimentado con los niveles de los tres días previos del río y la precipitación de los tres días anteriores; tenía como objetivo predecir el nivel del río para alertar a la población de posibles inundaciones. El segundo modelo fue entrenado con información de la temperatura media, el caudal registrado y la precipitación medida en la cuenca alta de los tres días anteriores; tenía como objetivo predecir el caudal para el siguiente día. De manera simultánea, se desarrollaron otros modelos matemáticos basados en polinomios para aproximar la relación entre las variables de entrada y salida. Los resultados para ambos tipos modelos fueron favorables, siendo los modelos basados en redes neuronales los que mejor resultados obtuvieron; no obstante, los autores destacan dentro de sus conclusiones el elevado tiempo de cómputo para el proceso de entrenamiento en comparación con los modelos matemáticos.

Table 2. Performance in Predicting Water Level for Heshui Catchment

Models	Calibration				Verification				Computation time (s)
	R^2	E	RMSE	MAPE	R^2	E	RMSE	MAPE	
MLR	0.898	0.898	0.1178	6.7185	0.900	0.890	0.1148	7.3727	5
QRS	0.923	0.923	0.1023	5.8478	0.926	0.926	0.0988	6.4761	5
NRS	0.925	0.925	0.1012	5.7429	0.932	0.932	0.0945	6.2977	20
ANN-GDA	0.914	0.914	0.1082	6.0689	0.920	0.920	0.1028	6.5330	1,110
ANN-LM	0.920	0.920	0.1037	5.8393	0.922	0.921	0.1016	6.4467	450

Table 4. Performance in Predicting Daily Discharge for Coquitlam Catchment

Models	Calibration				Verification				Computation time (s)
	R^2	E	RMSE	MAPE	R^2	E	RMSE	MAPE	
MLR	0.594	0.594	6.1419	298.231	0.650	0.650	4.0368	238.458	5
QRS	0.691	0.691	5.5382	236.668	0.722	0.720	3.6089	190.278	5
NRS	0.704	0.704	5.2450	233.841	0.721	0.719	3.6189	186.444	240
ANN-GDA	0.654	0.654	5.6795	267.500	0.721	0.720	3.6120	203.748	720
ANN-LM	0.642	0.640	5.7875	268.129	0.725	0.711	3.6685	216.597	1,470

Fig 2.8. Comparación de desempeño de distintos modelos (Yu. J . Et Al, 2014)

Ankit Chakravarti Et. Al. (2015) desarrollaron un modelo lluvia-escurrecimiento para una cuenca artificial dentro de un laboratorio, con el objetivo de medir el desempeño de esta tecnología en la descripción de procesos hidrológicos no lineales. Para el experimento se utilizaron distintos patrones de lluvia con diferentes duraciones y distintas configuraciones de la cuenca experimental con 4 pendientes distintas. El modelo utilizaba una arquitectura feedforward y como datos de entrada únicamente la precipitación inducida en la cuenca; y como dato de salida el caudal registrado. Los resultados del modelo fueron favorables en todos los casos, por lo cual los autores recomiendan su uso para la predicción en tiempo real de escurrimientos.

Frederik Kratzert Et. Al. (2018) utilizaron un modelo con arquitectura recurrente (ver *Capítulo III: Redes Neuronales Recurrentes R.N.R*) para crear un modelo lluvia-escurrecimiento en 241 cuencas de los Estados Unidos y comparar sus resultado con lo obtenidos por el modelo físico “*Sacramento Soil Moisture Accounting*” (SAC-SMA). Durante su investigación, los autores intentan demostrar cómo el funcionamiento de su modelo tiene paralelismos con respecto al proceso físico que intenta describir; no obstante, estas relaciones sean inferidas de los datos de calibración y, por tanto, no se muestren de manera explícita en la red.

Para el entrenamiento se utilizó la base datos de precipitación diaria, caudales registrados y deshielo de CAMELS dataset, con un registro de 15 años. Se realizaron 3 experimentos, el primero consistió en realizar el entrenamiento para cada una de las cuencas, utilizando una misma arquitectura; el segundo consistió en realizar un modelo regional agrupando las cuencas en 18 regiones hidrológicas, la premisa de este modelo era el que cada uno contara con el mayor número de información que le permitiera describir mejor el comportamiento de la cuenca, además de que pudiera ser utilizado para cuencas sin medición; el tercer experimento fue el de utilizar el modelo regional como modelo pre entrenado y volver a entrenarlos con los datos individuales de cada cuenca para comparar el desempeño con respecto a los modelos individuales.

Como resultado, en el primer experimento más de la mitad de los modelos obtuvieron un coeficiente de eficiencia de Nash-Sutcliffe de 0.65, y en todos los casos uno mayor a 0.63, lo que representa una aproximación adecuada de los modelos; además de que presentó un mejor desempeño que el modelo SAC-SMA, particularmente en cuencas áridas que no eran afectadas por deshielo. Los resultados del segundo experimento mostraron una mejora con respecto a los primeros en un 80 % de los casos, mientras que el otro 20 % empeoraba, atribuyendo esto a que estas regiones posiblemente no eran homogéneas, por lo que la información que transferían del modelo regional al local era errónea.

Finalmente, los resultados del último experimento fueron significativamente mejor que en los anteriores 2 modelos (15% de mejora en el coeficiente de Nash-Sutcliffe en promedio), particularmente en las cuencas que habían tenido un peor desempeño en los modelos regionales. De entre las conclusiones de los autores, se destaca la recomendación del uso de modelos regionales pre entrenados debido a su habilidad para transferir información; no obstante, también se señala la importancia de realizar una correcta regionalización para que esta transferencia sea efectiva.

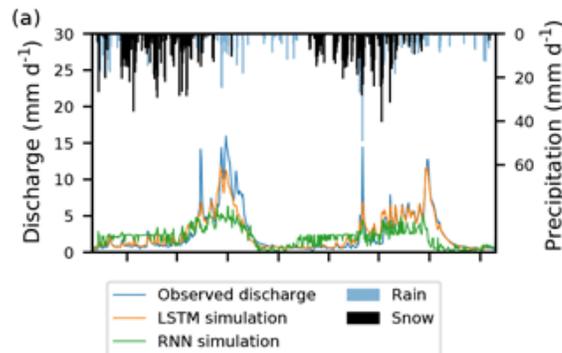


Fig 2.9. Comparación del Hidrograma modelado vs real para una de las cuencas de estudio (Yu. J. Et Al, 2014).

Frederik Kratzert et al. (2019) desarrollaron un modelo basado en redes neuronales con arquitectura recurrente enfocado en cuencas no aforadas; para ello realizaron diferentes experimentos, no obstante los mejores resultados se obtuvieron a través de un modelo similar al propuesto por Frederik Kratzert Et. Al. (2018), con la modificación de incluir información de 27 atributos de las cuencas como datos de entrada. Para entrenar el modelo se utilizó información de 531 cuencas ubicadas en Estados Unidos, agrupadas de manera aleatoria en 12 grupos; el modelo fue entrenado con los primeros 11 grupos y probado en el grupo 12. Este procedimiento se repitió 12 veces, con 10 modelos idénticos, los resultados fueron promediados y comparados con dos modelos de referencia : el *Sacramento Soil Moisture Accounting Model* (SAC-SMA) y el *NOAA National Water Model Reanalysis*. Los autores encontraron que su modelo tenía un desempeño igual o superior que el de los modelos de referencia en la gran mayoría de las cuencas; salvo por un conjunto de cuencas en las cuales su desempeño no fue satisfactorio. Los autores relacionan estos resultados atípicos a posibles errores en la medición de los atributos físicos con los cuales se alimentó la red.

Modelos de Pronóstico Hidrológico

Los modelos de pronóstico hidrológico son una herramienta que permiten estimar el volumen de escurrimiento en una cuenca para un determinado horizonte de tiempo. Estos son de suma importancia para la gestión integrada de recursos hídricos ya que; además de tener aplicaciones en la mitigación de desastres naturales relacionados con fenómenos hidrometeorológicos (inundaciones y sequías), son esenciales para la operación eficiente de infraestructura y la planeación eficaz de este recurso (WMO,2009). La escala de tiempo para la cual se pronostica puede variar de horas, días e inclusive años; esta depende principalmente del campo de aplicación del pronóstico. Sin embargo, factores tales como: la escala espacial, características del sitio de estudio (tiempo de concentración), así como la instrumentación e información disponible dentro de la cuenca pueden ser limitantes para establecer este horizonte (Sene,2009).

Existen diferentes enfoques para abordar el pronóstico dentro de una cuenca (Farmer & Vogel, 2016); no obstante, en este trabajo únicamente se abordará el enfoque determinista, debido a su aplicación dentro del campo de operación de infraestructura (Duong Vo & Gourbesville, 2016). Este enfoque tradicionalmente parte de un conjunto de datos hidrometeorológicos para, a través de un modelo lluvia-escurrimiento, conocer el volumen de agua que escurrirá en un punto dado de la cuenca dentro de un determinado horizonte de tiempo (Sene,2009).

En el caso de modelos de pronóstico hidrológicos cuyo horizonte es mayor al tiempo de concentración de la cuenca, estos requieren a su vez de pronósticos meteorológicos para su funcionamiento. Estos pronósticos tradicionalmente son generados a partir de modelos numéricos, los cuales resuelven las ecuaciones físicas que gobiernan el comportamiento de los procesos atmosféricos (ecuaciones de Navier-Stokes, ecuaciones de conservación de masa, energía y momentum, etc.) para determinar el estado de diversos parámetros atmosféricos; entre ellos, la precipitación esperada en una zona dentro de un determinado horizonte de tiempo (Kalnay E., 2002).

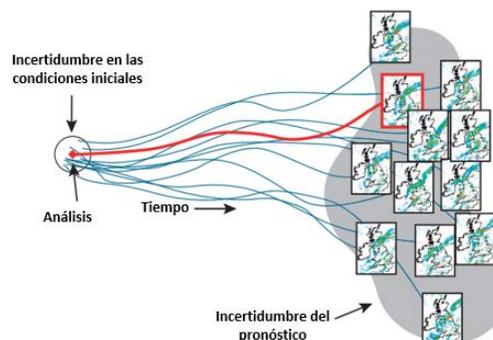


Fig 2.10 Esquema de la propagación de la incertidumbre en modelos meteorológicos (P. Bauer et al,2015).

Sin embargo, debido a la “naturaleza caótica” o no linealidad de dichos procesos atmosféricos (Selvam, 2010) , pequeños cambios en las condiciones iniciales pueden magnificarse y modificar considerablemente el resultado de salida (P. Bauer et al ,2015); por lo que aquellos modelos hidrológicos de pronóstico con un horizonte mayor al tiempo de concentración de la cuenca, tienen asociados una cierta incertidumbre relacionada con la precisión de la estimación meteorológica; la cual se magnifica conforme el horizonte de pronóstico incrementa (ver Fig 2.10). Además, debe adicionarse la incertidumbre propia del modelo lluvia-escurrimiento seleccionado.

Por lo qué, la principal limitación para implementar un modelo hidrológico de pronóstico es la disponibilidad de información meteorológica confiable para alimentarlo; es decir, que se cuente con información con una escala espacial y temporal que sirva para los fines del modelo, así como con un margen de error aceptable.

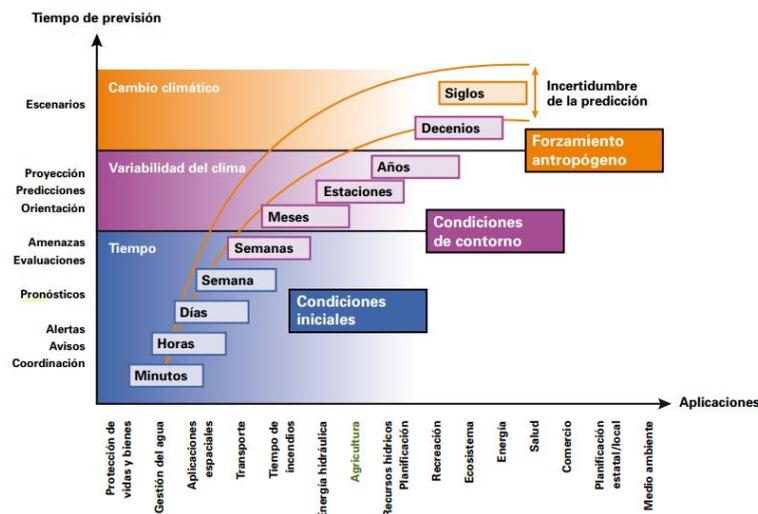


Fig 2.11 Esquema de los principales retos para el desarrollo de modelos de predicción para diferentes horizontes y su campo de aplicación (WMO, 2011)

En México, la recopilación de información y pronósticos hidrometereológicos está a cargo del Servicio Metereológico Nacional (SMN); y es publicada en boletines hidrometereológicos diarios por la Comisión Nacional del Agua (CONAGUA), quien es la dependencia federal encargada de la administración del agua en México (González Villarreal et al, 2014). Esta dependencia en 2021 cuenta con información de 12 Radares Meteorológicos, de los cuales 7 de ellos se encuentran ubicados en Estados Unidos y son operados por el *National Weather Service (NWS)* (CONAGUA,2021).

No obstante, esta dependencia también cuenta con información satelital que le permite alimentar dos modelos numéricos para el pronóstico del estado meteorológico, los cuales son (CENAPRED, 2017): el *Global Forecast System* (GFS), cuya resolución espacial es de 27.75 km^2 y tiene un horizonte de hasta 192 horas (Environmental Modeling Center, 2003); y el *Weather Research and Forecasting Model* (WRF), cuya resolución espacial es 12 km^2 y un horizonte de hasta 72 horas (Skamarock, 2008). Diversos estudios se han realizado para analizar la precisión de ambos modelos (Sergio Santana et al, 2012; CENAPRED, 2017); sin embargo, estos no son concluyentes sobre la superioridad de un modelo sobre otro y sus resultados varían dependiendo del sitio de estudio y el periodo de análisis. Por lo que se infiere que la precisión de dichos modelos puede variar dependiendo de las características del sitio (particularmente la orografía) y el tipo de tormenta que se presente (Houze, R. A. 1981).

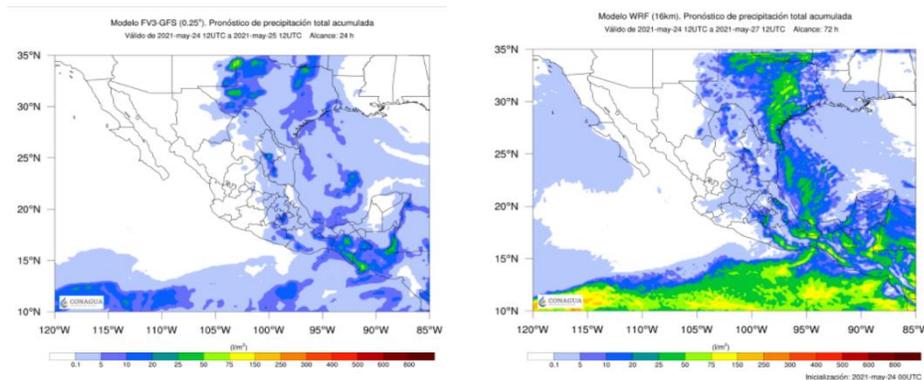


Fig 2.16 . Visualización de los modelos numéricos de pronóstico (SMN,2021)

Redes Neuronales Artificiales Aplicadas a Modelos de Pronóstico

Una alternativa al uso de modelos numéricos para realizar estimaciones sobre el pronóstico meteorológico es el uso de redes neuronales artificiales; esto debido a su facilidad para trabajar con procesos no lineales y grandes cantidades de información y datos (Subhajini,2018). Este tipo de modelos resultan especialmente útiles en sitios con escasa instrumentación o donde la información es difícil de obtener; ya que pueden alimentarse de distintos tipos de variables para analizar su tendencias y patrones de comportamiento, permitiendo anticipar el estado del tiempo (Elsafi, S. H. 2014). Adicionalmente, tiene la ventaja de que, dadas sus capacidades para asociar cualesquiera conjunto de variables, pueden integrar al proceso de pronóstico hidrológico (pronóstico meteorológico, modelo lluvia-escurrecimiento y tránsito por el cause) dentro un único modelo.

Estos modelos pueden alimentarse de variables atmosféricas (Fente, D. N., & Kumar Singh, D. ,2018) y/o de valores previos de las variables meteorológicas (Wu Et. Al., 2005; Frederik Kratzert Et. Al. ,2018) para realizar su estimación. Los principios de su funcionamiento son similares a los modelos Lluvia-Escurreimiento basados en R.N.A. antes mencionados; ya que tienen un funcionamiento de “caja negra” donde el análisis de las relaciones entre variables de entrada y salida se realizan dentro del modelo sin tener ningún paralelismo con las ecuaciones físicas que describen su comportamiento. No obstante, estos han sido utilizados por diversos autores y algunos otros ejemplos de su aplicación pueden ser consultados en: Mislan et al, 2015; Lee et al 2015; Abhishek et al, 2012.

Cabe mencionar que, dentro de esta bibliografía consultada, destaca el uso de un tipo especial de red neuronal artificial denominada “Recurrente”; particularmente en una configuración específica denominada LSTM (Long-Short Term Memory) (ver capítulo III: Redes Neuronales Recurrentes). Este tipo de redes están enfocadas para trabajar con series de tiempo; lo que les permite ser dinámicas; es decir, utilizar información sobre la evolución, patrones y tendencias de las distintas variables, tanto atmosféricas como hidrometeorológicas, para realizar su pronóstico.

Contexto del Sistema Cutzamala:

Las características y aspectos técnicos relacionados con el funcionamiento e infraestructura del Sistema Cutzamala (S.C.) se discuten a detalle dentro del Capítulo IV: “Descripción del Área de Estudio”; por lo que este apartado únicamente tiene como objetivo mostrar algunos aspectos generales relacionados al contexto y problemáticas identificadas dentro de la región en la bibliografía consultada.



Fig 2.12. Mapa de municipios y subcuencas (WBG,2015)

El S.C. está compuesto por seis subcuencas de captación (El Bosque, Villa Victoria, Valle de Bravo, Tuxpan, Ixtapan del Oro y Chilesdo- Colorines) las cuales transfieren parte de sus escurrimientos hacia la Zona Metropolitana del Valle de México (ZMVM). Estas subcuencas abarcan un área de 3, 419 km²; y se circunscriben dentro de 14 municipios ubicados en los estados de Michoacán y México (ver fig. 2.12), donde habitan 728,598 personas (INGEGI,2020). Salvo por Valle de Bravo, todos estos municipios se encuentran categorizados con un nivel medio-alto de marginación (carencias de acceso a educación, viviendas inadecuadas e ingresos insuficientes); un 42.1% de la población que habita esta región vive en condiciones de pobreza, de las cuales el 40% se encuentran en condiciones de pobreza extrema (WBG,2015).

La principal actividad económica dentro de esta región es la agricultura, particularmente en su modalidad de temporal; siendo el Maíz (71 % de la superficie cultivada) y la Avena Forrajera (20 % de la superficie cultivada) los cultivos más importantes dentro de la región (SAGARPA,2020). En 2020 la superficie cultivada dentro de las subcuencas de captación que conforman al S.C. fue de 121,180 Ha (ver fig. 2.13), de las cuales: 15,344 Ha (13 %) de riego y 105,863 Ha (87%) cultivo de temporal; lo que representa el 35 % de la superficie total de captación de las subcuencas (SAGARPA,2020). El turismo y la piscicultura, en menor medida,

también son actividades económicas importantes que se desarrollan dentro de esta zona (WBG,2015).

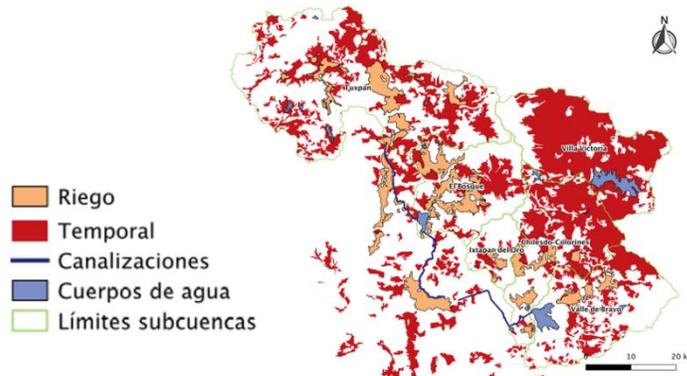


Fig 2.13. Mapa Zonas de Riego en el S.C. (WBG,2015)

En las últimas décadas, el incremento poblacional en la región y el consecuente incremento en las actividades productivas sin variaciones en su estructura (enfocadas principalmente hacia actividades agrícolas) han incrementado la presión sobre la explotación de recursos naturales, particularmente el agua (WBG.2015). Esto debido a que el incremento en la superficie de cultivo reduce la disponibilidad; adicionalmente en algunos casos, debido a los retornos del riego, se merma la calidad del recurso disponible. Esta situación ha planteado el escenario en el cual, los usos propios de las subcuencas podrían entrar en competencia con las extracciones que se realizan para abastecer de agua a la Z.M.V.M.

Debido a esto, a la importancia que tiene esta fuente para el abastecimiento de la Z.M.V.M., y a otros factores tales como: la alta exposición que tiene esta fuente de abastecimiento ante fenómenos o sucesos que pudiesen comprometer su operación (sequías, fallas de suministro eléctrico, fallas mecánicas, etc); el riesgo de que se genere una disminución en la disponibilidad de la cuenca producto del cambio climático; estos otros aspectos. Este sistema ha sido objeto de estudio por diversos autores, con el objetivo de generar políticas y planes que permitan gestionar de manera sustentable las subcuencas que lo conforman.

El Banco Mundial (2015), por ejemplo, realizó un estudio integral sobre el funcionamiento del S.C. en cada una de las subcuencas que lo integran, considerando aspectos ambientales, socioeconómicos, infraestructura, gobernanza, entre otros; con el objetivo de brindar un diagnóstico sobre el manejo de las subcuencas que lo integran y del funcionamiento del sistema en general.

De entre las conclusiones obtenidas dentro de este estudio destaca que: si bien, aún existe un equilibrio hídrico entre las entregas del sistema y las demandas internas de las subcuencas que lo integran, este mostraba una tendencia a estar comprometido en el corto plazo debido al incremento significativo en la demanda para uso agrícola en la zona. Otro aspecto identificado es la falta de control referente a extracciones clandestinas en el canal Tuxpan-El Bosque-Colorines, dónde se contabilizaron al menos 1,400 tomas irregulares.

Asimismo, se detectaron filtraciones en la cortina de la presa “El Bosque”; las cuales si bien, no representan un riesgo estructural para la estabilidad de la presa, se estima que generan pérdidas equivalentes a un caudal de entre $1\text{ m}^3/\text{s}$ y $2\text{ m}^3/\text{s}$; con valores extremos que pueden llegar a los $4\text{ m}^3/\text{s}$, dependiendo del nivel de la presa. Cabe mencionar que, de acuerdo con el estudio, estas filtraciones han sido aprovechadas por los habitantes aguas abajo de la presa para desarrollar actividades agrícolas.

Finalmente, se señala la carencia de instrumentación dentro del sistema y las subcuencas que lo conforman; ya que el número de estaciones climatológicas instaladas es insuficiente para el área e importancia del sistema, adicionalmente de que no se encontraron estaciones hidrométricas en operación al momento de realizar el estudio.

Por su parte, Escolero et al (2016) realizaron un estudio de vulnerabilidad para las fuentes de abastecimiento de agua potable para el Valle de México, incluido el S.C. Destacando el hecho de que el S.C. es la fuente con un mayor índice de vulnerabilidad debido a factores tales como: deficiencias en el mantenimiento y operación de la infraestructura; extracciones clandestinas, altos niveles de degradación hídrica en las cuencas de captación, presencia de contaminantes asociados a actividades agrícolas y descargas de agua residual, así como conflictos sociales y una baja eficiencia económica del sistema.

Bunge et al (2012) elaboraron un modelo dinámico conceptual del S.C. utilizando la herramienta VENSIM; donde se modeló tanto la disponibilidad, como distintos escenarios de demanda para los diferentes usos consultivos con el objetivo de estimar el déficit hídrico en cada una de las subcuencas que lo conforman. Al respecto los autores concluyeron que las subcuencas que presentan mayor exposición a presentar una condición de déficit hídrico, y que al mismo tiempo presentan un mayor crecimiento en su demanda son: El Bosque, Tuxpan e Ixtapan del Oro. Asimismo, destaca el efecto que tienen las tomas clandestinas en la disponibilidad del sistema; y lo importante que es contar con información de precipitación confiable para poder hacer estimaciones.

Hernández Vargas (2017) desarrolló un modelo conceptual del Sistema Cutzamala en WEAP a través del método de humedad del suelo; con el objetivo de evaluar su funcionamiento actual, así como escenario de cambio climático. Cabe mencionar que este modelo tuvo que ser diseñado con una escala temporal mensual para adaptarse a la escala temporal de la información disponible referente a cultivos; así mismo que fue calibrado con los registros disponibles (1970-2000) y podría requerir ser actualizado debido al intenso cambio global en la cuenca. Se destaca que entre las conclusiones, se menciona la falta de información y la incertidumbre asociada a esta como un impedimento para realizar el análisis bajo este método; particularmente en la cuenca Ixtapan del Oro.

	Punto de control	Gráfico	Acumulación	RRD	EV	INS	Calificación final
Calibración	EH Río Grande	A	A	MB	MB	MB	Muy Bueno
	EH Río Chiquito	A	A	MB	MB	MB	Muy Bueno
	EH Zitácuaro	A	A	MB	MB	MB	Muy Bueno
	Presa Valle de Bravo	A	A	S	MB	B	Bueno
	EH San José Malacatepec	A	IA	MB	MB	MB	Poco Satisfactorio
	Presa Villa Victoria	A	IA	S	B	S	Poco Satisfactorio
Validación	EH Río Grande	A	A	MB	B	B	Bueno
	EH Río Chiquito	A	A	B	S	B	Satisfactorio
	EH Zitácuaro	A	A	MB	B	B	Bueno
	Presa Valle de Bravo	A	A	MB	MB	MB	Muy Bueno
	EH San José Malacatepec	A	IA	MB	MB	MB	Poco satisfactorio
	Presa Villa Victoria	A	IA	B	MB	B	Poco satisfactorio

Fig 2.14 . Resultados del modelo desarrollado por Hernández Vargas (Hernández Vargas 2017)

Capítulo III Marco Teórico

Redes Neuronales Artificiales

Las redes neuronales artificiales (RNA) son una técnica, enmarcada dentro del aprendizaje automático y la inteligencia artificial, que emula el mecanismo mediante el cual aprenden los organismos biológicos (Aggarwal C.C, 2018); particularmente, a la manera en la cual es generado el conocimiento procedimental. Este tipo de conocimiento es también llamado “memoria muscular”, y forma parte de la memoria inconsciente, la cual está relacionada con el aprendizaje de tareas mecánicas tales como andar en bicicleta; las cuales se caracterizan por ser difíciles de verbalizar o explicar la manera en la cual se realizan, debido a que este conocimiento es generado en una región primitiva del cerebro separada el “consciente” (Zichlin M., 2011).

El cerebro humano está conformado por aproximadamente 10^{11} células nerviosas llamadas neuronas, las cuales son la unidad básica de procesamiento del sistema nervioso. Estas se relacionan entre si a través de conexiones sinápticas, las cuales permiten llevar a cabo todo tipo de acciones (movimiento motor, procesos biológicos, pensamiento, etc.) (NIDA ,2010). Cada que se realiza una acción, un determinado grupo de neuronas son activadas mediante estas conexiones; cuando dicha acción es reconfortante para la persona que la realiza (se cumple el objetivo esperado), el cerebro genera dopamina, el cual es un neurotransmisor que incrementa la plasticidad del cerebro (la capacidad de generar y fortalecer conexiones sinápticas) mejorando la conectividad de aquellas neuronas que permitieron llevar a cabo dicha acción (Barbara Knowlton et al, 2017).

De esta forma, a través de la repetición o el “entrenamiento”, el cerebro humano va adaptando las conexiones neuronales biológicas para mejorar su desempeño en la ejecución de dicha actividad; generando así lo que se denomina como “conocimiento procedimental”.

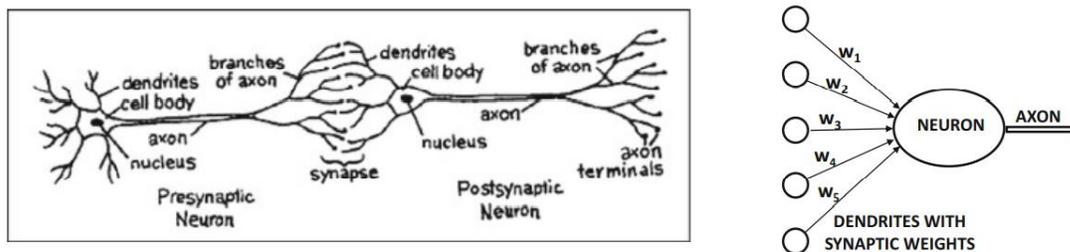


Fig 3.1. Comparación entre red neuronal biológica y artificial (Aggarwal C.C, 2018)

De manera análoga, las redes neuronales funcionan a partir de nodos, también denominados “neuronas”, los cuales son la unidad básica de procesamiento de la red. Estos nodos se relacionan entre si a través de conexiones, las cuales guían el flujo de la información; dichas conexiones tienen asociadas un factor de peso, el cual determina la relevancia que tiene la información provista por cada uno de los nodos para el proceso que está ejecutando el nodo subsecuente. El aprendizaje es generado mediante un proceso denominado “entrenamiento”, el cual consiste en procesar un determinado número de ejemplos; durante este proceso, la red va ajustando los pesos de las conexiones para aproximarse a la solución real, es decir, va aprendiendo la relación que existe entre los datos de entrada y salida (ver Aprendizaje en una RNA).

Los nodos en una red se agrupan en capas, las cuales pueden contener uno o más nodos. La red se compone por 3 tipos de capas: capa de entrada, la cual contiene los *inputs* o datos de entrada de la red; una o más capas ocultas, las cuales están encargadas de realizar lo cálculos matemáticos para mapear los inputs en outputs; y una capa de salida, la cual contiene los *outputs* o el resultado. Se denomina “profunda” a una red neuronal cuando esta contiene más de una capa oculta (Aggarwal C.C, 2018).

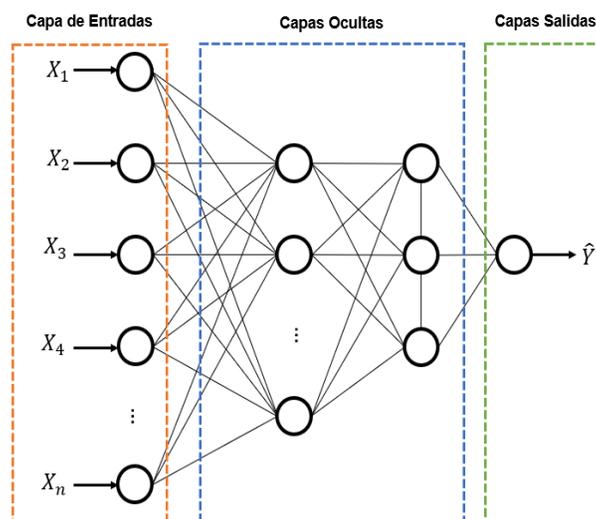


Fig 3.2. Diagrama Red Neuronal (Elaboración propia)

Dependiendo de la arquitectura de la red, es decir a la manera en que se configure el flujo de información y la interacción entre las distintas capas de la red, los nodos pueden recibir información de la capa anterior (arquitectura “*Feed-Forward*”), de los nodos contiguos dentro de una misma capa (arquitectura “*Recurrente*”), o de otra red (arquitectura “*Modular*”).

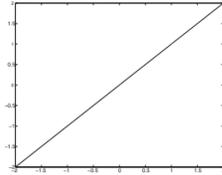
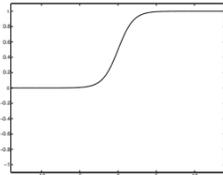
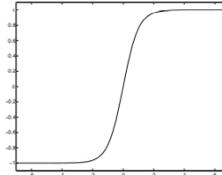
Cada nodo está programado para realizar la suma ponderada de la información recibida utilizando como referencia los pesos de cada una de las conexiones ligadas a él; posteriormente, a este resultado se le aplica una función de activación (ϕ). Estas funciones permiten determinar la activación de la neurona (si el valor de salida será diferente de 0) y transformar los valores de entrada con respecto al valor esperado de salida. Matemáticamente, cada nodo está representado de la siguiente manera:

$$\hat{h}_j = \phi(\bar{X} * \bar{W}_j - b_j) \quad Eq. 3.1$$

Ecuación 3.1. Donde: \hat{h}_j es el valor de salida del nodo j ; \bar{X} es el vector $[x_1, \dots, x_n]$ que contiene los datos de entrada; \bar{W}_j es el vector $[w_1, \dots, w_n]$, que contiene los pesos de cada una de las conexiones; b_j es el coeficiente de sesgo asociado al nodo j ; ϕ es la función de activación.

Existe una amplia variedad de funciones de activación (ver tabla 3.1), las cuales pueden dividirse en dos grupos: lineales y no lineales; en ambos casos las funciones deben tener la cualidad de ser diferenciable. No obstante, para el caso particular de las redes neuronales con más de una capa oculta (redes profundas), resulta necesario el uso de funciones no lineales, debido a que estas permiten agrupar nodos y que cada uno de estos se especialice en una dimensión específica de la relación que se pretende describir; ya que de lo contrario la red se simplificaría en un polinomio de grado uno con limitaciones en cuanto a su capacidad de mapeo de relaciones complejas (Sharma, S.,2017).

Tabla 3.1 Ejemplos para Funciones de Activación

Nombre	Identidad	Sigmoide	tanh
Función	$\phi(z) = z \quad Eq. 3.2$	$\phi(z) = \frac{1}{1 + e^{-z}} \quad Eq. 3.3$	$\phi(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad Eq. 3.4$
Derivada	$\phi'(z) = 1 \quad Eq. 3.5$	$\phi'(z) = \phi(1 - \phi) \quad Eq. 3.6$	$\phi'(z) = 1 - \phi^2 \quad Eq. 3.7$
Dominio	$(-\infty, \infty)$	$(0, 1)$	$[-1, 1]$
Gráfica			

Ecuaciones 3.2-3.7: Donde ϕ es la función de activación seleccionada, ϕ' es la primera derivada de la función ϕ , z representa los cálculos internos de cada nodo previos a la función de activación (ver ecuación 3.1) de manera que: $z = \bar{X} * \bar{W}_j - b_j$

Aprendizaje en una RNA

Dentro del campo de la inteligencia artificial y el aprendizaje automático, existen dos ramas asociadas al uso de RNA (Sandro Skansi,2018):

Aprendizaje supervisado: Consiste en entrenar a la red con un conjunto de datos “etiquetados”, es decir a los cuales se tiene conocimiento del valor esperado de salida; de manera que la red aprenda las relaciones existentes entre ambos conjuntos.

Aprendizaje no supervisado: El entrenamiento se realiza con un conjunto de datos “sin etiquetar”; por lo que la red aprende a inferir atributos dentro del conjunto de datos para agruparlos o codificarlos.

Dentro de este trabajo, únicamente se abordará el enfoque de “Aprendizaje Supervisado”; ya que por sus características es el que mejor ajusta al problema de crear un modelo lluvia-escorrimento y de pronóstico (ver capítulo II). Asimismo, cabe mencionar que, dentro del campo del aprendizaje automático, existe una tercera rama denominada “Aprendizaje por refuerzo”. Esta consiste en exponer a una IA (Inteligencia Artificial) o “agente” a realizar una determinada tarea dentro de un ambiente específico; de manera que, a través de un sistema de recompensas, el agente mejore su desempeño (R.S. Sutton & A.G. Barto 2018). A diferencia de los enfoques antes mencionados, en este no se cuenta con ningún conjunto de datos de entrenamiento; por lo que el aprendizaje se lleva a cabo mediante la interacción prueba-error (R.S. Sutton & A.G. Barto 2018). No obstante, este se enfoca en otro tipo de herramientas diferentes a las RNA, tales como: Redes de Markov, Programación Dinámica y Métodos de Monte Carlo (Richardson, M., Domingos,2006; R.S. Sutton & A.G. Barto 2018); por lo que tampoco será abordado a más a detalle dentro de este trabajo.

A su vez, los algoritmos de aprendizaje supervisado pueden dividirse en dos categorías de acuerdo con la tarea que realizan (Hastie T. et al, 2008):

Clasificación: Predice datos cualitativos. La red es entrenada con un conjunto de datos numéricos (de n dimensiones) para el cual, cada elemento tiene asociado algún atributo. De manera que la red aprende a relacionar un conjunto de datos cuantitativos a una categoría o clase.

Regresión: Predice datos cuantitativos. La red es entrenada con un conjunto de datos numéricos (de n dimensiones) asociados a un segundo grupo de datos numéricos (de n dimensiones), por lo que funcionan como un “aproximador universal”; es decir, adquieren la capacidad de “aproximar cualquier función continua entre cualesquiera pares de espacios euclidianos, uniformes o compactos” (Anastasis Kratsios, 2021).

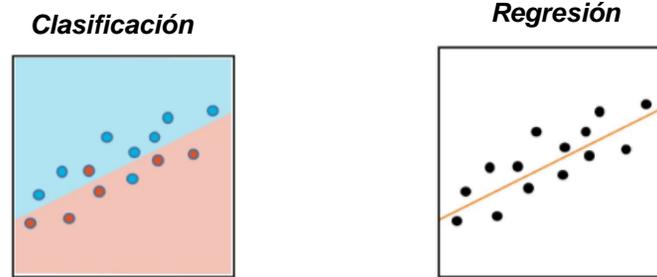


Fig 3.3. Ilustración problema de clasificación vs regresión. (Matanga, Yves, 2017)

En ambos casos se sigue el mismo proceso de “entrenamiento”; no obstante, a continuación se describe dicho proceso enfocado hacia problemas de regresión haciendo énfasis en las consideraciones necesarias para este:

El proceso de entrenamiento consiste en alimentar a la red con un conjunto de datos, para los cuales se conoce el valor esperado de salida. Este proceso tiene dos etapas (ver figura 3.4), las cuales se repiten de manera iterativa hasta alcanzar un valor óptimo (Rumelhart, D. E et al,1995):

Forward phase (Etapa de Análisis): En esta etapa, se ingresa un conjunto de datos a la red; generando una serie de cálculos en cada uno de los nodos que derivan en un resultado final. Posteriormente, este resultado es evaluado mediante un parámetro específico denominado “función de coste” (ver tabla 3.2). Durante la primera iteración los parámetros de cada uno de los nodos (los valores de \bar{W}_j y b_j) son generados de manera aleatoria, posteriormente se evalúan con los parámetros ajustados previamente.

Tabla 3.2 Diferentes funciones de coste para problemas de regresión

Error Medio Cuadrático (MSE)

$$C = (\hat{y}_i - y_i)^2 \text{ Eq.3.8}$$

Error Medio Absoluto (MAE)

$$C = |\hat{y}_i - y_i| \text{ Eq.3.9}$$

Huber Error

$$C = \begin{cases} (\hat{y}_i - y_i)^2 & \text{para } |\hat{y}_i - y_i| \leq \alpha \\ \alpha * \left(|\hat{y}_i - y_i| - \frac{1}{2} \alpha \right) & \text{para } |\hat{y}_i - y_i| > \alpha \end{cases} \text{ Eq.3.10}$$

Ecuaciones 3.8-3.10: Donde C es el error de la red, y_i es el valor real, \hat{y}_i es el valor estimado por la red; α es un parámetro arbitrario definido tal que $\alpha \in R^+$

Backward phase (Etapa de Retro-análisis): Durante esta etapa se obtiene la derivada parcial de la función de coste con respecto al vector \bar{W}_j y el coeficiente b_j en cada nodo; de manera que se pueda ajustar estos parámetros en sentido donde esta función se minimice; a esta técnica se le conoce como “Descenso del Gradiente”. Este proceso se genera de manera retrospectiva, es decir comenzando en el nodo de salida hacia los nodos de entrada.

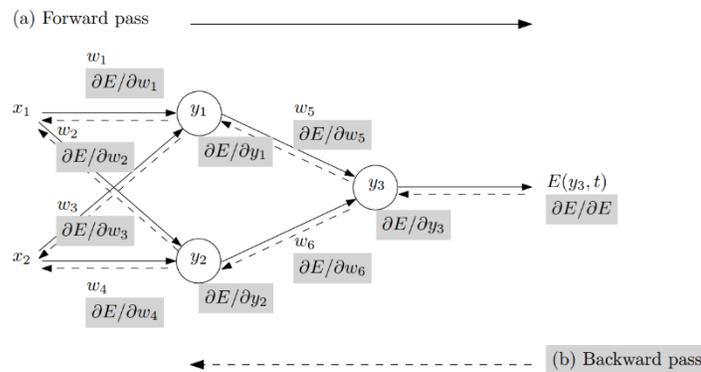


Fig 3.4. Ilustración del proceso de entrenamiento (Baydin et al, 2018)

Debido a que es un proceso iterativo, es necesario contar con una gran cantidad de datos que permitan al modelo ajustarse lo mejor posible; no obstante, esto puede incrementar considerablemente el tiempo de cómputo del proceso de entrenamiento. En la práctica, es recomendable que cada uno de estos pasos sea aplicado a un “mini-lote” de datos o “minibatch” (Maximilian Sieb et al, 2018), los cuales son una muestra representativa del conjunto de datos de entrenamiento. Esto permite tener un ajuste más robusto ante datos atípicos o “ruido”; así como reducir considerablemente los recursos computacionales requeridos y tiempo de cómputo durante el entrenamiento.

A esta técnica se le denomina “Mini Batch Gradient Descent” o “Descenso del Gradiente por Mini-lotes” (Khirirat, S et al , 2017); cabe mencionar que se denomina “mini-lote” para diferenciarse de la técnica “Batch Gradient Descent” o “descenso del gradiente por lote”; la cual consiste en computar en una iteración el total de datos para entrenamiento. No obstante, esta técnica no es común dentro del campo de las RNA; por lo que de acuerdo con la bibliografía consultada, por simplicidad suele denominársele únicamente como lote (*batch*) a cada una de las muestras del conjunto de datos.

Cabe mencionar que en la actualidad no existe una metodología exacta para determinar el tamaño o número de lotes en los cuales debe dividirse el conjunto de entrenamiento; por lo cual queda a consideración de cada autor, tomando en cuenta

el tipo de datos que maneja (si presenta alguna periodicidad) y la capacidad de cómputo (a mayor tamaño de lote, mayor será el requerimiento de memoria temporal) (Ian Goodfellow, 2016).

De este modo, una vez que se cuenta con el conjunto de lotes para realizar el entrenamiento; este procedimiento puede resumirse de la siguiente manera: Se ingresa cada uno de los valores contenidos en el primer lote de datos; para posteriormente calcular el error promedio de ajuste de acuerdo con la función de coste seleccionada (\bar{C}) (*Forward phase*). Se calcula el gradiente de esta función en cada uno de los nodos con relación a sus parámetros (\bar{W}_j y b_j); finalmente se ajustan estos parámetros de acuerdo con las ecuaciones 3.11 y 3.12 (*Backward phase*).

Este proceso se repite con el siguiente lote de datos, y así sucesivamente; cuando se ha terminado de procesar el total de los lotes de entrenamiento, se dice que se ha procesado una “época” (*epoch*). El proceso de entrenamiento puede continuar el número de “épocas” que sea necesario para minimizar la función de coste (\bar{C}) a un valor aceptable.

$$\bar{W}_j^{(s)} = \bar{W}_{jl}^{(s-1)} - \eta \frac{\partial \bar{C}}{\partial \bar{W}_j^{(s-1)}} \quad \text{Eq3.11} \qquad b_j^{(s)} = b_j^{(s-1)} - \eta \frac{\partial C}{\partial b_j^{(s-1)}} \quad \text{Eq3.12}$$

Ecuaciones 3.11 - 3.12. Donde $\bar{W}_j^{(s)}$ es el vector $[w_1, \dots, w_n]$, que contiene los pesos de cada una de las conexiones del nodo j en la iteración s ; $b_j^{(s)}$ es el coeficiente de sesgo asociado al nodo j en la iteración s ; C es la función de coste; y η es la tasa de aprendizaje de la red.

La tasa de aprendizaje (η) es un “hiperparámetro” (un parámetro externo al proceso de aprendizaje, el cual es seleccionado manualmente) cuyo valor puede variar entre 0 y 1. Este determina la magnitud del ajuste a los parámetros de la red; por lo cual está relacionado con la velocidad del entrenamiento, la precisión de la RNA y la cantidad de ciclos que deberá someterse a entrenamiento la red. Tasas de aprendizaje altas permitirán incrementar la velocidad de ajuste de la red, reduciendo el número de ciclos; no obstante, valores muy altos harán que nunca convergerá a un mínimo. Por el contrario, valores pequeños permitirán tener un ajuste más fino de los parámetros de la red; sin embargo, incrementarán exponencialmente el tiempo de entrenamiento y corren el riesgo de converger en un mínimo local (LeCun, Y. A, 2012).

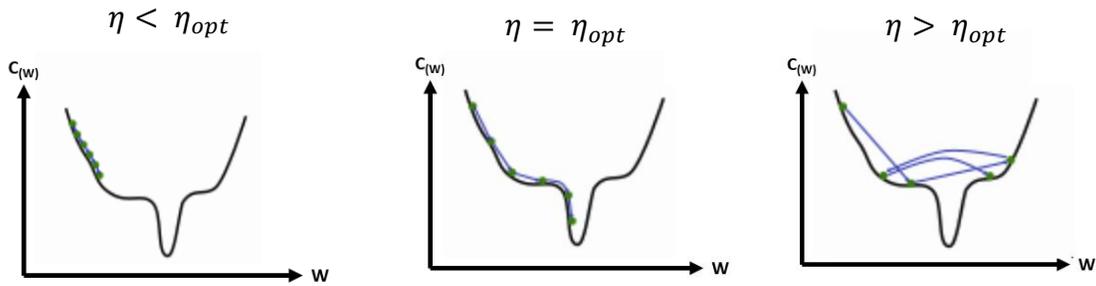


Fig 3.5. Comparación de optimización de la función de coste para diferentes tasas de aprendizaje. (B. D. Hammel, 2019)

Debido a la importancia de este hiperparámetro en el desempeño de la red, se han desarrollado diversos algoritmos para optimizar el ajuste de este (Sebastian Ruder,2017); siendo el más utilizado, dentro de la bibliografía consultada, “Adam” o “Adaptive Moment Estimation” (D. P. Kingma and J. Ba, 2014). Este permite ajustar de manera dinámica la tasa de aprendizaje para cada uno de los parámetros de la red, a partir del primer y segundo momento estadístico del gradiente. El algoritmo se describe a continuación:

- 1- Se calcula el primer (m_s) y segundo (v_s) momentos estadísticos a través de las ecuaciones 3.13 y 3.14 respectivamente. Durante la primera iteración m_{s-1} y v_{t-1} se les asigna el valor 0.

$$m_s = \beta_1 * m_{s-1} + (1 - \beta_1) \frac{\partial C}{\partial W_j^{(s-1)}} \quad Ec3.13$$

Ecuación 3.13: Donde m_s es el primer momento estadístico en la iteración s ($s=0 \rightarrow m_{s-1} = 0$; β_1 es la tasa de decrecimiento exponencial para los momentos estimados ($\beta_1 \in [0,1)$).

$$v_s = \beta_2 * v_{t-1} + (1 - \beta_2) \left(\frac{\partial C}{\partial W_j^{(s-1)}} \right)^2 \quad Ec3.14$$

Ecuaciones 3.14 : Donde v_s es el segundo momento estadístico en la iteración s ($s=0 \rightarrow v_{s-1} = 0$); β_2 es la tasa de decrecimiento exponencial para los momentos estimados ($\beta_2 \in [0,1)$).

- 2- Debido a que m_s y v_s son inicializados con un valor de 0, estos se encuentran sesgados por lo que requieren ser ajustados mediante las ecuaciones 3.15 y 3.16 respectivamente (ver D. P. Kingma and J. Ba, 2014)

$$\widehat{m}_s = \frac{m_s}{1 - \beta_1^s} \quad Ec3.15$$

$$\widehat{v}_s = \frac{v_t}{1 - \beta_2^s} \quad Ec3.16$$

Ecuaciones 3.15 - 3.16 : Donde \widehat{m}_s es el primer momento estadístico con corrección de sesgo; \widehat{v}_s es el segundo momento estadístico con corrección de sesgo; y β_1 y β_2 son las tasas de decrecimiento exponencial para los momentos estimados ($\beta_1, \beta_2 \in [0,1)$).

3- Los parámetros en cada nodo son actualizados mediante la ecuación 3.17:

$$\overline{W}_{j,l}^{(s)} = \overline{W}_{j,l}^{(s-1)} - \widehat{m}_s \left(\frac{\alpha}{\sqrt{\widehat{v}_t + \varepsilon}} \right) \quad Ec3.17$$

Ecuaciones 3.17 : Donde $\overline{W}_j^{(s)}$ es el vector $[w_1, \dots, w_n]$, que contiene los pesos de cada una de las conexiones del nodo j en la iteración s ; \widehat{m}_s es el primer momento estadístico con corrección de sesgo; \widehat{v}_s es el segundo momento estadístico con corrección de sesgo; α es el límite superior de la tasa de ajuste; ε es un coeficiente de ajuste inicial ($\varepsilon \rightarrow 0$). Se recomienda utilizar: $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ y $\varepsilon = 10^{-8}$ (D.P. Kingma and J. Ba, 2014). Las ecuaciones para el coeficiente de sesgo son análogas.

Problema de Sub-ajuste y Sobre-ajuste.

A la habilidad de la red para desempeñarse de manera adecuada ante nuevos datos de entrada (diferentes a aquellos con los cuales fue entrenado) se le denomina “generalización del conocimiento”. Durante el entrenamiento de una red, el objetivo es minimizar la función de coste; no obstante, este no necesariamente es representativo de la precisión de la red al exponerse a nuevos datos. A la brecha que existe entre la precisión del modelo durante el proceso de entrenamiento y su precisión al exponerse a nueva información se le denomina “error de generalización” (Ian Goodfellow et al,2016).

Para cuantificar el error de generalización en una red es necesario dividir al total del conjunto de datos disponibles en dos subconjuntos: el de entrenamiento, compuesto por un 80% - 90% de los datos, que será utilizado para ajustar los parámetros de la red; y el de validación, compuesto por el conjunto restante, que será utilizado para evaluar su desempeño (Ian Goodfellow et al,2016).

El término “*sub-ajuste*”, es empleado cuando la red no es capaz de aprender las relaciones intrínsecas entre los datos; es decir, no es capaz de reducir la función de coste a un valor aceptable. Mientras que el término “*sobreajuste*”, significa que la red no es capaz de generalizar dicho conocimiento; esto ocurre cuando la brecha entre el error promedio en el conjunto de entrenamiento y el de validación es demasiado grande (Ian Goodfellow et al,2016).

Entre los motivos que pueden propiciar el *subajuste* destaca una mala selección de hiperparámetros (funciones de activación, tasa de aprendizaje, etc.) así como limitaciones en cuanto a su *espacio hipotético* o a los grados de libertad que tenga la red para adaptarse a los datos (este parámetro está ligado al número de neuronas y capas existentes en la red). Por otro lado, el *sobreajuste* se debe esencialmente a un entrenamiento excesivo de la red, donde los parámetros se han ajustado a un grado tal que modelan el ruido propio de los datos (Haider Allamy, 2014).

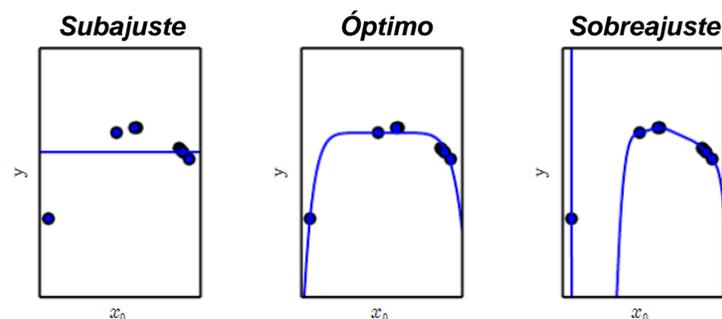


Fig 3.6. Comparación entre subajuste, óptimo y sobreajuste de la función de coste

(Ian Goodfellow et al,2016)

Redes Neuronales Recurrentes (RNR):

La arquitectura más simple en una RNA profunda es en configuración “*Feed-forward*”; donde cada uno de los nodos dentro de una capa envían información únicamente hacia la capa siguiente. Esta arquitectura permite inferir eficientemente la relación existente entre los datos de entrada y salida; sin embargo, supone que los datos de entrada son independientes y no existe ninguna relación entre las entradas anteriores y la actual (Samir B et al,2001).

Las redes neuronales recurrentes son un tipo especial de RNA en las cuales los nodos se retroalimentan (Fausett, 1994); es decir están configuradas de manera que las conexiones entre los nodos dentro de una capa generen un “ciclo”, lo que les permite tener una “memoria” e inferir también las relaciones existentes dentro del conjunto de datos de entrada.

Este tipo de redes son eficaces trabajando con datos secuenciales (series de tiempo y lenguaje); ya que gracias a su "memoria" pueden utilizar la información de entradas anteriores para inferir tendencias o patrones sobre su comportamiento, y extrapolar este conocimiento para describir el comportamiento de los datos de salida. Por lo que son útiles en tareas de predicción univariada y multivariada (*forecasting*) y de procesamiento de texto (Samir B et al,2001).

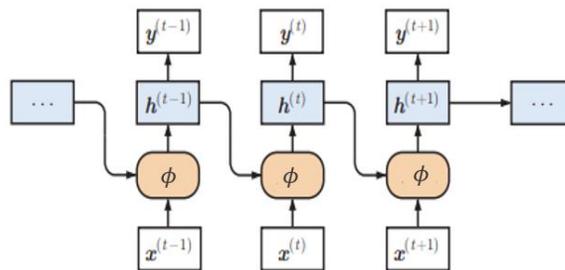


Fig 3.7 Ilustración del flujo de información en una red neuronal recurrente (Ming, Y. et al 2017)

Para trabajar con datos secuenciales, este tipo de redes requiere añadir una nueva dimensión al vector \bar{X} que contiene los datos de entrada para convertirlo en una matriz de dimensiones $v \times T$; donde v es el número de variables con los cuales se va a alimentar a la red, y T número de datos secuenciados previos que se tomarán en cuenta (para el caso particular de series de tiempo, T es la ventana temporal de datos previos). Dentro de la literatura, estas matrices de datos reciben el nombre de “tensores”.

Para crear una capa con arquitectura recurrente cada nodo dentro de esta se configura para trabajar como si fuera una capa en sí, con un número “ T ” de sub-nodos iguales (con los mismos parámetros \bar{W} y v , ver ecuación 3.18); a los cuales

se les denomina “Estados ocultos” (*hidden state*). Estos sub-nodos deben estar conectados entre sí de manera secuencial, tal que el subnodo (h^t) reciba como dato de entrada $x^{(t)}$ y el valor obtenido por el subnodo que lo precede dentro del arreglo (h^{t-1}) (ver Fig. 3.7). Matemáticamente, los subnodos en una red neuronal recurrente pueden expresarse de la siguiente manera (Ming, Y. et al 2017):

$$\hat{h}_j^{(t)} = \phi(\bar{x}^t * \bar{W} + v * \hat{h}_j^{(t-1)}) \quad \text{Eq 3.18}$$

Ecuación 3.18: Donde: $\hat{h}_j^{(t)}$ es el valor de salida del nodo j para el valor de la secuencia t ; \bar{x}^t es el vector $[x_1^t, \dots, x_v^t]$ que contiene los datos de entrada para el periodo t ; \bar{W} es el vector $[w_1, \dots, w_n]$, que contiene los pesos de cada una de las conexiones, v es el factor de peso para el estado $\hat{h}_j^{(t-1)}$; y ϕ es la función de activación.

La salida de un subnodo con arquitectura recurrente puede configurarse de dos maneras (ver fig 3.6): En modo “secuencia-secuencia”, donde a cada uno de los subnodos $\hat{h}_j^{(t)}$, además de estar conectados al subnodo $\hat{h}_j^{(t+1)}$, se conectan también a los nodos de la siguiente capa; de manera que se obtiene como resultado del procesamiento una segunda secuencia de datos de longitud T , esto permite agrupar más capas con arquitectura recurrente. El segundo esquema es en modo “secuencia-a uno”, donde los subnodos $\hat{h}_j^{(t)}$ únicamente se encuentra conectado al subnodo $\hat{h}_j^{(t+1)}$, y el único subnodo conectado a la siguiente capa es $\hat{h}_j^{(T)}$; de modo que el resultado de este tipo de arreglo es un único valor, por lo cual solo pueden ser anidada a una capa con arquitectura convencional (Ming, Y. et al 2017).

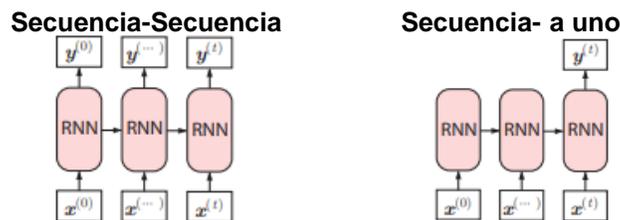


Fig 3.8 Ilustración Outputs (Ming, Y. et al 2017)

El entrenamiento de una RNR es similar al de una red neuronal convencional, ya que ambos utilizan los mismos algoritmos mencionados anteriormente (ver Aprendizaje en una RNA). No obstante, en estas redes se utiliza una variante del proceso de retropropagación del error denominada “*Backpropagation Through Time*” (BPTT) o “retropropagación a través del tiempo” (Werbos,1990). Este consiste en “desplegar” a cada uno de los nodos con arquitectura recurrente e ir calculando el gradiente de la función de coste con respecto a los parámetros de cada uno de los subnodos $\hat{h}_j^{(t)}$; debido a que todos ellos comparten los mismos

parámetros, el ajuste de ellos se realiza con respecto al error acumulado, es decir con respecto a la derivada de la función de coste con respecto a $\hat{h}_j^{(1)}$.

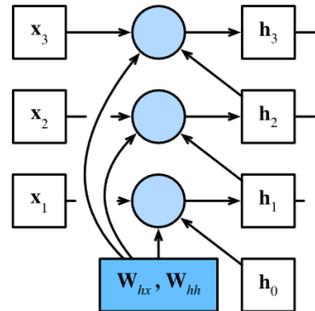


Fig 3.9 Esquema del desdoblamiento de un nodo con arquitectura recurrente (Zhang et al, 2021)

La principal limitación de este tipo de arquitectura es que tiende a presentar un problema denominado “*Vanishing Gradient*” o “Desvanecimiento del Gradiente” (Hochreiter,1998). Este consiste en que, durante la etapa de retro-análisis del proceso de entrenamiento, el cálculo del gradiente para minimizar la función de coste en las capas más lejanas a la capa de salida depende del valor del gradiente en las capas más cercanas (regla de la cadena); por lo que conforme la red se hace más profunda, el valor del gradiente en las últimas capas tiende a cero, lo que produce que los nodos dejen de actualizar sus parámetros. Cuando esto ocurre, el algoritmo ya no está aprendiendo.

En el caso particular de las redes con arquitectura recurrente, este problema se presenta cuando T es muy grande, (Sutskever Ilya,2013); ya que debido al proceso de BPTT, el gradiente en el subnodo $\hat{h}_j^{(1)}$ tiende a cero fácilmente. Esto limita a este tipo de arquitectura a trabajar con ventanas de tiempo muy extensas y a anidar varias de estas capas.

Redes LSTM (Long Short-Term Memory)

La arquitectura LSTM es un tipo especial de red neuronal recurrente (RNR), diseñada para superar la debilidad de las RNR tradicionales para trabajar con ventanas de tiempo extensas (T). Para ello, añade un elemento a cada subnodo denominado “estado de la neurona” o “cell state” (C_t), el cual funciona como una memoria interna que le permite almacenar información para realizar predicciones (Hochreiter, Sepp & Schmidhuber, Jürgen,1997).

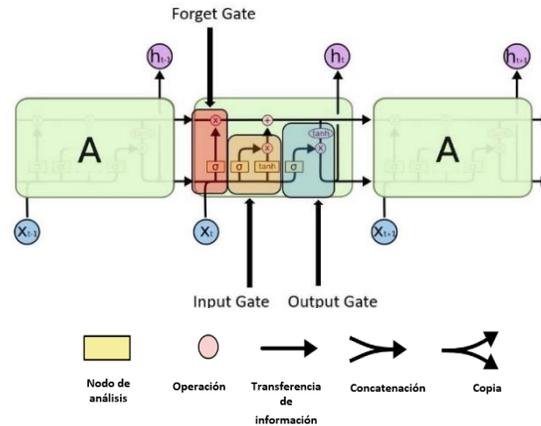


Fig 3.9 Ilustración de Compuertas nodo LSTM (Olah Christopher, 2015)

Este sistema de memoria se basa en un mecanismo de compuertas lógicas en cada subnodo (ver fig 3.9) conformado por: Compuerta de olvido (*Forget gate*), Compuerta de entrada (*Input gate*) y Compuerta de salida (*Output gate*). Estas compuertas controlan el flujo de la información; y operan de manera similar a los nodos conectados dentro de una RNA. La manera específica en la cual se configuran este arreglo permite a este tipo de arquitectura trabajar con ventanas de tiempo extensas; ya que las funciones de activación seleccionadas para cada una de estas compuertas mantienen el valor del gradiente estable durante el proceso de aprendizaje (Hochreiter, Sepp & Schmidhuber, Jürgen,1997.)

A continuación, se describe el funcionamiento de este mecanismo:

En la primera etapa, la red decide qué información almacenada ya no es relevante para futuras predicciones (Kratzert, F,2018). Esta decisión es tomada mediante la compuerta de olvido, descrita matemáticamente por la siguiente ecuación:

$$f_t = \sigma(W_f * \bar{x}^t + U_f \bar{h}^{t-1} + b_f)_{Eq\ 3.21}$$

Ecuación 3.21: Donde f_t es un vector con valores entre (0,1); σ es la función sigmoide (ec.3.3); W_f y U_f son matrices con los coeficientes de peso para esta compuerta, b_f es un coeficiente de sesgo; \bar{x}^t y \bar{h}^{t-1} son los vectores que contienen los datos de entrada y la información del subnodo h^{t-1} .

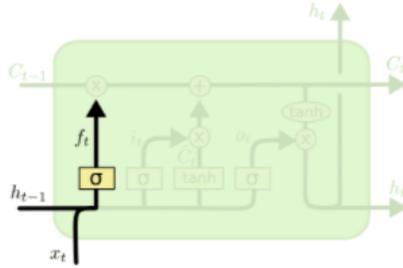


Fig 3.9 Ilustración flujo de información: Compuerta de Olvido (Olah Cristopher, 2015)

Posteriormente, la “compuerta de entrada” decide qué nueva información almacenar en dos etapas: En la primera se calcula un vector potencial de actualización (\tilde{c}_t) con los nuevos datos de entrada mediante la siguiente ecuación:

$$\tilde{c}_t = \tanh(W_{\tilde{c}} * \bar{x}^t + U_{\tilde{c}} \bar{h}^{t-1} + b_{\tilde{c}})_{Ec.3.22}$$

Ecuación 3.22: Donde \tilde{c}_t es un vector con valores entre (-1,1); \tanh es la función tangente hiperbólica (ec.3.4); $W_{\tilde{c}}$ y $U_{\tilde{c}}$ son matrices con los coeficientes de peso para esta compuerta, $b_{\tilde{c}}$ es un coeficiente de sesgo.

En la segunda etapa se calcula el valor de la “compuerta de entrada” (i_t), la cual decide qué información es relevante actualizar (Kratzert, F,2018) mediante la siguiente ecuación:

$$i_t = \sigma(W_i * \bar{x}^t + U_i \bar{h}^{t-1} + b_i)_{Ec.3.23}$$

Ecuación 3.23: Donde i_t es un vector con valores entre (0,1); σ es la función sigmoide (ec.3.3); W_i y U_i son matrices con los coeficientes de peso para esta compuerta, b_i es un coeficiente de sesgo.

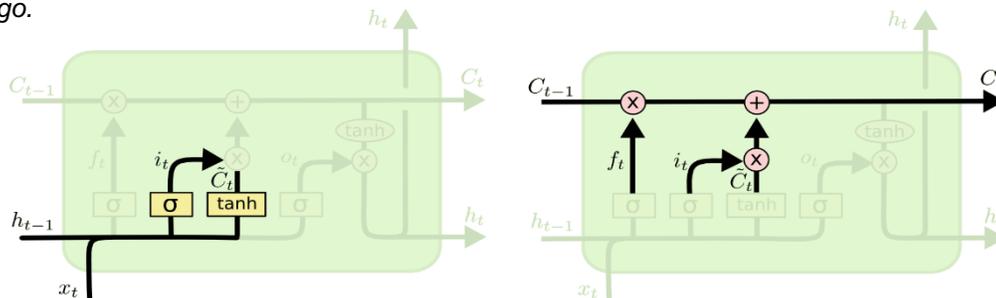


Fig 3.10 Ilustración flujo de información: Compuerta de Entrada (Olah Cristopher, 2015)

Finalmente se actualiza el estado de la “memoria” mediante la siguiente ecuación:

$$c_t = f_t * C_{t-1} + i_t * \tilde{c}_t_{Ec.3.24}$$

Ecuación 3.24: Donde c_t es el estado de la memoria para la secuencia t; f_t es el vector de la compuerta de olvido (Ec.3.21); C_{t-1} es el estado anterior de la neurona; i_t es el vector de la compuerta de entrada (Ec 3.23); y \tilde{c}_t es el vector potencial de actualización (Eq. 3.22).

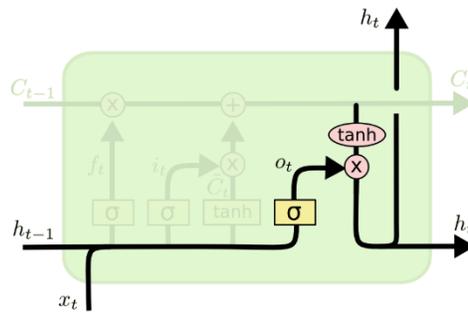


Fig 3.11 Ilustración flujo de información: Compuerta de Entrada (Olah Christopher, 2015)

La compuerta de salida determina qué información del estado actual de la memoria (Kratzert, F,2018) es relevante para determinar el valor de salida del nodo j en la iteración t ; mediante la siguiente manera:

$$o_t = \sigma(W_o * x^t + U_o * \hat{h}_j^{(t-1)} + b_o)_{Ec 3.25}$$

Ecuación 3.25: Donde O_t es el vector de pesos de la compuerta de salida; σ es la función sigmoide (ec.3.3) ; $\hat{h}_j^{(t-1)}$ es el valor de salida del nodo j para el valor de la secuencia anterior ($t-1$) ; W_o y U_o son matrices con los coeficientes de peso para esta compuerta, b_o es un coeficiente de sesgo.

De modo que, el valor de salida del subnodo t está dado por:

$$\hat{h}_j^{(t)} = o_t * \tanh(c_t)_{Ec 3.26}$$

Ecuación 3.26: Donde O_t es el vector de pesos de la compuerta de salida; \tanh es la función de tangente hiperbólica (ec 3.4) ; y c_t es el estado actual de la “memoria” .

Capítulo IV: Metodología

Análisis de la Zona de Estudio

El sistema Cutzamala es un complejo hídrico de almacenamiento, conducción, potabilización y distribución de agua para uso público-urbano e industrial. En promedio, provee de $14.2\text{m}^3/\text{s}$ de agua, a la Ciudad de México ($9.3\text{m}^3/\text{s}$) y al Estado de México ($4.9\text{m}^3/\text{s}$). Lo anterior lo convierte en la segunda fuente de abastecimiento más importante para el Valle de México, al aportar más de una cuarta parte de su abastecimiento total (WBG,2015).

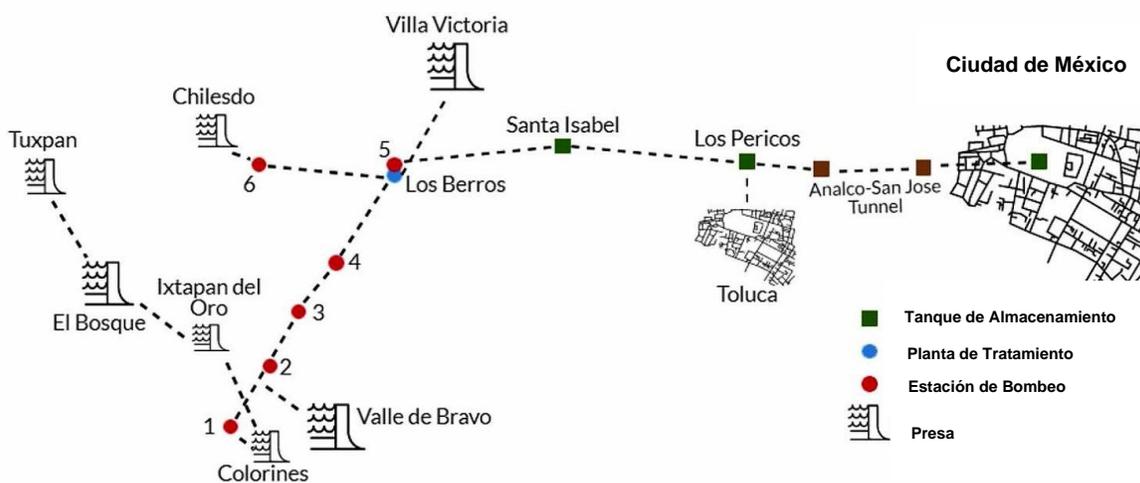


Fig 4.1 Esquema Sistema Cutzamala (WBG,2015).

Construido en tres etapas entre finales de los setenta y principios de los noventa, el sistema está conformado por 3 presas de almacenamiento (El Bosque, Valle de Bravo y Villa Victoria), 4 presas derivadoras (Tuxpan, Ixtapan del Oro, Colorines y Chilesdo), 6 estaciones de bombeo, un acueducto de 127 km y una planta potabilizadora de agua (Los Berros) (WBG,2015).

Su funcionamiento parte en la presa “Tuxpan”, la cual tiene una capacidad de 5hm^3 ; esta canaliza el agua de los ríos Grande y Chiquito hacia la presa “El Bosque” ubicada al sureste a través de una serie de canales abiertos, acueductos y túneles. La presa “El Bosque”, concebida inicialmente para la generación de energía eléctrica, fue añadida durante la tercera etapa de construcción en 1993; tiene una capacidad de 202hm^3 y recibe los escurrimientos del río Zitácuaro y los transferidos de la presa “Tuxpan”, esta a su vez transfiere agua a través de una serie de canales abiertos hacia la presa “Colorines”.

La presa “Ixtapan del Oro”, añadida durante la tercera etapa y con una capacidad de 0.5 Hm³, capta parte de los escurrimientos de la subcuenca “Ixtapan del Oro” y los canaliza hacia el canal Bosque-Colorines. La presa Colorines, la cual tiene una capacidad de 1.5 Hm³, recibe los escurrimientos del río Tilostoc y los transferidos por el canal Bosques-Colorines; esta se encuentra a una altitud de 1,628 m.s.n.m lo que la convierte en la más baja del sistema, por lo que a través de la planta de bombeo 1 (PB1) el agua es bombeada 1100 m verticales hasta la planta potabilizadora “Los Berros”, la cual recibe y trata la totalidad del agua que transfiere el S.C.

El sistema también está integrado por la presa “Valle de Bravo”, con una capacidad de 394.4Hm³, la cual recibe los escurrimientos de los ríos Amanalco, El Molino, Santa Mónica y González; y los transfiere directamente hacia la planta potabilizadora “Los Berros” mediante un sistema de plantas de bombeo. La presa Villa Victoria, con una capacidad de 186 Hm³, la cual almacena los escurrimientos de los arroyos San Diego, La compañía y El Molino, y también envía parte de los escurrimientos mediante el canal Ing. Héctor Martínez de Meza hacia la potabilizadora “Los Berros”. Finalmente la presa “Chilesdo”, la cual capta parte de los escurrimientos del río El Salitre y los transfiere mediante la planta de bombeo 6 (PB6) hacia la planta potabilizadora “Los Berros”.

Durante este trayecto, existen una serie de derivaciones para atender distintos usos consultivos propios de la región; entre los que destaca: tomas ubicadas en el canal Tuxpan - El Bosque para abastecer el módulo 7 de la unidad Hidalgo, perteneciente al Distrito de Riego 045-Tuxpan; así como otras ubicadas dentro del canal El Bosque – Colorines para abastecer las Unidades de Riego “La Florida-Mora” y “Susupuato”.

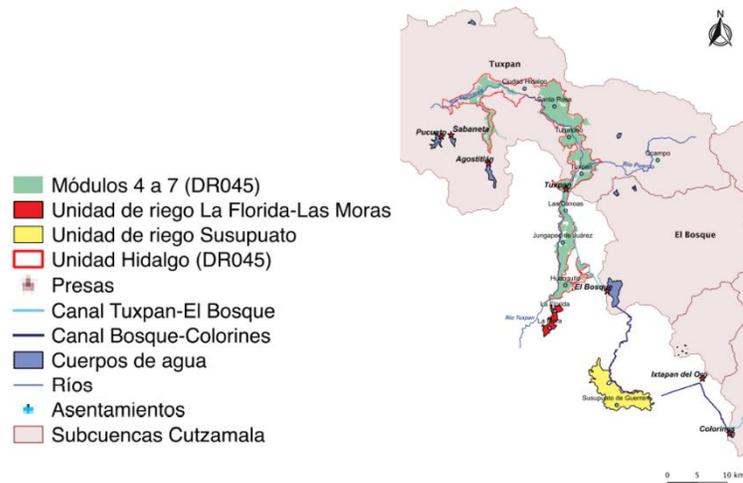


Fig 4.2 Ubicación de las principales extracciones del S.C. (WBG,2015)

Instrumentación en la Zona de Estudio

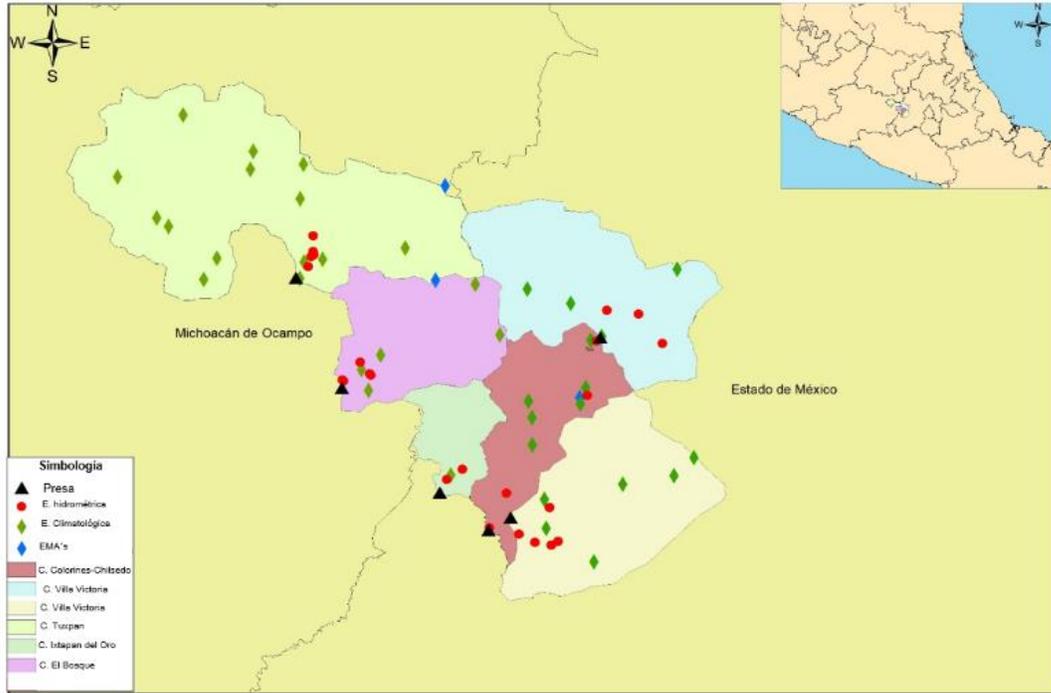


Fig 4.3 Mapa Sistema Cutzamala (Elaboración propia)

En esta región se encuentran ubicadas tres estaciones meteorológicas automáticas (EMA's) (ver tabla 4.1). Estas estaciones registran información referente a velocidad y dirección del viento, lluvia, humedad relativa, radiación solar, temperatura y presión; generando archivos con los promedios de esta información en intervalos de 10 minutos (SMN,2020).

Tabla 4.2 Estaciones Meteorológicas Automáticas (EMA's)

Nombre	Cuenca
Valle de Bravo	Valle de Bravo
M.Monarca I	Tuxpan
M.Monarca II	Chilesdo-Colorines

Asimismo, se tiene registro de 36 estaciones climatológicas tradicionales; de las cuales actualmente sólo 20 se encuentran en operación (SMN,2020); los períodos y longitud de registro de estas estaciones son variables, siendo el registro más antiguo a partir de 1921(ver tabla 4.2). Estas estaciones únicamente registran información de: precipitación acumulada en 24 horas, temperatura del aire (8:00 am), así como la temperatura máxima y mínima del día (SMN,2020).

Tabla 4.1 Estaciones climatológicas en el sitio

ID	Nombre	Periodo de Registro	Estado	Cuenca	ID	Nombre	Periodo de Registro	Estado	Cuenca
16097	Presa Pucuat	1950-1985; 1992-2015	Operando	Tuxpan	15378	San Pablo Malacatepec	1988-1997; 1999-2007; 2012-2015	Operando	Colorines-Chilesdo-
16098	Presa Sabaneta	1952-1968; 1970-2015	Operando	Tuxpan	15391	Villa Victoria	1993-1999; 2001-2015	Operando	Villa Victoria
16235	Huajumbaro	1980-2015	Operando	Tuxpan	16020	Ciudad Hidalgo	1959-1984; 1986-2015	Suspendida	Tuxpan
16258	Presa Tuxpan	1982-1994; 2003-2009; 2011-2015	Operando	Tuxpan	16071	Los Azufres	1964-1970; 1972-1984; 1986-1988;	Suspendida	Tuxpan
16514	Jaripeo	1988-2012 ;2014-2015	Operando	Tuxpan	16079	Melchor Ocampo	1961-1972; 1984-1988	Suspendida	Tuxpan
16002	Agostitlan	1950-1988; 1992-2006	Operando	Tuxpan	16092	Presa Agostitlan	1961-1984	Suspendida	Tuxpan
15197	Cuesta del Carmen	1974-1985; 1995-2006; 2012-2014	Operando	El Bosque	16099	Presa Tuxpan (CFE)	1952-1980; 1991-1994	Suspendida	Tuxpan
16036	Presa EL Bosque	1952-1988 ;20012006; 2009-2015	Operando	El Bosque	16134	Tuxpan	1949-1963; 1966-1977	Suspendida	Tuxpan
16192	La Encarnación	1975-1980; 1987-2005; 2008-2015	Operando	El Bosque	16152	Ciudad Hidalgo (SMN)	1922-198	Suspendida	Tuxpan
15066	Palizada	1958-2000; 2007-2015	Operando	Villa Victoria	16213	Irimbo	1978-1981; 1986-1995	Suspendida	Tuxpan
15133	Presa Villa Victoria	1961-1989; 2001-2006; 2008-2009; 2011-2015	Operando	Villa Victoria	16148	Zitacuaro	1921-1983	Suspendida	El Bosque
15140	Presa Chilesdo	2002-2009; 2011-2015	Operando	Colorines-Chilesdo-	16189	C.R.E.D.F.F.A.	1974-1976	Suspendida	El Bosque
15174	Palo Mancornado	1979-1986; 1992-2002; 2004-2014	Operando	Valle de Bravo	15035	Ixtapan del Oro	1961-1986	Suspendida	Ixtapan del Oro
15165	Valle de Bravo	1967-1969; 1973-1974; 2001-2005; 2011-2015	Operando	Valle de Bravo	15005	Amanalco de Becerra	1965-1982; 1988-1992	Suspendida	Valle de Bravo
15199	Mina Vieja	1974-1990; 1992-2014	Operando	Villa Victoria	15130	Valle de Bravo (CFE)	1969-1989	Suspendida	Valle de Bravo
15265	Campamento Los Berros	1989-1993; 2004-2009; 2011-2015	Operando	Chilesdo-Colorines	15131	Villa de Allende	1961-1985	Suspendida	Chilesdo-Colorines
15368	El Fresno	1989-2015	Operando	Valle de Bravo	15322	Asunción Donato Guerra	1982-1987; 1989-1998	Suspendida	Chilesdo-Colorines
15374	Agua Bendita	1982-1985; 1991-2002; 2008-2015	Operando	Valle de Bravo	15334	Palo Seco	1982-1994	Suspendida	Villa Victoria

Se tiene registro de 27 estaciones hidrométricas dentro de la cuenca y en los canales del sistema; sin embargo, ninguna de estas estaciones se encuentra en operación (CONAGUA,2020). De manera similar a las estaciones climatológicas, los períodos y longitudes de los registros son variables; iniciando en los años cuarenta y en algunos casos abarcando hasta 2010 (ver tabla 4.3). Estas estaciones tienen registro de gastos instantáneos (hidrogramas), del promedio de estos (gastos medios diarios) y en algunos casos mediciones de sedimentos.

Tabla 4.3 Estaciones Hidrométricas en el sitio

ID	Nombre	Periodo de Registro	Estado	Cuenca	ID	Nombre	Periodo de Registro	Estado	Cuenca
18001	SAN JOSE MALACATEPEC	1945-1969; 1971-1977	Suspendida	Chilesdo-Colorines	18489	EL SALTO	1950-1964;	Suspendida	Valle de Bravo
18169	EL MOLINO	1949-1964	Suspendida	Valle de Bravo	18536	EL MOLINO	1953-1958; 1983-1985	Suspendida	Villa Victoria
18274	VERTEDOR GONZALEZ	1957-1964	Suspendida	Valle de Bravo	18545	SALIDAS PRESA VILLA VICTORIA (C.F.E.)	1962-1963; 1966-1967	Suspendida	Villa Victoria
18275	SANTA MONICA	1957-1964	Suspendida	Valle de Bravo	18546	EL MONUMENTO	1968-1975	Suspendida	Chilesdo-Colorines
18279	CARRIZAL	1958-1964; 1978-1985; 1988-1991	Suspendida	Valle de Bravo	18171	EL ORO	1949-1958	Suspendida	El Bosque
18462	LA COMPAÑIA	1953-1964	Suspendida	Villa Victoria	18285	LLEGADA LA BOSQUE (KM 21+740)	1958-1972; 1974-1985	Suspendida	El Bosque
18464	EL RAMAL	1953-1957; 1961-1964	Suspendida	Villa Victoria	18343	LA GARITA	1953-1958	Suspendida	El Bosque
18465	EL DURAZNO	1966-1970	Suspendida	Chilesdo-Colorines	18461	ZITACUARO	1949-1977	Suspendida	El Bosque
18472	LA FLORIDA	1961-1977; 1981-1985	Suspendida	El Bosque	18458	TURUNDEO	1947-1957	Suspendida	Tuxpan
18256	CANAL CONDUCCION	1955-1961; 1963-1970; 1978-1985	Suspendida	Ixtapan del Oro	18459	RIO GRANDE	1947-1981; 1983-1997; 2003-2010	Suspendida	Tuxpan
18466	IXTAPA DEL ORO	1944-1959	Suspendida	Ixtapan del Oro	18460	RIO CHIQUITO	1951-2010	Suspendida	Tuxpan
18213	TUXPAN	1947-1954	Suspendida	Tuxpan	18460	Conducción Tuxpan-Bosque	1961-1964	Suspendida	Tuxpan

Planteamiento del Modelo:

Ante la falta de estaciones hidrométricas en operación dentro de las subcuencas de captación del S.C., se propone el uso de un modelo lluvia-escurrimiento que permita hacer un monitoreo de los ingresos al sistema. De manera que se cuente con información adicional que permita hacer cambios y/o ajustes en las políticas de operación de este, así como detectar comportamientos irregulares en los ingresos relacionados con posibles obstrucciones o aprovechamientos irregulares.

La propuesta de este modelo se basa en el uso de redes neuronales artificiales; ya que se ajusta a las características del sitio de estudio donde no se cuenta con información actualizada para calibrar otro tipo de modelos, pero si con registros con longitud suficientemente para entrenar al modelo. Adicionalmente, se propone utilizar la misma estructura del modelo para realizar el pronóstico de los ingresos futuros al sistema, utilizando una ventana de tiempo de hasta 5 días en el futuro; por lo cual se utilizará una red con arquitectura recurrente en su modalidad LSTM, para poder trabajar con series de tiempo y brindar información adicional al modelo.

Para desarrollar y probar este modelo, se seleccionó como cuenca piloto “El Bosque”; ya que dada la problemática de filtraciones que presenta la cortina de la presa ubicada en este sitio, se consideró que un modelo con las características propuestas podría ayudar a establecer mejores políticas de operación que permitan reducir las pérdidas.

No obstante, cabe mencionar que este modelo fue diseñado de manera en que este pueda ser escalable en el futuro, y pueda ser utilizado en las demás cuencas del S.C; así como cualquier otra cuenca donde se cuente con información de los registros. Por lo cual, se diseñó de manera que la información requerida para alimentar a el modelo fuese de fácil acceso y estuviese disponible para cualquier otro sitio. Por ello, también se utilizó únicamente software y librerías de código abierto, de manera que este pueda ser probado, implementado e inclusive mejorado por cualquier otro actor interesado. (El código de programación se encuentra disponible dentro del anexo 1.)

Selección de Datos y Variables

Dentro de la cuenca “El Bosque”, se seleccionó como piloto la estación hidrométrica Zitácuaro (18461) (CONAGUA,2021); la cual mide los escurrimientos del Río Zitácuaro, principal afluente de la cuenca. Adicionalmente se utilizó información de las estaciones climatológicas “Presa El Bosque” (16036), “Zitácuaro” (16148), “Presa Tuxpan” (16099), “Laguna del Fresno” (16061) y “Juangapeo” (16058) (CICESE, 2020); debido a que eran las estaciones más cercanas (a una distancia no mayor de 80 km de la cuenca) que compartían el mismo periodo de registro con la estación hidrométrica (ver Fig 4.4).

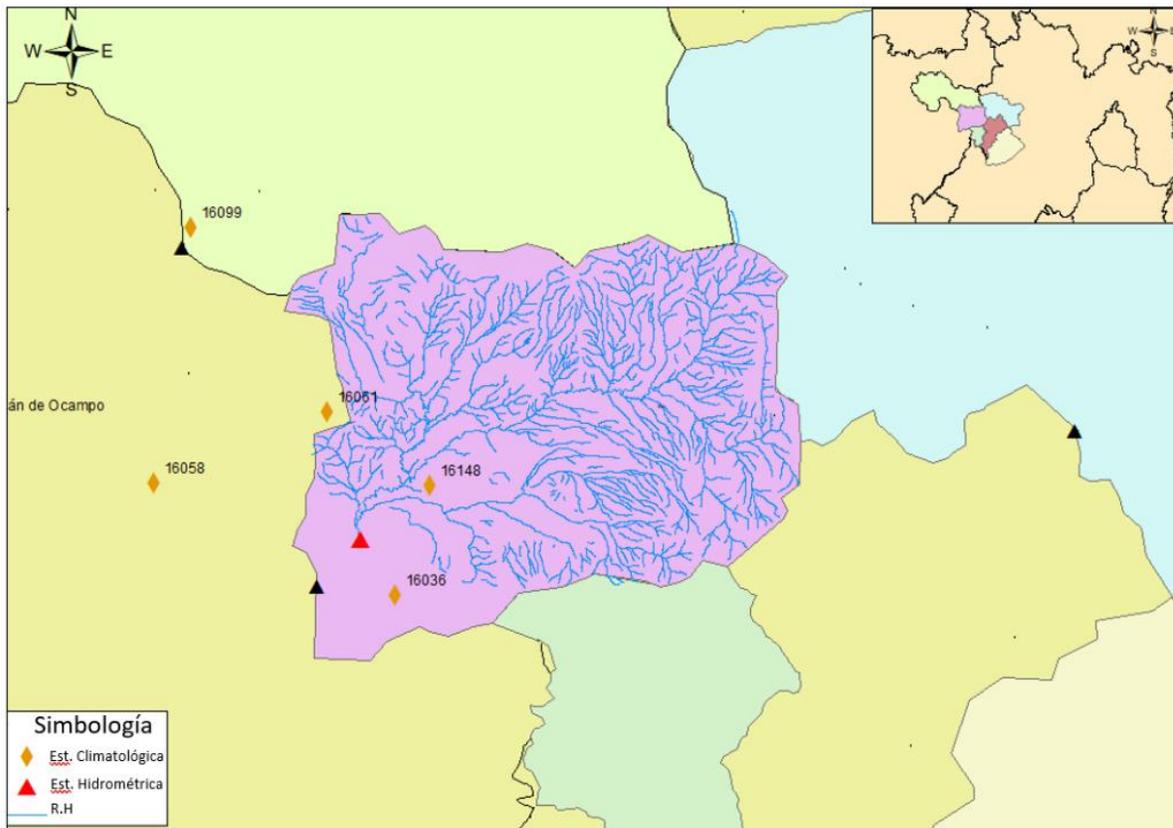


Fig 4.4 Mapa Estaciones de Estudio (Elaboración propia)

La información extraída de estas estaciones fue: en el caso de las hidrométricas, el gasto medio diario para estimar los volúmenes de los escurrimientos que llegan a la presa “El Bosque” como parte de las aportaciones del río Zitácuaro; y por parte de las estaciones climatológicas, únicamente se utilizó los datos registrados de precipitación acumulada en 24 horas. Cabe mencionar que no se aplicó ninguna técnica para estimar la precipitación dentro de la cuenca ya que es algo que el modelo inferirá durante el entrenamiento.

Dada la periodicidad en el comportamiento de los esguerrimientos, se añadió como dato el día del año asociado a la medición de precipitación y esguerrimiento. Para ello, se utilizó la función seno y coseno para describir esta variable; de manera que se pudiese representar la frecuencia sin añadir una jerarquía numérica específica a cada día. Lo anterior, debido a que el día 289 del año (octubre) no es más relevante que el día 170 (junio) o el día 28 (enero) para la predicción del esguerrimiento.

De manera que se estableció como información de entrada para el modelo:

- 1- La medición de la precipitación acumulada en cada una de las estaciones climatológicas antes mencionadas para los días “d-T”*; sin considerar el día en que se realiza el pronóstico.
- 2- El volumen de ingreso registrado en los días “d-T”, calculado a partir de la medición del caudal medio registrado en la estación “Zitacuaro”.
- 3- Día del año asociado a cada uno de los valores antes mencionados.

***Nota:** Se realizaron diversas pruebas para determinar el valor óptimo de la ventana “T” de datos previos (ver “Diseño de la Red”).

La información que se espera que produzca el modelo es:

- 1- El volumen de esguerrimiento para el día “d”; el cual permitirá tener una estimación de los volúmenes que llegan a la presa, dado que esta estación ya no se encuentra en operación.
- 2- La predicción de los volúmenes de esguerrimiento acumulados para un horizonte de 5 días en el futuro (d+4); el cual permitirá prever el comportamiento de los esguerrimientos.

Preprocesamiento de Datos

Análisis de Datos:

Se realizó un análisis estadístico (ver tabla 4.4) de los datos de estudio (escurrimiento medio diario y precipitación acumulada en 24 horas) y se agrupó esta información de manera mensual. Esto con el objetivo de tener visión general del comportamiento de los datos; Así como de detectar posibles datos atípicos asociados a errores de medición, ya que estos pueden afectar a la precisión del modelo.

Se realizó una verificación puntual en aquellos datos máximos que excedían el 4^{to} percentil (ver fig 4.5). No obstante, no se encontró evidencia suficiente en ninguno de los casos de que se tratara de algún error de medición; ya que, para las estaciones climatológicas los datos máximos se encontraban acompañados de un comportamiento similar en las estaciones vecinas, y en el caso de la estación hidrométrica, los datos se encontraban en medio de un periodo de crecida en el nivel del río.

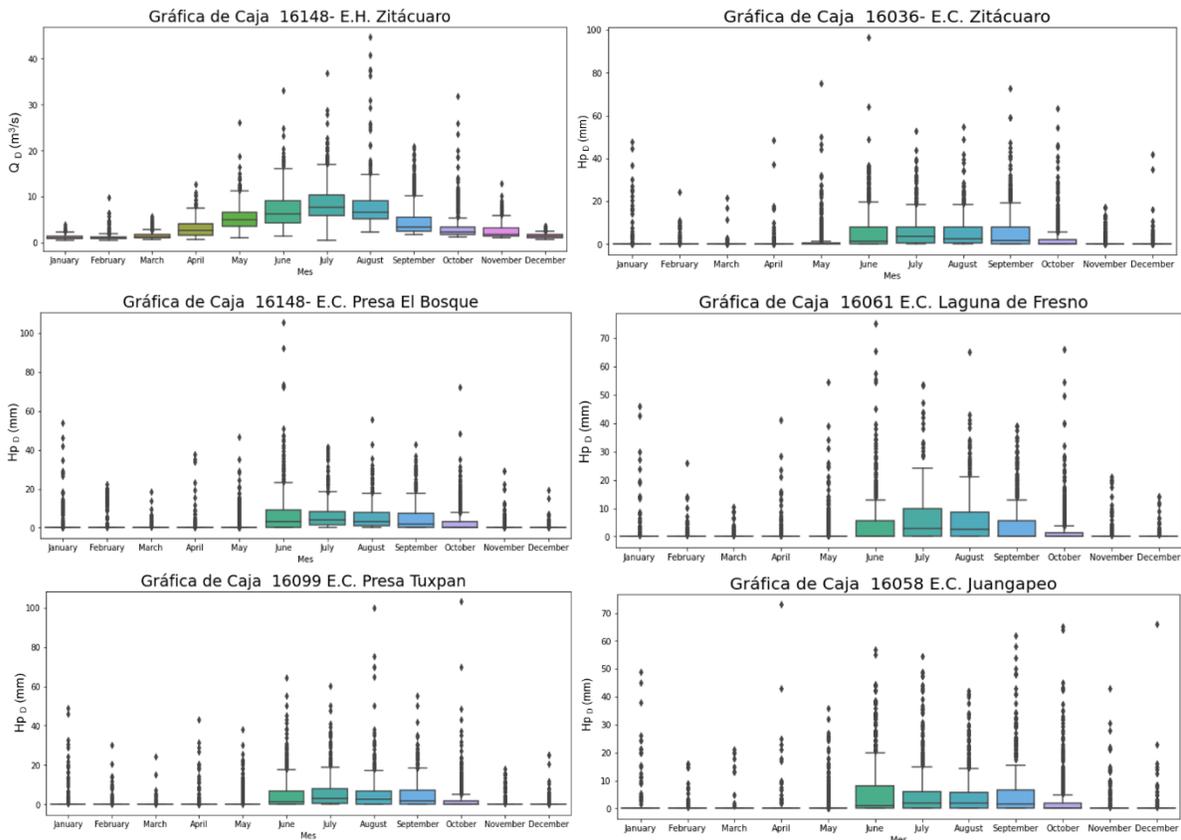


Fig 4.5 Gráficas de Caja de Información por Mes (Elaboración Propia)

Tabla 4.4 Estadísticos

Enero						
Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	1.30 m ³ /s	0.75 mm	1.41 mm	0.61 mm	0.97 mm	0.77 mm
Mediana	1.20 m ³ /s	0.00 mm				
Mínimo	0.50 m ³ /s	0.00 mm				
Máximo	3.90 m ³ /s	47.80 mm	53.70 mm	46.00 mm	49.00 mm	48.80 mm
Des. Est	0.46	4.41	5.31	3.75	4.46	4.37
Coef. Asim	1.32	7.31	5.54	8.45	6.86	7.33
Coef. Curtosis	2.80	59.58	38.33	81.92	55.96	61.57
C.V.	0.35	5.90	3.77	6.12	4.59	5.68
Febrero						
Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	1.25 m ³ /s	0.26 mm	1.82 mm	0.23 mm	0.25 mm	0.36 mm
Mediana	1.14 m ³ /s	0.00 mm				
Mínimo	0.50 m ³ /s	0.00 mm				
Máximo	9.85 m ³ /s	24.40 mm	22.30 mm	26.00 mm	16.00 mm	30.00 mm
Des. Est	0.78	1.65	4.89	1.63	1.57	2.16
Coef. Asim	5.37	9.10	2.50	11.02	8.07	8.77
Coef. Curtosis	41.00	105.27	4.70	146.48	70.49	94.02
C.V.	0.63	6.42	2.68	7.18	6.34	6.00
Marzo						
Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	1.60 m ³ /s	0.12 mm	0.18 mm	0.14 mm	0.23 mm	0.15 mm
Mediana	1.29 m ³ /s	0.00 mm				
Mínimo	0.75 m ³ /s	0.00 mm				
Máximo	5.83 m ³ /s	21.70 mm	18.70 mm	10.50 mm	21.00 mm	24.30 mm
Des. Est	0.83	1.29	1.26 mm	0.91	1.96	1.33
Coef. Asim	1.81	13.91	9.89	8.18	9.02	13.80
Coef. Curtosis	3.29	206.80	117.67	73.41	82.99	222.37
C.V.	0.52	10.38	6.89	6.56	8.46	8.81

Abril

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	3.10 m ³ /s	0.39 mm	0.48 mm	0.66 mm	0.53 mm	0.65 mm
Mediana	2.70 m ³ /s	0.00 mm				
Mínimo	0.82 m ³ /s	0.00 mm				
Máximo	12.80 m ³ /s	48.50 mm	37.50 mm	41.00 mm	73.00 mm	43.00 mm
Des. Est	1.85	3.01	3.14	3.16	4.24	3.56
Coef. Asim	1.28	12.12	9.18	7.58	12.30	7.52
Coef. Curtosis	2.39	168.32	92.77	71.89	181.91	65.63
C.V.	0.60	7.72	6.61	4.75	8.05	5.43

Mayo

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	5.34 m ³ /s	2.07 mm	1.45 mm	1.65 mm	1.39 mm	1.46 mm
Mediana	5.00 m ³ /s	0.00 mm				
Mínimo	1.04 m ³ /s	0.00 mm				
Máximo	26.11 m ³ /s	75.00 mm	46.50 mm	54.50 mm	36.00 mm	38.00 mm
Des. Est	2.86	6.43	4.47	4.92	4.25	4.12
Coef. Asim	1.44	5.74	4.96	5.32	4.51	4.24
Coef. Curtosis	5.82	44.77	32.22	38.22	23.74	22.84
C.V.	0.54	3.11	3.08	2.98	3.06	2.82

Junio

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	7.28 m ³ /s	5.92 mm	7.07 mm	4.71 mm	6.16 mm	5.43 mm
Mediana	6.33 m ³ /s	1.50 mm	3.00 mm	0.00 mm	1.00 mm	1.55 mm
Mínimo	1.45 m ³ /s	0.00 mm				
Máximo	33.15 m ³ /s	96.40 mm	105.30 mm	75.00 mm	57.00 mm	64.20 mm
Des. Est	4.09	9.62	11.70	9.21	9.93	8.86
Coef. Asim	1.42	3.22	3.68	3.45	2.16	2.59
Coef. Curtosis	3.62	18.00	19.85	15.88	4.74	8.62
C.V.	0.56	1.63	1.65	1.95	1.61	1.63

Julio

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	8.61 m ³ /s	5.96 mm	6.04 mm	6.09 mm	5.66 mm	6.00 mm
Mediana	7.80 m ³ /s	3.60 mm	4.20 mm	3.00 mm	2.00 mm	3.30 mm
Mínimo	0.60 m ³ /s	0.00 mm				
Máximo	36.79 m ³ /s	53.00 mm	41.50 mm	53.50 mm	54.50 mm	60.10 mm
Des. Est	4.30	7.47	6.66	8.26	9.15	8.10
Coef. Asim	1.63	2.23	2.20	2.22	2.55	2.54
Coef. Curtosis	5.17	6.46	6.72	6.95	7.18	8.62
C.V.	0.50	1.25	1.10	1.36	1.62	1.35

Agosto

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	7.96 m ³ /s	5.33 mm	5.33 mm	5.92 mm	4.50 mm	5.66 mm
Mediana	6.74 m ³ /s	2.50 mm	3.20 mm	2.50 mm	1.80 mm	2.50 mm
Mínimo	2.45 m ³ /s	0.00 mm				
Máximo	44.66 m ³ /s	54.80 mm	55.50 mm	65.00 mm	42.00 mm	100.00 mm
Des. Est	4.92	7.15	6.48	8.19	7.10	9.76
Coef. Asim	3.59	2.56	2.58	2.24	2.58	4.47
Coef. Curtosis	18.58	9.62	10.74	7.30	7.66	29.19
C.V.	0.62	1.34	1.22	1.38	1.58	1.72

Septiembre

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	4.93 m ³ /s	5.74 mm	5.09 mm	4.14 mm	5.34 mm	5.06 mm
Mediana	3.50 m ³ /s	1.80 mm	2.00 mm	0.00 mm	1.50 mm	2.00 mm
Mínimo	1.85 m ³ /s	0.00 mm				
Máximo	20.88 m ³ /s	72.70 mm	43.00 mm	39.00 mm	62.00 mm	55.20 mm
Des. Est	3.64	9.15	7.24	7.15	9.38	7.56
Coef. Asim	1.89	2.90	2.06	2.33	2.86	2.41
Coef. Curtosis	3.19	11.69	4.59	5.80	9.63	7.96
C.V.	0.74	1.59	1.42	1.73	1.76	1.49

Octubre

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	3.50 m ³ /s	2.88 mm	2.92 mm	2.47 mm	3.13 mm	2.93 mm
Mediana	2.33 m ³ /s	0.00 mm				
Mínimo	1.25 m ³ /s	0.00 mm				
Máximo	31.80 m ³ /s	63.20 mm	72.30 mm	66.00 mm	65.00 mm	103.00 mm
Des. Est	3.21	7.08	6.49	6.68	7.93	7.89
Coef. Asim	3.82	4.25	4.43	4.79	3.95	6.26
Coef. Curtosis	20.64	23.09	30.78	30.33	19.36	59.53
C.V.	0.92	2.46	2.22	2.71	2.54	2.70

Noviembre

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	2.63 m ³ /s	0.51 mm	0.65 mm	0.47 mm	0.74 mm	0.55 mm
Mediana	1.95 m ³ /s	0.00 mm				
Mínimo	1.04 m ³ /s	0.00 mm				
Máximo	12.90 m ³ /s	26.50 mm	29.20 mm	23.00 mm	43.00 mm	18.00 mm
Des. Est	1.63	2.27	2.89	2.42	3.57	2.32
Coef. Asim	1.95	6.30	6.62	6.71	6.81	4.94
Coef. Curtosis	5.07	48.57	51.34	48.96	56.66	25.39
C.V.	0.62	4.44	4.43	5.10	4.79	4.21

Diciembre

Estadístico	EH 18461	EC 16036	EC 16148	EC 16061	EC 16058	EC 16099
Media	1.57 m ³ /s	0.34 mm	0.24 mm	0.22 mm	0.39 mm	0.38 mm
Mediana	1.48 m ³ /s	0.00 mm				
Mínimo	0.72 m ³ /s	0.00 mm				
Máximo	3.70 m ³ /s	42.00 mm	19.50 mm	14.20 mm	66.00 mm	25.00 mm
Des. Est	0.56	2.52	1.36	1.28	3.21	2.03
Coef. Asim	1.10	12.83	8.76	8.06	15.77	8.78
Coef. Curtosis	1.18	189.38	97.71	72.97	302.26	92.29
C.V.	0.35	7.52	5.73	5.77	8.29	5.42

Integración de Conjuntos de Datos

Con el objetivo de poder validar el modelo se dividió al registro de datos en dos subconjuntos : El primero, denominado “Conjunto de Entrenamiento”, comprende a los registros de 1952 a 1967 (5805 datos) y fue utilizado para el proceso de entrenamiento de la red en la medición del subajuste; el segundo conjunto, denominado “Conjunto de Validación”, comprende al periodo 1968-1970 (830 datos) y fue utilizado durante el proceso de validación de resultados para probar el desempeño del modelo y corroborar que no presentará sobreajuste. A continuación, se muestran los datos que servirán como *inputs* para el modelo:

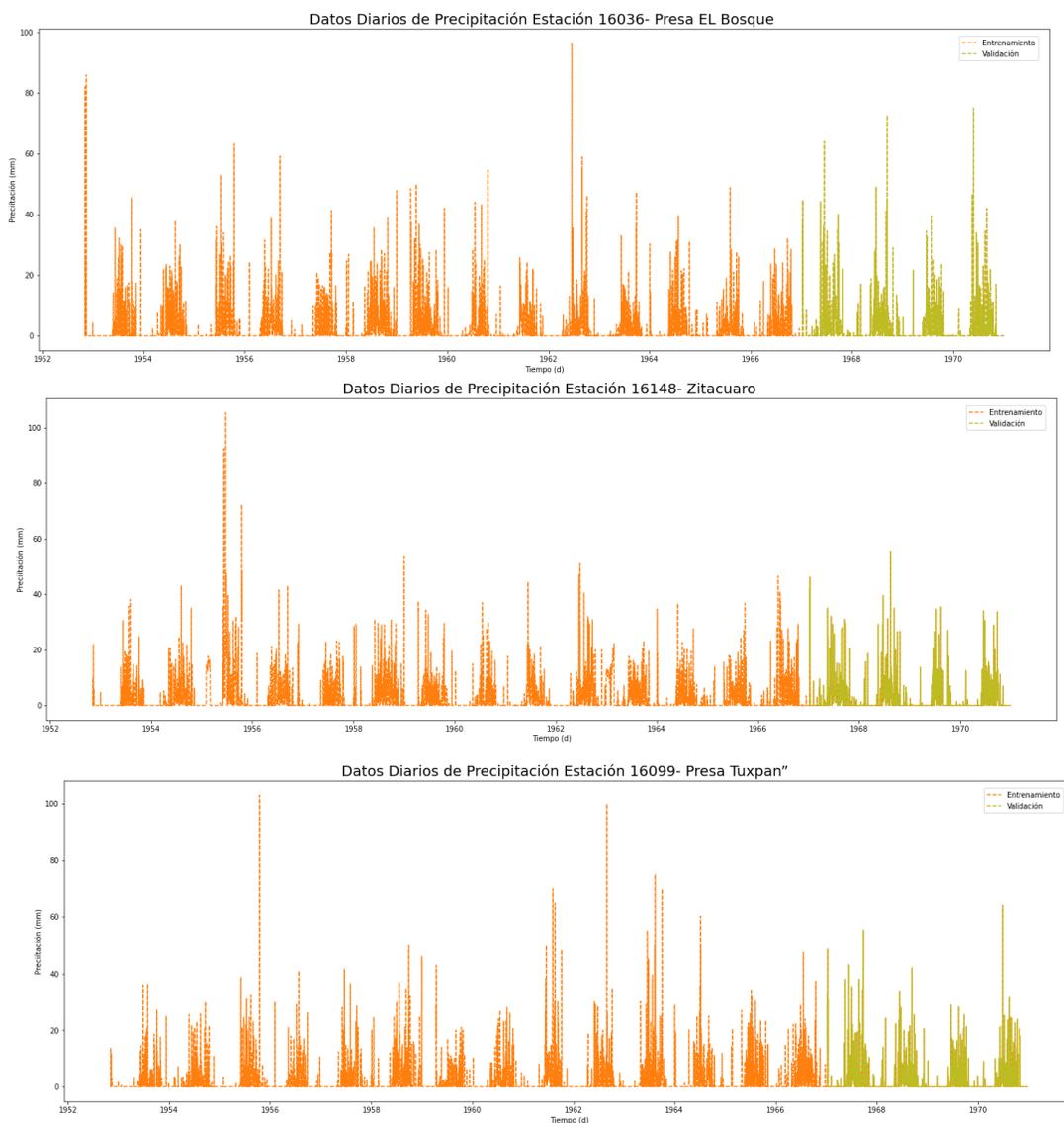


Fig 4.6 Inputs Modelo (Elaboración Propia)

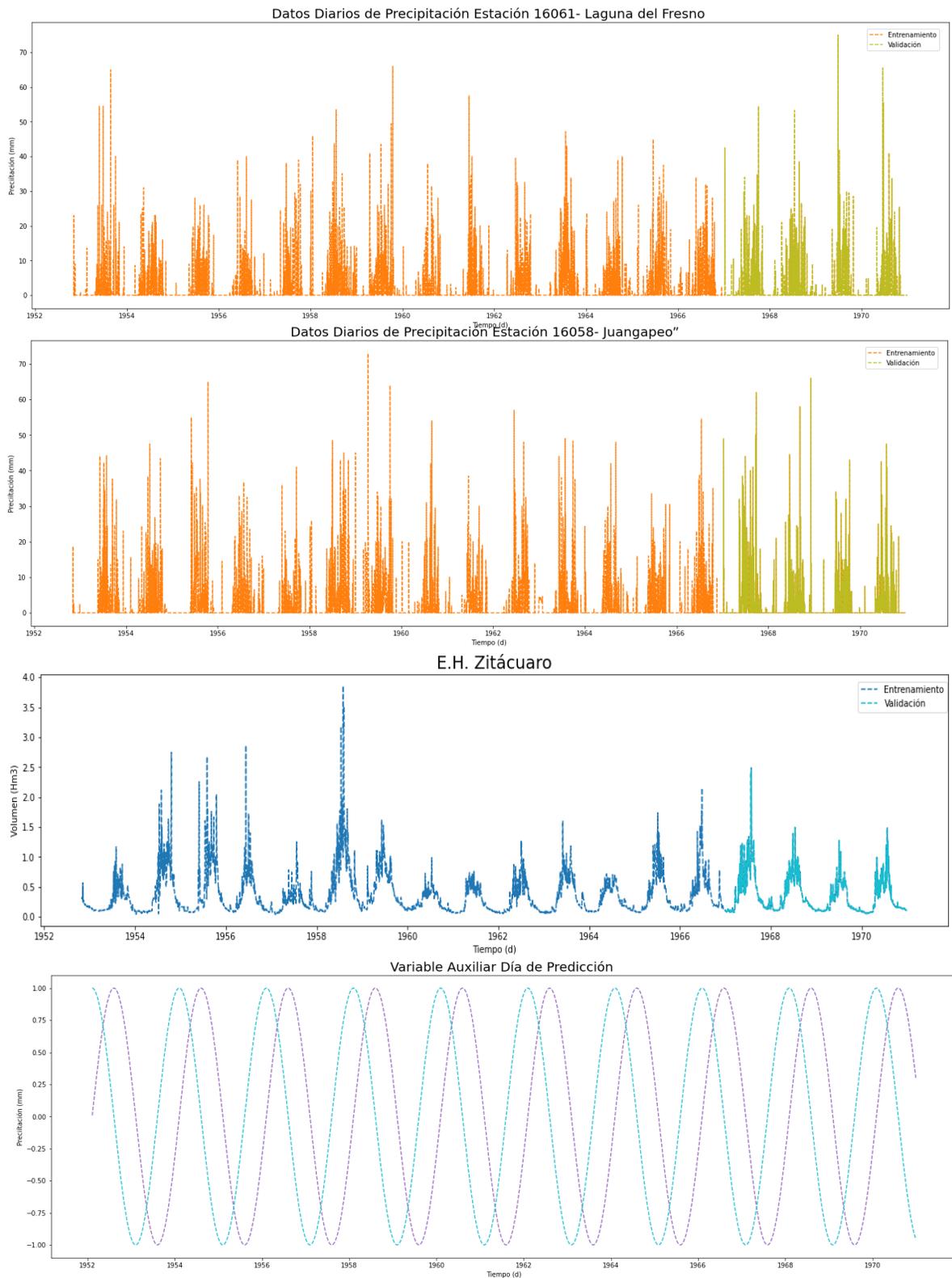


Fig 4.7 Inputs Modelo (Elaboración Propia)

Con respecto a los datos de salida para entrenar al modelo, se conformaron las siguientes series con información de: El volumen de esguimiento medido en el día “d”; el volumen de esguimiento acumulado para los siguientes dos días (d+1), para los siguientes 3 (d+2), y así sucesivamente hasta los siguientes 5 días (d+4). Estas series, se dividieron de la misma manera que los datos de entrada.

A continuación se presenta únicamente la serie de datos de validación, con el objetivo de mostrar con mayor detalle el comportamiento en el ruido de estos datos; ya que, por ejemplo, si se compara la serie de esguimiento en el día d con respecto a “d+4”, se puede observar que esta última tiene una menor cantidad de “picos” y presenta un comportamiento considerablemente más estable, y por tanto predecible.

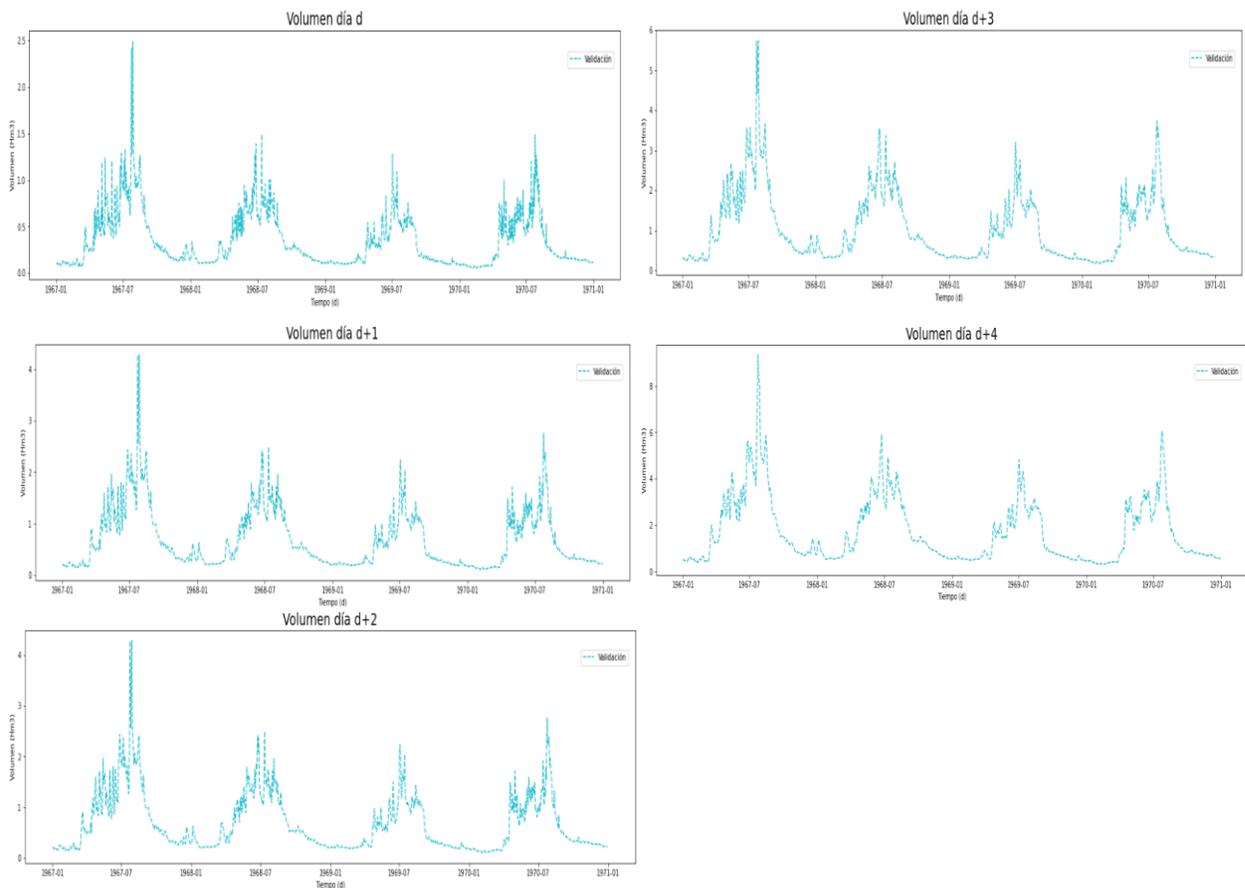


Fig 4.8 Reducción del ruido: outputs del modelo (Elaboración Propia)

Escalado de Atributos

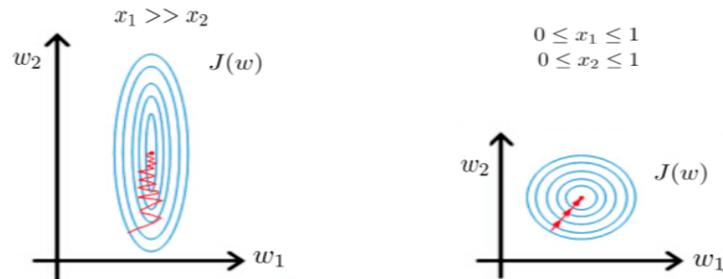


Fig 4.9 Comparación de la trayectoria de optimización (Ian Goodfellow,2017)

Es importante que todos los atributos o variables con los que se alimente al modelo se encuentren dentro de un mismo orden de magnitud; ya que de lo contrario se generan distorsiones al momento de optimizar los parámetros de la red durante el proceso de entrenamiento, afectando la sensibilidad del modelo y haciéndolo dependiente de los órdenes de magnitud de las variables (Ian Goodfellow,2017).

Adicionalmente es recomendable que el orden de los atributos se encuentre cerca del rango de -1 y 1, ya que esto permite al algoritmo converger de manera más eficiente al simplificar los cálculos y reducir los tiempos de cómputo para el entrenamiento de la red (Ian Goodfellow,2017) (ver fig 4.9).

Es necesario escalar de manera conjunta los datos con los cuales trabajará el modelo (tanto el conjunto de entrenamiento como el de validación). Existen diversos métodos para realizar este escalado de los atributos; siendo el más utilizado dentro de la bibliografía consultada la estandarización (Eq.4.1), el cual elimina la media y escala los datos con respecto a la desviación estándar, de manera que se centran en 0 con una desviación estándar cercana a uno.

$$Z_i = \frac{x_i - \bar{x}_i}{\sigma} \text{ Eq.(4.1)}$$

Ecuación 4.1. Donde Z_i la variable escalada, x_i es la variable original, \bar{x}_i es la media del conjunto de entrenamiento original y σ la desviación estándar.

No obstante, esta transformación solo es recomendable cuando los datos tienen una distribución normal y es sensible a valores extremos, por lo que no fue utilizada para este modelo ya que los datos de precipitación diaria y caudal medio diario no siguen esta distribución (ver tabla 4.4).

Se encontró que, una técnica más adecuada para el tipo de datos de este estudio es el “estimador robusto” o “RobustScaler” (Eq 4.2). Este remueve la media y escala los datos de acuerdo con el rango entre el primer cuartil y el tercer cuartil del conjunto de entrenamiento, este tiene la ventaja de ser menos susceptible a valores extremos (por lo que ajustó bien para los datos de precipitación) y que no está condicionado a que los datos pertenezcan a una determinada distribución, por lo que fue el estimador que se utilizó para este modelo.

$$Z_i = \frac{x_i - x_{md}}{x_{0.75} - x_{0.25}} \text{ Eq.(4.2)}$$

Ecuación 4.2. Donde Z_i la variable escalada, x_i es la variable original, x_{md} es la media del conjunto de entrenamiento original, $x_{0.75}$ y $x_{0.25}$ son los cuantiles 3 y 1 respectivamente

Cabe mencionar que existen otras técnicas de escalado consultadas en la bibliografía. Sin embargo, son más susceptibles a los datos extremos o tienen condicionantes que, para el caso particular de este modelo, hacían que su desempeño fuera menor a la transformación elegida. No obstante, a continuación de describen brevemente:

MinMaxScaler: es una transformación que escala las características a valores entre 0 y 1, esta transformación responde bien si la desviación estándar es pequeña y cuando una distribución no es gaussiana; no obstante, es sensible a valores extremos.

$$Z_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \text{ Eq.(3.3)}$$

Ecuación 3.3. Donde Z_i la variable escalada, x_i es la variable original, x_{min} es el valor mínimo de del conjunto de datos x ; x_{max} es el valor máximo del conjunto de datos x .

MaxAbsScaler escala cada atributo con respecto al valor absoluto máximo de la serie, arrojando valores que únicamente se encuentran entre 0 y 1, este estimador también es sensible a valores extremos.

$$Z_i = \left\| \frac{x_i}{x_{max}} \right\|$$

Ecuación 3.4. Donde Z_i la variable escalada, x_i es la variable original, x_{max} es el valor máximo del conjunto de datos x .

Diseño de la Red

El modelo de red neuronal se programó en lenguaje de Python utilizando las librerías Keras (Chollet, F. et al, 2015) y TensorFlow (Martín Abadi et al, 2015); las cuales son de código abierto. Estas librerías contienen funciones y procedimientos preprogramados llamados API's (*Application Programming Interface*) enfocados al diseño de modelos de aprendizaje automatizado. Estos API's se dividen dos tipos de modelos: el primero es para la creación de "modelos secuenciales", los cuales se generan creando una secuencia simple de capas de nodos, que se van apilando y donde solo existe un "tensor" o matriz de datos de entrada y salida. El segundo tipo de modelos es llamado "funcional" y permite la creación de modelos más complejos; los cuales se pueden anidar entre sí, de manera que parte de las salidas de una de las capas sirva como entrada para un segundo modelo. (Martín Abadi et al, 2015). Este modelo se programó utilizando las paqueterías enfocadas hacia el diseño de redes con arquitectura secuencial.

Adicionalmente se utilizó como librerías auxiliares para el manejo y análisis de datos las librerías Pandas (Reback J et al 2020); Numpy (Charles R, et al, 2020) y Matplotlib (Hunter, J. D, 2007).

Como se mencionó dentro del capítulo III, el proceso de entrenamiento de una red requiere del cómputo de una serie de operaciones complejas de manera iterativa. En una computadora de escritorio convencional, estos cálculos se llevan a cabo dentro del CPU (*Central Processing Unit*) de manera secuencial (uno a uno); debido a la cantidad y tipo de cálculos requeridos, este proceso puede llegar a demorar varias horas e inclusive días (aún en un CPU de alta gama), lo cual solía limitar la aplicación de este tipo de técnicas. Algunas computadoras adicionalmente cuentan con un GPU (*Graphic Processing Unit*), este tipo de procesador permite realizar varios cálculos de manera paralela; y en algunos casos, cuentan con módulos especializados enfocados para trabajar con R.N.A. Sin embargo, este tipo de procesadores especializados suelen tener un costo elevado.

Debido a esto, en este trabajo se recurrió al uso de Google Colaboratory (Google, 2021). Esta es una herramienta, basada en *Cloud Computing*, que permite tener acceso al uso de GPUs y TPUs (*Tensor Processing Unit*, similar al GPU pero enfocado únicamente para actividades relacionadas con el aprendizaje automatizado) de sus servidores (Kaz Sato, 2021). Esta herramienta tiene una versión gratuita, la cual fue suficiente para los requerimientos de este modelo.

Una vez que se ha definido el tipo de modelo y el entorno de programación se establece el número y tipo de capas, así como el número de neuronas para cada capa. Para este modelo se propuso la implementación de capas tipo Long Short Term Memory (LSTM), debido a las características mencionadas anteriormente.

Se propuso una arquitectura de dos capas intermedias del tipo LSTM, con 25 y 20 nodos respectivamente; donde los nodos de la primera capa estuvieran programados de manera “Secuencia-Secuencia”, para poder aplicar otra capa recurrente y aprovechar su capacidad de análisis; y la segunda capa LSTM mediante una configuración “Secuencia-a uno”, de manera que se obtuviese un único resultado.

A estas capas se les aplicó la técnica de “Dropout” la cual inactiva un porcentaje de los nodos de manera aleatoria en cada iteración del proceso de entrenamiento; esto con el objetivo de evitar que el modelo sobreajuste o que existan nodos que no estén trabajando, este se estableció en 20 % de inactivación de los nodos. Finalmente se añadió una capa que fungiese como capa de salida, con un único nodo, y cuya función de activación fuese la función identidad, de manera que los resultados no estuviesen limitados a un rango específico.

Cabe mencionar que la arquitectura propuesta está basada en la utilizada por Kratzert, F et al (2018), quienes desarrollaron una red compuesta por dos capas LSTM de 20 nodos cada una para la predicción de caudales en diferentes cuencas ubicadas en Estados Unidos; no obstante, realizando pruebas se encontró un mejor desempeño de la red añadiendo nodos a la primera capa. En este sentido es conveniente mencionar que el número de nodos y capas incrementa considerablemente el tiempo de computó durante el proceso de entrenamiento.

Finalmente se estableció la ventana de datos (T) que recibirá el modelo en cada iteración para realizar el pronóstico; dentro del trabajo de Kratzert, F et al (2018) este parámetro se definió como 365 días, no obstante se encontró que para este modelo ventanas tan amplias daban un comportamiento demasiado robusto que impedía que se ajustase a los valores picos. Realizando pruebas se encontró que el valor óptimo de T para este conjunto específico de datos era de 10 días.

El resultado final fue un modelo con 7,101 parámetros para ajustar durante el proceso de entrenamiento.

```

Model: "sequential"
Layer (type)      Output Shape      Param #
-----
lstm (LSTM)       (None, 5, 25)    3400
lstm_1 (LSTM)     (None, 20)       3680
dense (Dense)     (None, 1)        21
-----
Total params: 7,101
Trainable params: 7,101
Non-trainable params: 0
    
```

Fig 4.10 Impresión del Resumen del Modelo

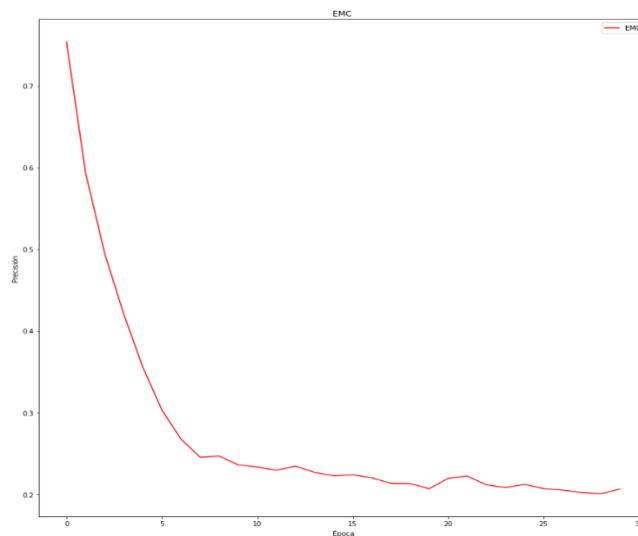
Entrenamiento

El entrenamiento de la red se realizó mediante la técnica de descenso del gradiente por mini-lotes; se propuso un tamaño de 365 datos para cada uno de estos mini-lotes (*Batches*), de manera que se abarcase un ciclo completo del comportamiento periódico de los datos. Generando así 25 mini-lotes o lotes de entrenamiento.

Posteriormente se seleccionó como función de coste el error medio cuadrático (EMC) (Ec.3.8); ya que, dentro de la bibliografía consultada, es la función más común aplicada a este tipo de modelos. Finalmente, para ajustar la tasa de aprendizaje durante el entrenamiento se utilizó el algoritmo “Adam”; utilizando como sus hiper-parámetros del algoritmo:

$$\alpha = 0.001, \beta_1=0.9, \beta_2 = 0.999 \text{ y } \varepsilon = 10^{-8}.$$

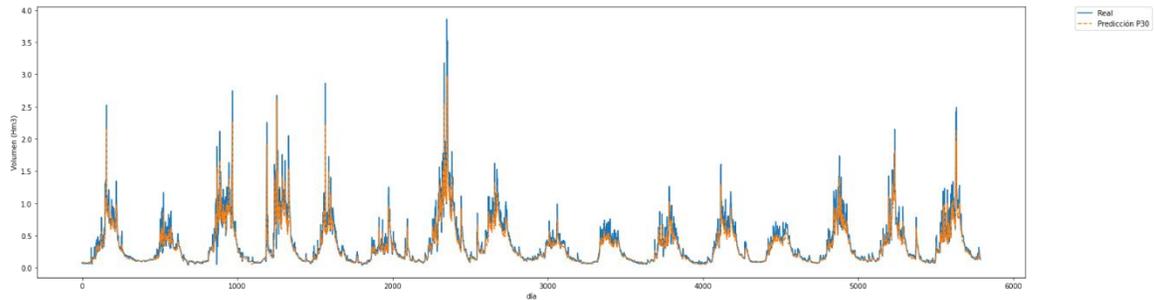
El proceso de entrenamiento se repitió por 30 “épocas”; es decir, los 25 lotes en los cuales se dividió el conjunto de entrenamiento fueron procesados 30 veces para ajustar los parámetros de la red. En la figura 4.9 se muestra la evolución de la función de coste con respecto a la “época” (*Epoch*) de entrenamiento, se puede apreciar cómo esta deja de variar significativamente a partir de la época 25; por lo que no se realizan más ciclos para prevenir que la red presente sobreajuste.



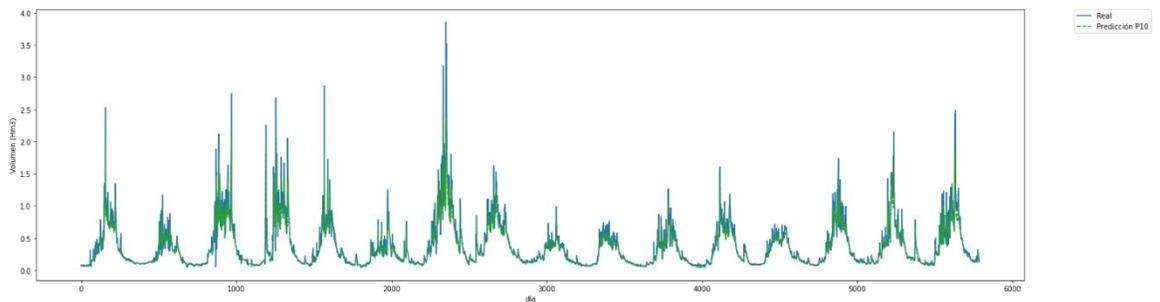
Fia 4.11 Evolución de EMC

En la figura 4.12 se muestra el proceso de evolución en el comportamiento del modelo para la predicción de un día en el futuro para el conjunto de entrenamiento; la manera en la cual se ajusta el modelo conforme se incrementan las épocas de entrenamiento es idéntica para los otros conjuntos de datos.

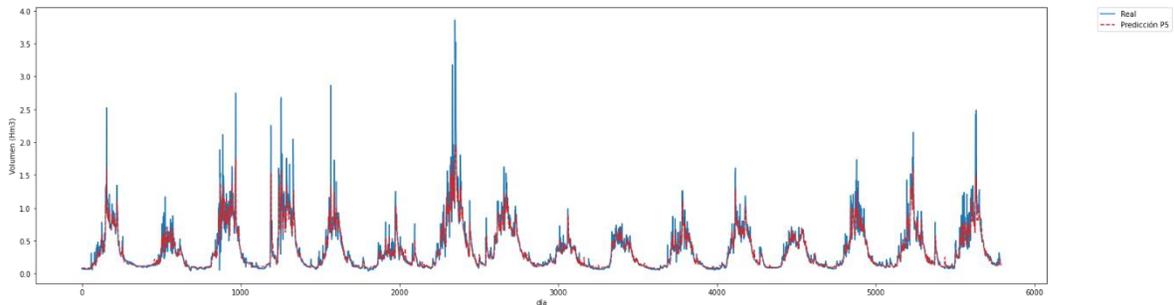
Época 30



Época 10



Época 5



Época 1

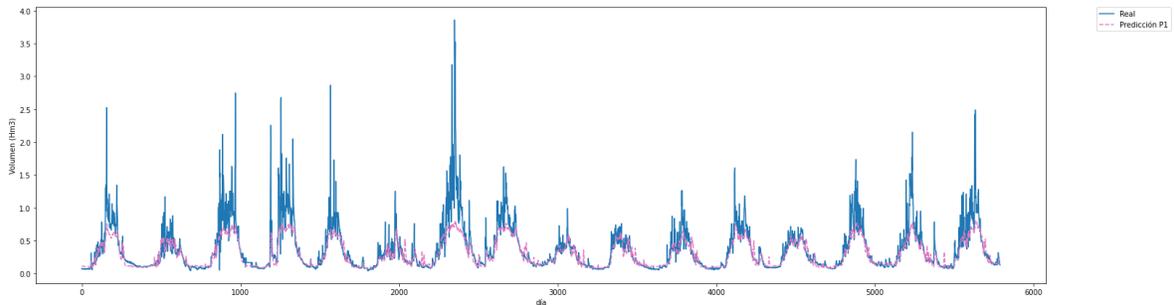


Fig 4.12 Evolución del Ajuste del Modelo para predicción de un día

Validación del Modelo

Al finalizar el entrenamiento se probó al modelo utilizando el conjunto de entrenamiento. Cabe recordar que este conjunto de datos no había sido procesado anteriormente por el modelo, por lo que es información nueva para la cual el modelo no se ha ajustado específicamente. También se volvió a procesar el conjunto de datos de entrenamiento. Finalmente se re escalaron los datos a su dimensión original para su análisis. Se eligió indicador del ajuste el error medio cuadrático ya que este fue el indicador utilizado durante el entrenamiento para ajustar los parámetros de la red.

Tabla 4.5 Comparación MSE (Datos Escala Original - Hm³)

MSE/Días	1	2	3	4	5
Entrenamiento	0.0279	0.0983	0.2121	0.3852	0.5653
Validación	0.0171	0.0554	0.1833	0.3681	0.5338

En general, en ninguno de los cinco casos se encontró evidencia de que el modelo estuviese sobre ajustando; ya que en todo el error asociado al conjunto de validación fue menor al del conjunto de entrenamiento. Lo cual se atribuye a que los datos del conjunto de entrenamiento tienen un menor ruido. Esto también se puede apreciar en la evolución de la relación entre el error del conjunto de entrenamiento con respecto al de validación, ya que en términos relativos este disminuye considerablemente con respecto al horizonte de pronóstico; dado que, como se mostró en la figura 4.7, el conjunto de datos para una ventana de 5 días tiene asociado un menor ruido que el de solo uno.

Finalmente se señala que, aun que el error incrementa considerablemente conforme la ventana de tiempo de predicción se amplía, también lo hace el orden de magnitud de la predicción, dado que se está pronosticando el volumen acumulado de escurrimiento; por lo que este incremento no necesariamente está directamente relacionado con un decremento en el desempeño del modelo.

A continuación se muestran los datos graficados de las estimaciones y los valores reales para ambos conjuntos de datos.

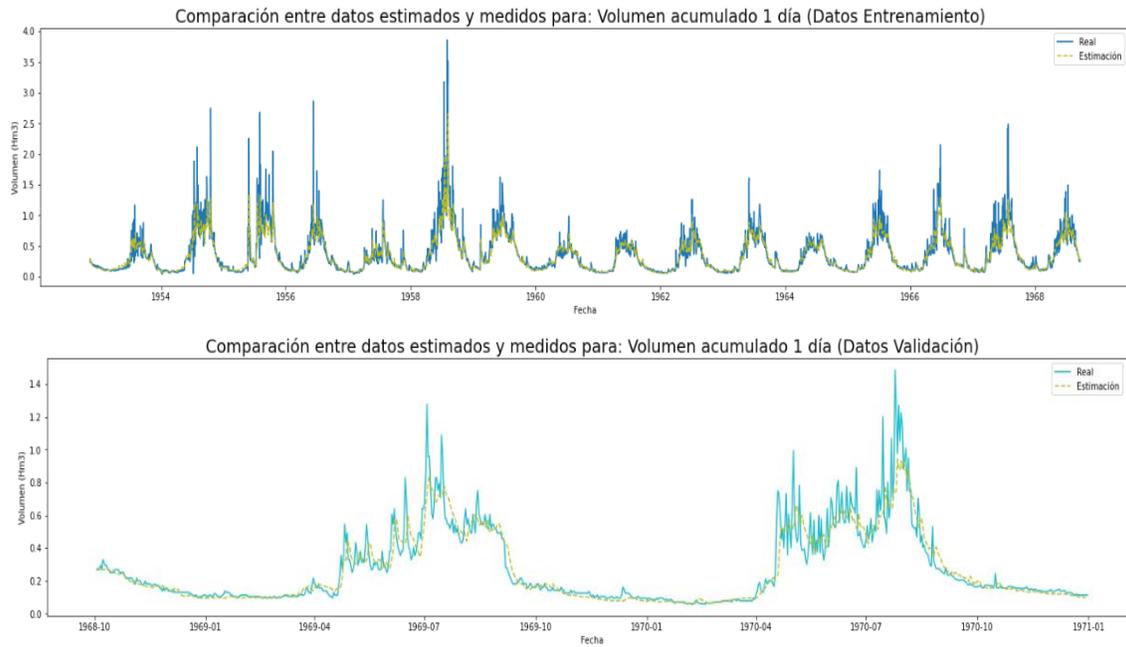


Fig 4.13 Valores Estimados vs Reales: Ventana 1 día

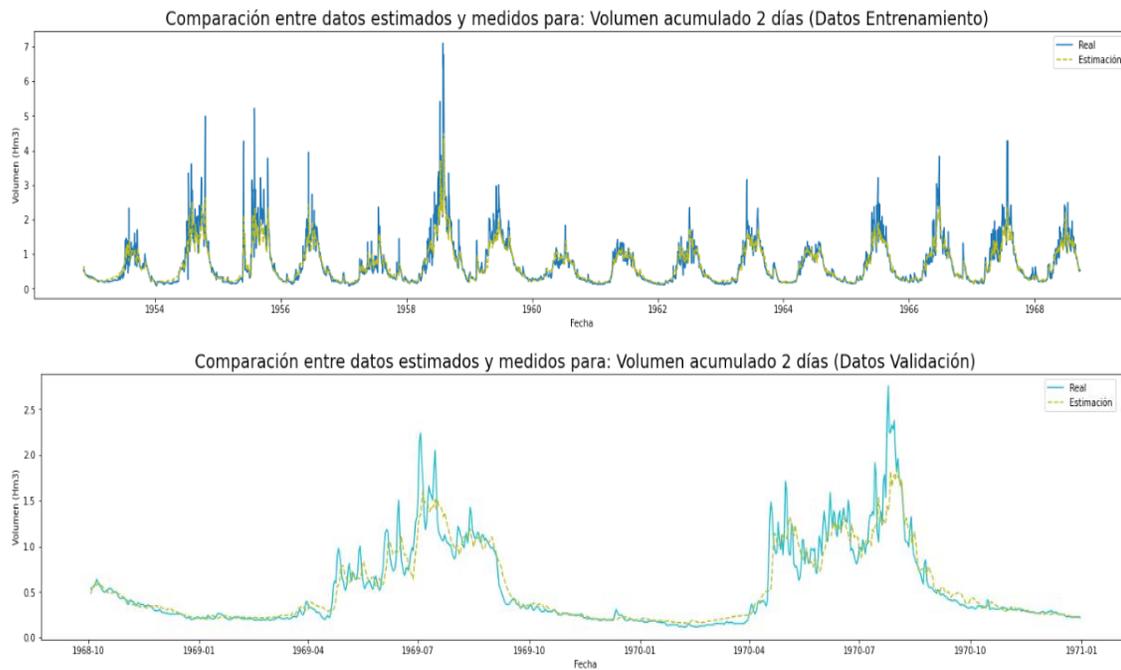


Fig 4.14 Valores Estimados vs Reales: Ventana de 2 días

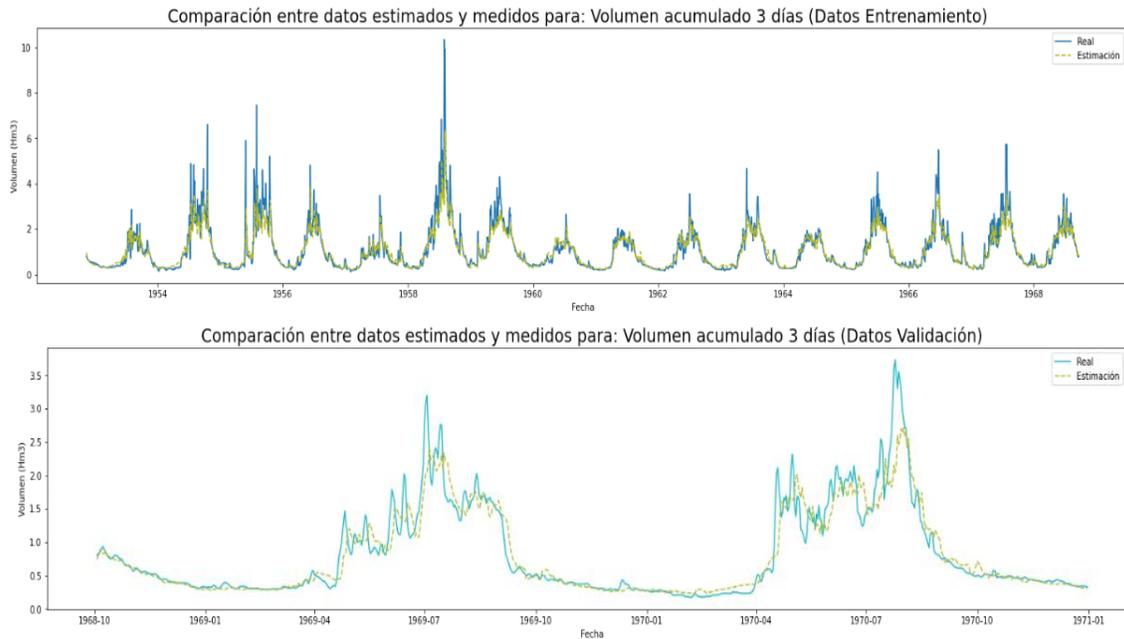


Fig 4.15 Valores Estimados vs Reales: Ventana de 3 días

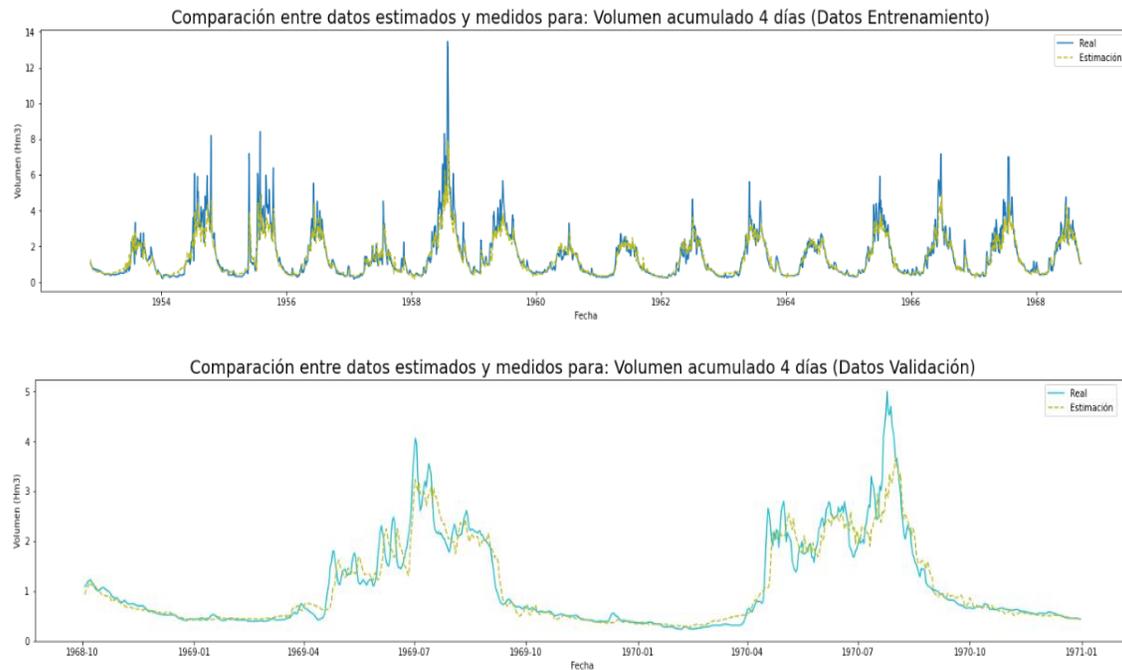


Fig 4.16 Valores Estimados vs Reales: Ventana de 4 días

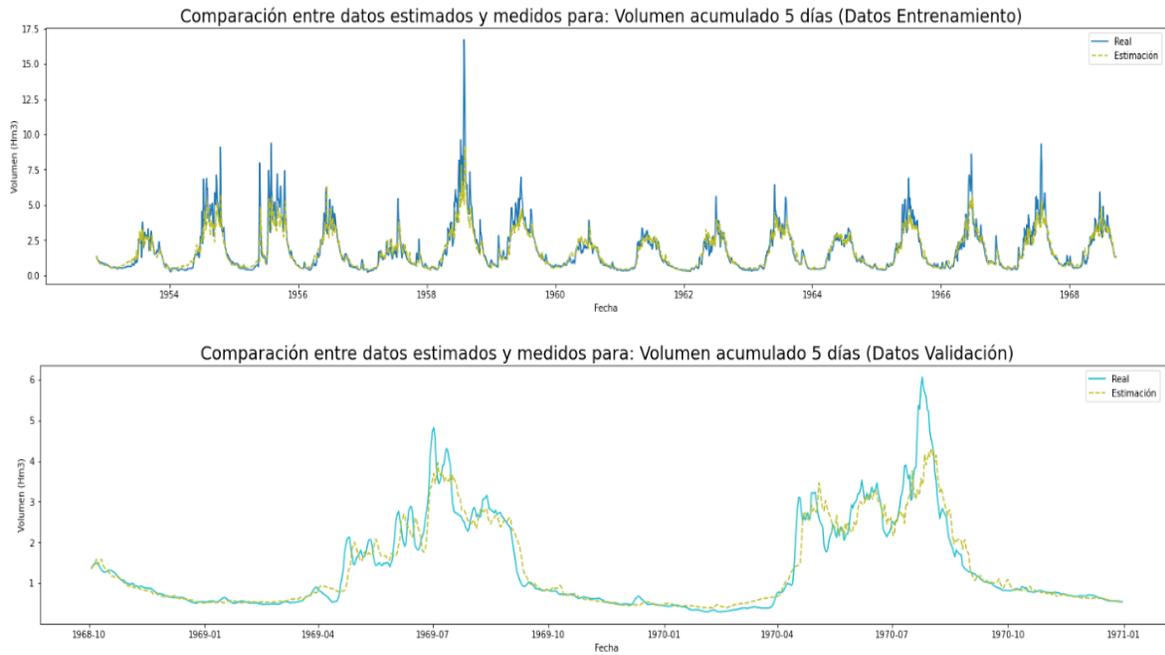


Fig 4.17 Valores Estimados vs Reales: Ventana de 5 días

Capítulo V: Análisis de Resultados

Predicción horizonte 1 día.

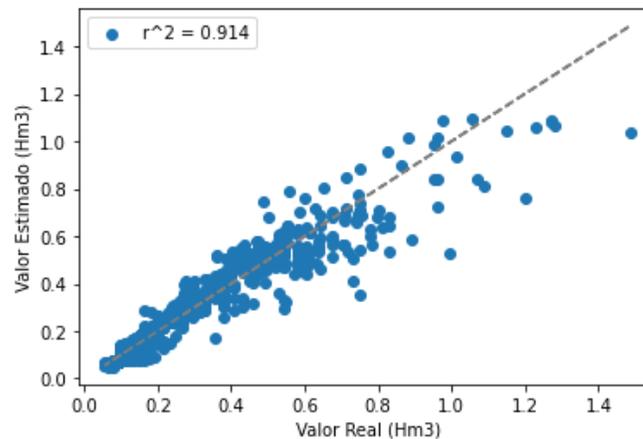


Fig 5.1 Comparación entre los Valores Estimados y Reales del conjunto de validación.

En la figura 4.13 se mostró el ajuste del modelo para los datos de un día en el futuro. Se aprecia que en general, el desempeño del modelo es adecuado; ya que describe adecuadamente el comportamiento de los datos, con un error aceptable en cuanto su magnitud. Sin embargo, este presenta dificultad para describir los valores pico y generalmente los subestima; esto se puede ver con mayor detalle en la figura 5.1, donde se ilustra cómo incrementa la dispersión de los datos en los valores más altos, y cómo estos están sesgados hacia la parte inferior de la línea de perfecto ajuste, lo que indica que en prácticamente todos los picos extremos el modelo subestimó estos valores. Otro fenómeno que se identificó en el comportamiento del modelo fue que generalmente este se acoplaba de manera satisfactoria al volumen real medido; no obstante, existía algunos incrementos que el modelo no reconocía y se ajustaba de manera tardía.

En la figura 5.2 se muestra el comportamiento del modelo con respecto a los valores reales para el periodo de lluvias del año 1969, así como el registro de las estaciones climatológicas. Se puede observar que no existe registro de precipitación para el mes de abril (04) y aun así se presenta un incremento en los escurrimientos a finales de este; lo que indica un subregistro de la precipitación en la cuenca.

En la Fig. 4.2 se observó que existía muy pocas estaciones climatológicas en la parte alta de la cuenca, y las pocas existentes su registro difería del de las demás estaciones climatológicas, así como del de la hidrométrica utilizada. Por lo cual, esta información no pudo ser utilizada para alimentar al modelo.

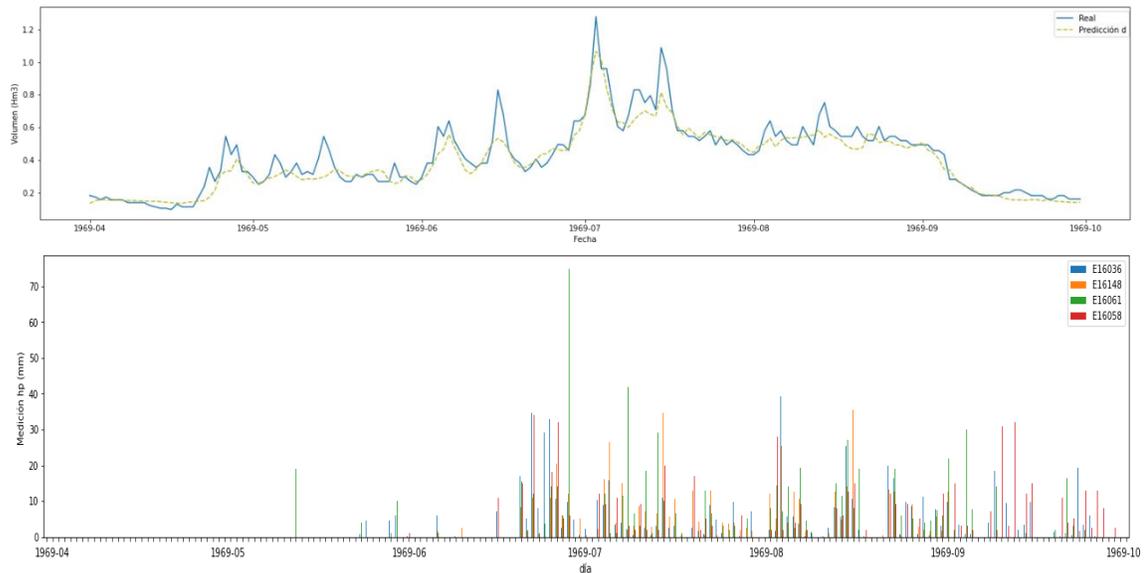


Fig 5.2 Comparación entre los Valores Estimados y Reales del conjunto de validación.

De manera que se intuye que estos incrementos en los esgurrimientos se deben a eventos puntuales de precipitación en la parte alta de la cuenca, de la cual no se tiene registro. No obstante, gracias a que el modelo también utiliza la información de las mediciones previas de esgurrimiento, este logra adaptarse rápidamente a estos cambios y describir de manera satisfactoria el comportamiento general del esgurrimiento.

En contraste, el incremento que se presenta a principios del mes de julio (07) es descrito con bastante precisión ya que este se origina por la precipitación registrada. Por lo cual, se infiere que el problema del modelo para describir adecuadamente algunos de los valores extremos podrían estar relacionado también con el subregistro de la precipitación. De manera que el modelo únicamente este transformando los valores de la precipitación de la parte baja de la cuenca a esgurrimiento, subestimando el total de este debido a la falta de información.

Una segunda hipótesis sobre este comportamiento es debido a la escala temporal; ya que al ser una escala diaria, la resolución no permite captar la evolución en el comportamiento de los datos y muestra un comportamiento más accidentado o irregular.

Predicción horizonte 2 días.

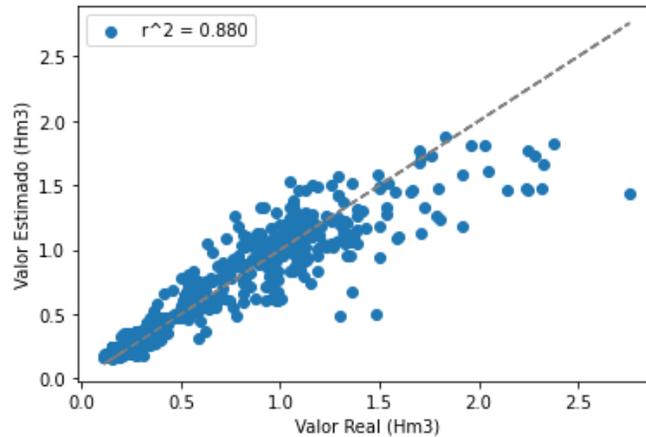


Fig 5.3 Comparación entre los Valores Estimados y Reales del conjunto de validación.

En la figura 4.14 se muestra el ajuste del modelo para ambos conjuntos de datos (entrenamiento y validación) con un horizonte de dos días en el futuro. Se observa que, de manera similar al caso anterior, tiende a subestimar los valores. Esto, debido a que parten de la misma estructura.

No obstante, describe de manera satisfactoria la tendencia en el comportamiento de los datos. Aun cuando no puede predecir los cambios abruptos en el comportamiento del escurrimiento, describe con bastante precisión la tendencia general en el comportamiento de estos. Adicionalmente, gracias a la información previa que utiliza para alimentarse y al tamaño del horizonte de predicción, este parece adaptarse rápidamente a los cambios abruptos del comportamiento de los datos.

En la figura 5.4 se muestra el comportamiento del modelo con respecto a los valores reales para el mismo periodo que el mostrado en la figura 5.2. Se observa como en los puntos en los cuales falla la predicción del modelo son los mismos que en el caso anterior. De igual manera se observa como el modelo ajusta su predicción gracias a que se retroalimenta.

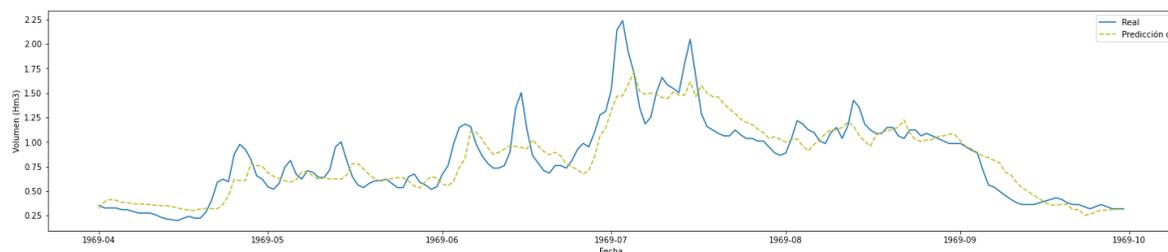


Fig 5.4 Comparación entre los Valores Estimados y Reales del conjunto de validación.

Predicción horizonte 3 días.

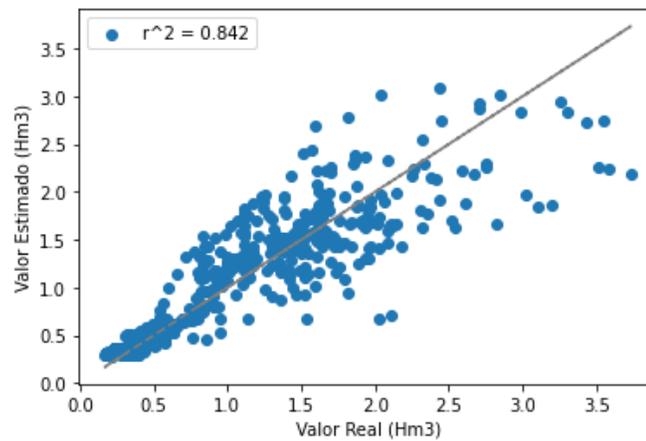


Fig 5.5 Comparación entre los Valores Estimados y Reales del conjunto de validación.

En la figura 4.16 se muestra el ajuste del modelo para ambos conjuntos de datos (entrenamiento y validación) con un horizonte de tres días en el futuro. En este caso el error comienza a magnificarse, aún que sigue prediciendo con cierta exactitud la tendencia en el comportamiento general de los datos. Por lo que podría seguir considerándose como indicador para tomar decisiones con respecto a la operación de la presa “El Bosque”.

En la figura 5.6 se muestra el comportamiento del modelo con respecto a los valores reales para el mismo periodo que el mostrado en la figura 4.17. Los puntos donde más se desfasa la estimación con respecto al valor real corresponden a los mismo que en las estimaciones anteriores.

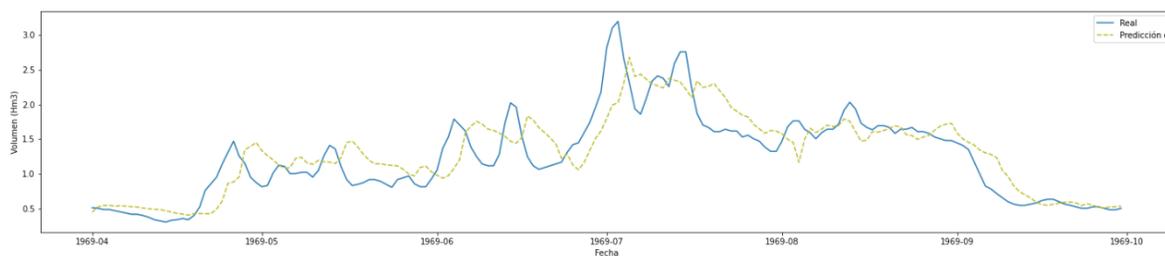


Fig 4.5.6 Comparación entre los Valores Estimados y Reales del conjunto de validación.

Predicción horizonte 4 días.

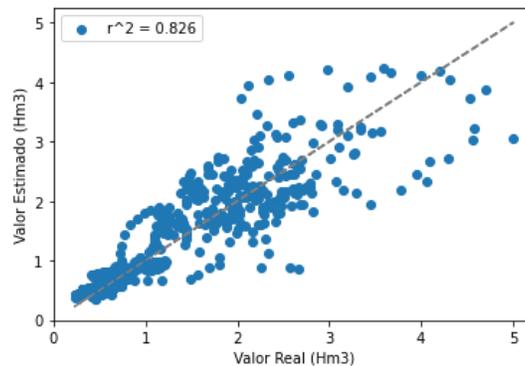


Fig 4.24 Comparación entre los Valores Estimados y Reales del conjunto de validación.

En la figura 4.24 se puede observar el incremento en la dispersión de las predicciones; esto debido a que como se observa en la figura 4.17 aún cuando los datos están suavizados (tiene cambios menos abruptos en su comportamiento), existen varios picos a los cuales el modelo no puede anticiparse debido a que no cuenta con información suficiente para ello; y dado que la ventana de tiempo es demasiado larga el modelo tiene una mayor dificultad para ajustarse. Aún así se considera que la información provista podría servir como un indicador robusto del comportamiento general a mediano plazo para fines de planeación.

En la figura 4.25 se muestra el comportamiento del modelo con respecto a los valores reales para el mismo periodo que el mostrado en la figura 4.17. Llama la atención que, tal como también se puede observar en la figura 4.24, para este horizonte de pronóstico el modelo ya no tiende a subestimar los escurrimientos. Se observa también este comportamiento robusto en cuanto a la estimación de la tendencia de los datos.

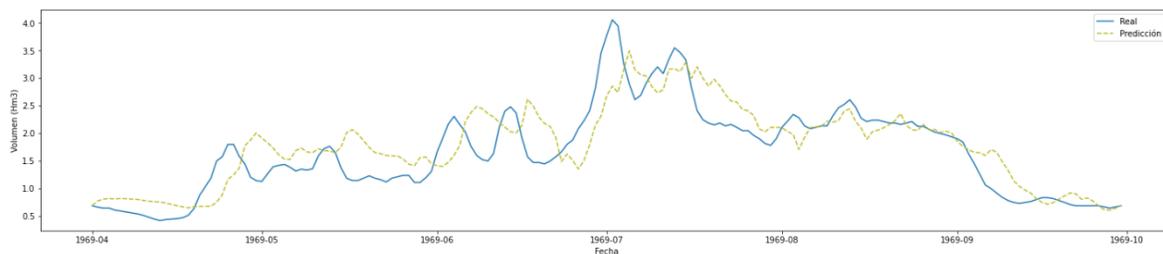


Fig 4.25 Comparación entre los Valores Estimados y Reales del conjunto de validación.

Predicción horizonte 5 días.

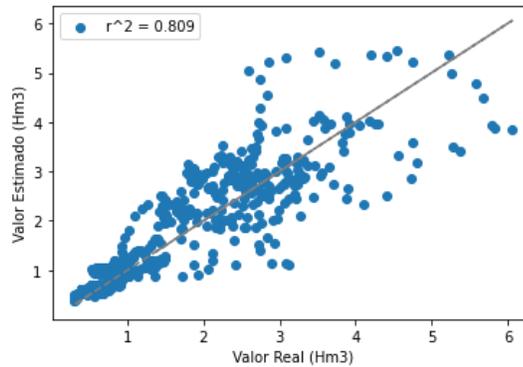


Fig 4.26 Comparación entre los Valores Estimados y Reales del conjunto de validación.

En la figura 4.26 se puede observar, de manera similar que para la estimación de 4 días en el futuro, el incremento en la dispersión de las predicciones causado por el incremento en la ventana temporal. Aun así, el modelo parece seguir brindando una estimación aceptable del comportamiento general de los datos para la ventana temporal que se emplea.

En la figura 4.26 se muestra el comportamiento del modelo con respecto a los valores reales para el mismo periodo que el mostrado en la figura 4.17. Los puntos donde más se desfasa la estimación con respecto al valor real corresponden a los mismo que en las estimaciones anteriores.

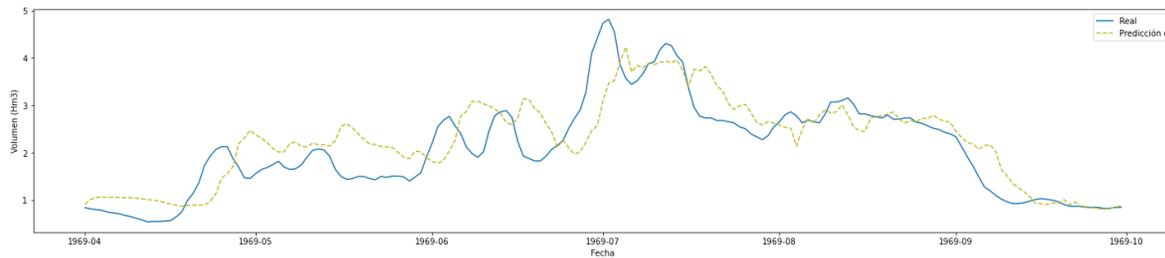


Fig 4.25 Comparación entre los Valores Estimados y Reales del conjunto de validación.

Capítulo VI Conclusiones

Los procesos involucrados dentro del ciclo hidrológico son sumamente complejos y difíciles de describir; y, si bien en los últimos años el avance científico y tecnológico ha generado importantes avances teóricos sobre entendimiento de las dinámicas de estos distintos procesos (atmosféricos, hidráulicos, geológicos, etc), la brecha existente entre estos conocimientos y su aplicación dentro de modelos prácticos que permitan obtener información precisa sobre una determinada variable hidrológica para la toma de decisiones es aún importante.

Lo anterior es, en gran medida, debido a que en pocos sitios existe la instrumentación necesaria para registrar las variables que requiere un modelo sofisticado. Esta situación se agrava en países como México, donde si bien ha existido un esfuerzo importante por tener registros confiables y de fácil acceso a la población; debido a factores geográficos, económicos, políticos y sociales estos aún resultan insuficientes para muchas regiones.

No obstante, contar con una herramienta que permita describir y anticipar el comportamiento de variables hidrológicas tales como el escurrimiento resultan imprescindible para la gestión adecuada de recursos. Particularmente ante el contexto actual en que vivimos donde existe una clara tendencia en aumento hacia la demanda de recursos hídricos; y una disponibilidad con perspectiva en el mediano plazo a disminuir debido a la intervención antropogénica en las cuencas (modificaciones en los usos de suelo, contaminación del agua, forzamiento climático, etc.).

Por ello, se han desarrollado y perfeccionado diversas herramientas que, desde diferentes enfoques (empíricos, conceptuales y físicos), permiten analizar, describir y predecir el comportamiento de variables tales como el escurrimiento en una cuenca. Cada una de estas tiene ciertas ventajas y desventajas dependiendo de las consideraciones propias del enfoque seleccionado. Es importante reconocer que cada uno de estos enfoques tiene un contexto específico de aplicación, por lo que es de suma importancia previo a la selección de cualquiera de estas herramientas realizar el estudio y caracterización del sitio, de manera que se pueda seleccionar aquella herramienta que mejor ajuste a las características específicas del sitio.

Una de estas herramientas, la cual fue el objeto de estudio de este trabajo, es el uso de redes neuronales artificiales. Estas, en su aplicación general dentro del campo de los modelos hidrológicos, ofrecen la ventaja de ser sumamente versátiles; esto debido a su capacidad de mapear cualesquiera dos conjuntos de datos que guarden una relación. Por lo cual tiene aplicaciones potenciales no solo para el desarrollo de modelos lluvia- escurrimiento o de pronóstico, sino también pueden

utilizarse para el relleno de registros de datos (a partir del registro de las estaciones vecinas), análisis de sequías (determinar, a partir de un modelo de clasificación, si los patrones en los registros recientes de precipitación corresponden o no a un periodo de sequía), entre otros.

Este enfoque, al igual que los demás, tiene limitaciones ya que requiere de registros confiables de variables para su aplicación; por lo que, en el contexto de su aplicación para el desarrollo de modelos lluvia-esguerrimiento o de pronóstico, únicamente puede ser utilizado en sitios con algún grado de instrumentación.

Adicionalmente, se encontró durante el análisis bibliográfico sobre el uso de redes neuronales artificiales en el campo de la hidrología que existe una brecha significativa de conocimiento entre aquellos trabajos cuya base parte de un enfoque orientado hacia las ciencias de la computación y aquellos orientados hacia la hidrología. Es decir, más allá de las diferencias propias del enfoque de análisis de cada uno de estos campos, se encontraron discrepancias metodológicas en cuanto al diseño de las redes. Por ejemplo, generalmente los trabajos orientados hacia el campo de la hidrología no consideraban dividir al conjunto de entrenamiento para validar el modelo, o los registros utilizados tienden a ser más cortos que los recomendados por aquellos trabajos cuya base corresponde al campo de las ciencias computacionales.

En este sentido, probablemente una de las principales áreas de oportunidad de esta herramienta es que hace falta una colaboración más estrecha entre ambos campos del conocimiento; así como un mayor desarrollo de trabajos de investigación referentes a este tema que permita homologar las consideraciones metodológicas. Cabe mencionar actualmente ya se observa una tendencia de incremento en el número de publicaciones que combina estos dos campos; sin embargo aún hace falta seguir incrementando este número.

Un último punto referente al uso de esta herramienta que ha limitado su aplicación dentro de la investigación es el hecho de que, tal como se abordó en el capítulo III, es imposible trazar un paralelismo entre los procesos internos de la red para inferir la relación entre cualesquiera dos conjuntos de variables y los procesos físicos que las rigen. Esto es por qué en el trasfondo de la red, esta nunca pretende acercarse a las ecuaciones matemáticas que describen dichos procesos físicos, sino únicamente ajustar los parámetros de su estructura interna para ofrecer una respuesta “mecánica” que se asemeje al resultado esperado. En este sentido queda a consideración de cada investigador/autor el ponderar los resultados que pueden obtenerse a través del uso de esta herramienta con respecto al beneficio que puede obtener en el entendimiento del fenómeno de estudio.

Con respecto a la aplicación de esta herramienta para el pronóstico de los esguimientos dentro del S.C. En términos generales parece ser una opción que se adapta bastante bien a los requerimientos del sistema; ya que es una herramienta económica que, en principio, podría ser escalada fácilmente a las demás cuencas para atender la falta de medición hidrométrica actual.

Adicionalmente, la implementación de esta herramienta dentro de la cuenca tiene un gran potencial para complementarse con otras tecnologías tales como internet de las cosas, árboles de decisiones para la automatización del sistema, entre otras más pertenecientes al enfoque de gestión inteligente del agua. Lo que podría incrementar considerablemente la eficiencia en pro de la explotación sustentable de las cuencas que lo conforman; a la vez de reducir los costos asociados a su operación.

Por parte del análisis puntual de los resultados obtenidos por el modelo: Se concluye que, de acuerdo con los indicadores evaluados durante la validación, el desempeño de este fue satisfactorio para las distintas escalas temporales de pronóstico; aún con la falta de instrumentación suficiente en la parte alta de la cuenca. Lo anterior, gracias al uso de la arquitectura LSTM, que permitió utilizar valores previos del estado del cauce para detectar cuando existían aportaciones distintas a las registradas por las estaciones climatológicas, para adaptarse a esta nueva condición de “flujo base”.

De igual manera se destaca que si bien, la precisión del modelo disminuye considerablemente conforme se amplía la ventana temporal, aún en el caso más extremo (5 días) sigue proveyendo de información importante para la toma de decisiones como lo es la tendencia en cuanto a las aportaciones futuras. Esto gracias a que se utilizó como variable para predecir el volumen acumulado de esguimientos. Ya que haber tratado de predecir eventos puntuales (anticipar y dimensionar una tormenta) basado únicamente en los registros previos de las estaciones hubiera sido complicado; sin embargo, el uso del volumen acumulado resulta en una variable más estable y por tanto fácil de pronosticar, sin dejar de ser una variable relevante para la operación del sistema.

Finalmente cabe mencionar que para su implementación real dentro del sistema, sería necesario validar los resultados haciendo mediciones en sitio; así como comparar el error medio de cada uno de los horizontes de pronóstico con las curvas elevación-área-capacidad de la presa, con el objetivo de dimensionar el efecto que podría tener el error del modelo con la variación en el nivel del embalse y consecuentemente en las pérdidas por filtraciones de la cortina.

Como último aspecto se plantea las siguientes líneas de investigación para trabajos futuros:

1. El implementar este modelo para las demás estaciones hidrométricas en la cuenca y evaluar su desempeño.
2. El implementar el modelo en un sitio donde pueda ser alimentado con información atmosférica; con el objetivo probar su desempeño en la predicción del mediano plazo.
3. Acoplar al modelo con un prototipo de sistema de automatización en la operación de la presa “El Bosque” para optimizar su funcionamiento.

Referencias:

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87(1-2), 45–59.

Aggarwal C.C. (2018) An Introduction to Neural Networks. In: *Neural Networks and Deep Learning*. Springer, Cham. https://doi.org/10.1007/978-3-319-94463-0_1

Alan Yeakley, J., Ervin, D., Chang, H., Granek, E.F., Dujon, V., Shandas, V. and Brown, D. (2016). Ecosystem services of streams and rivers. In *River Science* (eds D.J. Gilvear, M.T. Greenwood, M.C. Thoms and P.J. Wood). <https://doi.org/10.1002/9781118643525.ch17>

Allamy, Haider. (2014). *Methods to Avoid Over-Fitting and Under-Fitting in Supervised Machine Learning (Comparative Study)*.

ALLEN, R. G; PEREIRA, Luis S.; RAES, D. y SMITH, M. (1998). *Crop evapotranspiration: Guidelines for computing crop water requirements*. FAO Irrigation and Drainage Paper 56. Food an Agriculture Organization of the United Nations, FAO. Rome.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. *Artificial neural networks in hydrology. I: Preliminary concepts*. *J Hydrol Eng ASCE*. 2000 Apr; 5(2):115–231

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. *Artificial neural networks in hydrology. II: Hydrologic applications*. *J Hydrol Eng ASCE*. 2000 Apr; 5(2):124–37.2

Assegid, Yirgalem & Naja, Ghinwa & Rivero, Rosanna & Melesse, Assefa. (2012). *Water Quality Monitoring Using Remote Sensing and an Artificial Neural Network*. *Water, Air, & Soil Pollution*. 223. 4875-4887. 10.1007/s11270-012-1243-0.

Ayana, E. K., & Srinivasan, R. (2019). Impact of the Grand Ethiopian Renaissance Dam (GERD) and climate change on water availability in Sudan. *Extreme Hydrology and Climate Variability*, 137–149. doi:10.1016/b978-0-12-815998-9.00012-9

B. D. Hammel. (23-03-2019) . What learning rate should I use?. Recuperado de: <http://www.bdhammel.com/learning-rates/>

Banko, M., Brill, E., 2001. Mitigating the paucity-of-data problem: Exploring the effect of training corpus size on classifier performance for natural language processing, in: *Proceedings of the First International Conference on Human*

Language Technology Research, Association for Computational Linguistics, USA.doi:10.3115/1072133.1072204.

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47–55. doi:10.1038/nature1495

Beven, K. J. (2012). *Rainfall-Runoff Modelling: The Primer* (2nd ed.): Wiley-Blackwell.

Bunge, V., Martínez, J. y Ruiz-Bedolla, K. (2012) “Escenarios de la dinámica hídrica de la región de aporte del sistema Cutzamala”. Documento de Trabajo de la Dirección General de Ordenamiento Ecológico y Conservación de Ecosistemas, Instituto Nacional de Ecología y Cambio Climático, México. Disponible en: http://inecc.gob.mx/descargas/cuencas/doc_trabajo_dinamica_hidrica_cutzamala.pdf

CENAPRED. (2017). Comparación de los modelos numéricos GFS, NAM y WRF del Instituto Mexicano de Tecnología del Agua y del Servicio Meteorológico Nacional de la variable de precipitación, 2015. interno, CENAPRED, Ciudad de México.

Chakravarti, Ankit & Joshi, Nitin & Panjiar, Himanshu. (2015). Rainfall Runoff Analysis using Artificial Neural Network. *Indian Journal of Science and Technology*.

Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke & Travis E. Oliphant. Array programming with NumPy, *Nature*, 585, 357–362 (2020), DOI:10.1038/s41586-020-2649-2

CHEN, S. M., WANG, Y. M., & TSOU, I. (2013). Using artificial neural network approach for modelling rainfall–runoff due to typhoon. *Journal of Earth System Science*, 122(2), 399–405.

Chollet, F., & others. (2015). Keras. <https://keras.io>.

Chow, V. T., D. R. Maidment, and L. W. Mays, *Applied Hydrology*, McGraw-Hill, Singapore, 1988.

Christian Refsgaard, J., Storm, B., & Clausen, T. (2010). Système Hydrologique Européen (SHE): review and perspectives after 30 years

development in distributed physically-based hydrological modelling. *Hydrology Research*, 41(5), 355–377

CONAGUA. (2021) Visor de Radares Meteorológicos. Recuperado de: <https://smn.conagua.gob.mx/es/observando-el-tiempo/radares-meteorologicos-separador/visor-radares> [11/05/2021].

Corzo, Gerald & Solomatine, Dimitri. (2007). Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. *Hydrological Sciences Journal-journal Des Sciences Hydrologiques - HYDROLOG SCI J.* 52. 491-507. 10.1623/hysj.52.3.491.

D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *International Conference on Learning Representations*, 2015.

Devi, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A Review on Hydrological Models. *Aquatic Procedia*, 4, 1001-1007. doi: <http://dx.doi.org/10.1016/j.aqpro.2015.02.126>

E.D. White, Z.M. Easton, D.R. Fuka, A.S. Collick, E. Adgo, M. McCartney, S.B. Awulachew, Y.G. Selassie, T.S. Steenhuis. Development and application of a physically based landscape water balance in the SWAT model *Hydrol. Process.*, 25 (6) (2011), pp. 915-925

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48:71–99.

Elsafi, S. H. (2014). Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*, 53(3), 655–662. doi:10.1016/j.aej.2014.06.010

Environmental Modeling Center (2003). “The GFS Atmospheric Model”. NCEP Office Note 442, Global Climate and Weather Modeling Branch, EMC, Camp Springs, Maryland, 14 pp.

Escolero, O., Kralisch, S., Martínez, S., & Perevochtchikova, M. (2016). Diagnóstico y análisis de los factores que influyen en la vulnerabilidad de las fuentes de abastecimiento de agua potable a la Ciudad de México, México. *Boletín De La Sociedad Geológica Mexicana*, 68(3), 409-427. Retrieved February 28, 2021, from <http://www.jstor.org/stable/24921552>

Farmer, W. H., and Vogel, R. M. (2016), On the deterministic and stochastic use of hydrologic models, *Water Resour. Res.*, 52, 5619– 5633, doi:10.1002/2016WR019129.

Fausett, L., *Fundamentals of Neural Networks*, Prentice Hall, Englewood Cliffs, NJ, 1994.

Fente, D. N., & Kumar Singh, D. (2018). Weather Forecasting Using Artificial Neural Network. 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). doi:10.1109/icicct.2018.8473167

Gauch, M., Mai, J., & Lin, J. (2020). The Proper Care and Feeding of CAMELS: How Limited Training Data Affects Streamflow Prediction. *Environmental Modelling & Software*, 104926. doi:10.1016/j.envsoft.2020.104926

González Villarreal F.J; Carrillo Sosa J.J y Gutiérrez Aviña G .(2014). Aplicación de un Pronóstico Hidráulico en la Ciudad de Villahermosa en el Estado de Tabasco, México. Congreso Latinoamericano De Hidráulica. Santiago, Chile. Agosto , 2014.

Google (2021) Google Colaboratory ¿ Qué es Colaboratory?. Disponible en (21/05/2021)
:https://colab.research.google.com/notebooks/intro.ipynb#scrollTo=5fCEDCU_qrC0

H.Y. GUPTA, K. HSU, and S. SOROOSHIAN (2000) “ Effective And Efficient Modeling For Streamflow Forecasting” : Artificial Neural Networks in Hydrology, 7-22. 2000 Kluwer Academic Publishers.

Half, H. M., and Azmoodeh, M. (1993). “Predicting runoff from rainfall using neural networks.” *Proc., Engrg. Hydrol., ASCE*, New York, 760–765.

Hargreaves, G. H. and Allen, R. G. 2003. History and evaluation of Hargreaves equation. *J. Irrigation Drainage Engineering.-ASCE*. 129:53-63

Hastie, T., Tibshirani, R., & Friedman, J. (2008). Overview of Supervised Learning. *Springer Series in Statistics*, 9–41. doi:10.1007/978-0-387-84858-7_2

Henriksen, H. J., Troldborg, L., Højberg, A. L. & Refsgaard, J. C. 2008 Assessment of exploitable groundwater resources of Denmark by use of ensemble resource indicators and numerical groundwater-surface water model. *J. Hydrol.* 348(1–2), 224–240

Hernández Vargas Miguel Angel (2017) Implementación de la Herramienta Weap al Sistema Cutzamala (Tesis Maestría) Universidad Nacional Autónoma de México.

Hochreiter, Sepp & Schmidhuber, Jürgen. (1997). Long Short-term Memory. *Neural computation*. 9. 1735-80. 10.1162/neco.1997.9.8.1735.

Horton, R. E. (1933). The Role of infiltration in the hydrologic cycle. *Transactions, American Geophysical Union*, 14(1), 446. doi:10.1029/tr014i001p00446

Houze, R. A. (1981). Structures of atmospheric precipitation systems: A global survey. *Radio Science*, 16(5), 671–689. doi:10.1029/rs016i005p00671

Hunter, J. D.: Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.*, 9, 90–95, 2007

Ian Goodfellow, Yoshua Bengio, & Aaron Courville (2016). *Deep Learning*. MIT Press

INEGI (2020). Censo Población y Vivienda. Consultado en (01/04/2021) : <https://www.inegi.org.mx/programas/ccpv/2020/#Tabulados> .

J. D. Salas, M. Markus, And A.S. Tokar (2000) “Streamflow Forecasting Based On Artificial Neural Networks “: *Artificial Neural Networks in Hydrology*, 23-51. 2000 Kluwer Academic Publishers.

J. Wen, J.L. Zhao, S.W. Luo, Z. Han, The improvements of BP neural network learning algorithm, in *Proceedings of 5th International Conference on Signal Processing (IEEE Press, 2000)*, pp. 1647–1649

Johnson, M. S., Coon, W. F., Mehta, V. K., Steenhuis, T. S., Brooks, E. S., & Boll, J. (2003). Application of two hydrologic models with different runoff mechanisms to a hillslope dominated watershed in the northeastern US: a comparison of HSPF and SMR. *Journal of Hydrology*, 284(1-4), 57-76. doi: 10.1016/j.jhydrol.2003.07.005

Kalnay, E. (2002). The continuous equations. In *Atmospheric Modeling, Data Assimilation and Predictability* (pp. 32-67). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511802270.003

Kaz Sato (2021) .What makes TPUs fine-tuned for deep learning? . Disponible en (01/05/2021): <https://cloud.google.com/blog/products/ai-machine-learning/what-makes-tpus-fine-tuned-for-deep-learning>

Knowlton, Barbara & Siegel, Alexander & Moody, Teena. (2017). *Procedural Learning in Humans*. 10.1016/B978-0-12-809324-5.21085-7.

Kratsios, Anastasis. (2021). Universal Approximation Theorems. *Annals of Mathematics and Artificial Intelligence*. Doi: 10.1007/s10472-020-09723-1

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, *Hydrol. Earth Syst. Sci.*, 22, 6005–6022, <https://doi.org/10.5194/hess-22-6005-2018>, 2018.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S., & Nearing, G. S. (2019). Towards Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning. *Water Resources Research*. doi:10.1029/2019wr026065

Kratzert, Frederik & Klotz, Daniel & Brenner, Claire & Karsten, Schulz & Herrnegger, Mathew. (2018). Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks. *Hydrology and Earth System Sciences Discussions*. 1-26. 10.5194/hess-2018-247.

LeCun, Y. A., Bottou, L., Orr, G. B., & Müller, K.-R. (2012). Efficient BackProp. *Neural Networks: Tricks of the Trade*, 9–48. doi:10.1007/978-3-642-35289-8_3

Li, Z., & Fang, H. (2016). Impacts of climate change on water erosion: A review. *Earth-Science Reviews*, 163, 94–117.

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Rafal Jozefowicz, Yangqing Jia, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Mike Schuster, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015. Software available from tensorflow.org.

Matanga, Yves. (2017) Analysis of Control Attainment in Endogenous Electroencephalogram Based Brain Computer Interfaces. DOI: 10.13140/RG.2.2.10493.05608.

Maximilian Sieb, Byeongjoo Ahn, Anqi Yang (2018) .Convex Optimization . Colección personal de Ryan Tibshirani .UC Berkeley EECS dept

Miller, M. P., Johnson, H. M., Susong, D. D., & Wolock, D. M. (2015). A new approach for continuous estimation of baseflow using discrete water quality data: Method description and comparison with baseflow estimates from two existing approaches. *Journal of Hydrology*, 522, 203–210. doi:10.1016/j.jhydrol.2014.12.039

Ming, Y., Cao, S., Zhang, R., Li, Z., Chen, Y., Song, Y., & Qu, H. (2017). Understanding Hidden Memories of Recurrent Neural Networks. 2017 IEEE Conference on Visual Analytics Science and Technology (VAST). doi:10.1109/vast.2017.8585721

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, 26(7), 1465–1473. doi:10.1029/wr026i007p01465

Ngoc Duong Vo, Philippe Gourbesville; Application of deterministic distributed hydrological model for large catchment: a case study at Vu Gia Thu Bon catchment,

Vietnam. *Journal of Hydroinformatics* 1 September 2016; 18 (5): 885–904. doi: <https://doi.org/10.2166/hydro.2016.138>

NIDA. (2010, March 1). *The Brain: Understanding Neurobiology Through the Study of Addiction*. Retrieved from <https://archives.drugabuse.gov/publications/brain-understanding-neurobiology-through-study-addiction> on 2021

Olah, Christopher (27 de Agosto de 2015) *Understanding LSTM Networks*. Github . Disponible en (01/03/2021) <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Oyebode, O., & Stretch, D. (2018). Neural network modeling of hydrological systems: A review of implementation techniques. *Natural Resource Modeling*, e12189

P. J. Werbos, "Backpropagation through time: what it does and how to do it," in *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550-1560, Oct. 1990, doi: 10.1109/5.58337.

Peters, T. (2016). *Water Balance in Tropical Regions*. *Tropical Forestry Handbook*, 391–403. doi:10.1007/978-3-642-54601-3_40

Pujol Reig Lucas (2009) *Predicción De Caudales En Tiempo Real En Grandes Cuencas Utilizando Redes Neuronales Artificiales* (Tesis Doctoral) Universidad Politécnica de Valencia. Doi : 10.4995/Thesis/10251/6422

R.S. Sutton, A.G. Barto (2018), *Reinforcement Learning: An Introduction* , Segunda Edición .MIT Press, Cambridge. ISBN 978026203924

Ramón Domínguez Mora, Gabriela Esquivel Garduño, Baldemar Méndez Antonio, Alejandro Mendoza Reséndiz, Maritza Liliana Arganis Juárez, Eliseo Carrizosa Elizondo. 2008 *Manual del Modelo para Pronóstico de Esguimiento*. Universidad Nacional Autónoma De México,

Reback, J., McKinney, W., Jbrockmendel, Bossche, J. V. D., Augspurger, T., Cloud, P., Gfyoung, Sinhrks, Klein, A., Roeschke, M., Hawkins, S., Tratner, J., She, C., Ayd, W., Terji Petersen, Garcia, M., Schendel, J., Hayden, A., MomIsBestFriend, ... Mortada Mehyar. (2020). *pandas-dev/pandas: Pandas 1.0.3 (v1.0.3)* [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.3715232>

Refsgaard, J. C., Sørensen, H. R., Mucha, I., Rodak, D., Hlavaty, Z., Banský, L., Klucovská, J., Topolska, J., Takac, J., Kosc, V., Enggrob, H. G., Engesgaard, P., Jensen, J. K., Fiselier, J., Griffioen, J. & Hansen, S. 1998 An integrated model for the Danubian Lowland—methodology and applications. *Water Resour. Manage.* 12, 433–465

Richardson, M., Domingos, P. Markov logic networks. *Mach Learn* 62, 107–136 (2006). <https://doi.org/10.1007/s10994-006-5833-1>

Rijtema, P. E. (1965). An analysis of actual evapotranspiration (Doctoral dissertation, Pudoc).

Roshni, T., Jha, M.K. & Drisya, J. Neural network modeling for groundwater-level forecasting in coastal aquifers. *Neural Comput & Applic* 32, 12737–12754 (2020). <https://doi.org/10.1007/s00521-020-04722-z>

Rumelhart, D. E., Durbin, R., Golden, R., and Chauvin, Y. (1995). Backpropagation: The basic theory. In Chauvin, Y. and Rumelhart, D. E., editors, *Backpropagation: Theory, architectures, and applications*, pages 1–34. Lawrence Erlbaum, Hillsdale, New Jersey.

SAGARPA (2020) . Anuario Estadístico de la Producción Agrícola. Consultado en (01/04/2021) : <https://nube.siap.gob.mx/cierreagricola/>

Saha, P. P., Zeleke, K., & Hafeez, M. (2019). Impacts of land use and climate change on streamflow and water balance of two sub-catchments of the Murrumbidgee River in South Eastern Australia. *Extreme Hydrology and Climate Variability*, 175–190.

Samir B. Unandkat, Malina M. Ciocoiu & Larry R. Medsker (2001) *Recurrent Neural Networks: Design and Applications*. Introduction to RNN, doi: 10.1201/9781003040620

Santana, Julio Sergio & Mendoza Uribe, Indalecio & Mateos, Efraín & González, Ricardo. (2012). PRONÓSTICO DE LLUVIA POR CUENCA HIDROLÓGICA EN MÉXICO.

Selvam, A.Mary. (2010). Nonlinear Dynamics and Chaos: Applications in Atmospheric Sciences. *Journal of Advanced Mathematics and Applications*. 1. 10.1166/jama.2012.1014.

Sharma, S. (2017). Activation functions in neural networks. towards data science. Consultado en (01/04/2020): <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.Y. Huang, W. Wang and J.G. Powers (2008). “A Description of the Advanced Research WRF Version 3”. NCAR Technical note NCAR/TN-475+STR, Boulder, CO, USA.

Subhajini A. C. (2018) Application of Neural Networks in Weather Forecasting in *International Journal of Weather, Climate Change and Conservation Research*. Vol.4, No.1, pp.8-18, February 2018.

Sutskever Ilya (2013) Training Recurrent Neural Networks (Tesis Doctorado)Universidad de Toronto .

Tokar, A. S., & Markus, M. (2000). Precipitation-Runoff Modeling Using Artificial Neural Networks and Conceptual Models. *Journal of Hydrologic Engineering*, 5(2), 156–161.

Üneş, Fatih. (2010). Dam Reservoir Level Modeling by Neural Network Approach: A Case Study. *Neural Network World*. 20. 461-474.

Vaze, J., Jordan, P., Beecham, R., Frost, A., Summerell, G. (2012). Guidelines for RainfallRunoff Modelling:Towards best practice model application (pp. 47).

Wit, C. T. (1958). Transpiration and crop yields (No. 64.6). *Verslagen van Landbouwkundige Onderzoekingen*.

Wittenberg, H. (1999). Baseflow recession and recharge as nonlinear storage processes. *Hydrological Processes*, 13(5), 715–726. doi:10.1002/(sici)1099-1085(19990415)13:5<715::aid-hyp775>3.0.co;2-n

World Meteorological Organization. (2009). *Guide to Hydrological Practices Volume II Management of Water Resources and Application of Hydrological Practices*. ISBN :978-92-63-10168-6

World Meteorological Organization. (2011). *WMO Strategic Plan 2012–2015*. WMO-No. 1069, Geneva

World Meteorological Organization. (2017). *Guidelines for Nowcasting Techniques* ISBN : 978-92-63-11198-2

Wu, J. S., Han, J., Annambhotla, S., & Bryant, S. (2005). Artificial Neural Networks for Forecasting Watershed Runoff and Stream Flows. *Journal of Hydrologic Engineering*, 10(3), 216–222

Yang, W.-Y., Li, D., Sun, T., & Ni, G.-H. (2015). Saturation-excess and infiltration-excess runoff on green roofs. *Ecological Engineering*, 74, 327-336. doi: <http://dx.doi.org/10.1016/j.ecoleng.2014.10.023>

Yu, J., Qin, X., Larsen, O., & Chua, L. H. C. (2014). Comparison between Response Surface Models and Artificial Neural Networks in Hydrologic Forecasting. *Journal of Hydrologic Engineering*, 19(3), 473–481.

Zichlin, M. (2011). Procedural Memory. *Encyclopedia of Clinical Neuropsychology*, 2033–2034. doi:10.1007/978-0-387-79948-3_1143

Anexo: Código del Modelo

Importar Librerías

```
import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import tensorflow as tf
from sklearn import preprocessing
from sklearn.preprocessing import Normalizer
from tensorflow.keras.optimizers import SGD, Adam
```

Importar Datos

```
df= pd.read_csv("ElBosque5.csv",parse_dates=["Fecha"],index_col="Fecha",dayfirst=True)
```

Preprocesamiento

Generar conjunto de entrenamiento y validación

```
train_size = 5805
test_size = len(df) - train_size
train, test = df.iloc[0:train_size], df.iloc[train_size:len(df)]
print(len(train), len(test))
#f_columns = ["Volumen", "E16036", "E16148", "DiaSen", "DiaCos"]
f_columns = ["Volumen", "E16036", "E16148", "E16061", "E16058", "E16099", "DiaSen", "DiaCos"]
#f_columns = ["Volumen"]
f_lbs=["Volumenlb4"]

def create_dataset(X, y, time_steps=1):
    Xs, ys = [], []
    for i in range(len(X) - time_steps):
        v = X.iloc[i:(i + time_steps)].values
        Xs.append(v)
        ys.append(y.iloc[i + time_steps])
    return np.array(Xs), np.array(ys)
```

Escalar datos

```
transformera = RobustScaler().fit(train[f_columns])
transformerb = RobustScaler().fit(train[f_lbs])
train_df = transformera.transform(train[f_columns])
test_df =transformera.transform(test[f_columns])
train_lb = transformerb.transform(train[f_lbs])
test_lb = transformerb.transform(test[f_lbs])
```

Diseño de la Red

```
time_steps = 10
# reshape to [samples, time_steps, n_features]
DStrain=pd.DataFrame(train_df)
DStrain[3]=train["DiaSen"].to_numpy()
DStrain[4]=train["DiaCos"].to_numpy()
DStest=pd.DataFrame(test_df)
DStest[3]=test["DiaSen"].to_numpy()
DStest[4]=test["DiaCos"].to_numpy()
```

```
X_train, y_train = create_dataset(DStrain,pd.DataFrame (train_lb), time_steps)
X_test, y_test = create_dataset(DStest,pd.DataFrame (test_lb), time_steps)
print(X_train.shape, y_train.shape)

tf.keras.backend.clear_session()
model = tf.keras.models.Sequential([
    # Shape [batch, time, features] => [batch, time, lstm_units]
    tf.keras.layers.LSTM(50,dropout=0.2, return_sequences=True),
    #tf.keras.layers.Dropout(0.2),
    tf.keras.layers.LSTM(50,dropout=0.2, return_sequences=False),

    #tf.keras.layers.Dense(units=2,activation="relu"),
    # Shape => [batch, time, features]
    tf.keras.layers.Dense(units=1)
])

model.compile(loss='mean_squared_error', optimizer=Adam(lr=0.001),metrics=['mse'])
```

Entrenamiento

```
history = model.fit(
    X_train, y_train,
    epochs=30,
    batch_size=365,
    validation_split=0.20,
    shuffle=False,
)
```

Validación

```
y_pred=model.predict(X_test)
y_prede=model.predict(X_train)
y_ser=model.predict(X_train)
#y_ser=(train_lb_std[0]*y_ser)+train_lb_mean[0]
y_ser2=transformerb.inverse_transform(y_ser).flatten()
y_pt=transformerb.inverse_transform(y_pred).flatten()
print((tf.keras.metrics.mean_squared_error(pronot,prono2t).numpy()))
```