



Universidad de San Carlos de Guatemala
Escuela de Ciencias Físicas y Matemáticas
Departamento de Física

**USO DE HIDROESTIMADORES Y REDES
NEURONALES ARTIFICIALES PARA PRONOSTICAR
NIVELES DE AGUA EN RÍOS DE LA VERTIENTE DEL
OCÉANO PACÍFICO DE GUATEMALA**

Pablo Daniel Martínez Zeceña

Asesorado por Dr. Enrique Pazos Avalos

Guatemala, Julio de 2024

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS

**USO DE HIDROESTIMADORES Y REDES
NEURONALES ARTIFICIALES PARA
PRONOSTICAR NIVELES DE AGUA EN RÍOS DE
LA VERTIENTE DEL OCÉANO PACÍFICO DE
GUATEMALA**

TRABAJO DE GRADUACIÓN
PRESENTADO A LA JEFATURA DEL
DEPARTAMENTO DE FÍSICA
POR

PABLO DANIEL MARTÍNEZ ZECEÑA
ASESORADO POR DR. ENRIQUE PAZOS AVALOS

AL CONFERÍRSELE EL TÍTULO DE
LICENCIADO EN FÍSICA APLICADA

GUATEMALA, JULIO DE 2024

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
ESCUELA DE CIENCIAS FÍSICAS Y MATEMÁTICAS



CONSEJO DIRECTIVO INTERINO

Director	M.Sc. Jorge Marcelo Ixquiac Cabrera
Representante Docente	Arqta. Ana Verónica Carrera Vela
Representante Docente	M.A. Pedro Peláez Reyez
Representante de Egresados	Lic. Urías Amitaí Guzmán García
Representante de Estudiantes	Elvis Enrique Ramírez Mérida
Representante de Estudiantes	Oscar Eduardo García Orantes
Secretario	M.Sc. Freddy Estuardo Rodríguez Quezada

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

Director	M.Sc. Jorge Marcelo Ixquiac Cabrera
Examinador	Dra. Ana Beatriz Cosenza Muralles
Examinador	Dr. Enrique Pazos Avalos
Examinador	Lic. Luis Eduardo Mack Alvizures
Secretario	M.Sc. Freddy Estuardo Rodríguez Quezada

Ref. D.DTG. 008-2024
Guatemala 09 de julio de 2024

El Director de la Escuela de Ciencias Físicas y Matemáticas de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del jefe de la Licenciatura en Física Aplicada, al trabajo de graduación titulado: **"USO DE HIDROESTIMADORES Y REDES NEURONALES ARTIFICIALES PARA PRONOSTICAR NIVELES DE AGUA EN RÍOS DE LA VERTIENTE DEL OCEANO PACÍFICO DE GUATEMALA"**, presentado por el estudiante universitario, Pablo Daniel Martínez Zeceña, autoriza la impresión del mismo.

IMPRÍMASE.



"ID Y ENSEÑAD A TODOS"



M.Sc. Jorge Marcelo Ixquiac Cabrera
Director

*Dedico este trabajo a Dios
a mis padres, Juan José y Marlenny,
a mi hermano, José,
a mi esposa Carolina,
y a mi hija Olivia*

AGRADECIMIENTOS

A mis padres, quienes me han apoyado toda la vida y quienes con su esfuerzo y amor me dieron la oportunidad de soñar y alcanzar las metas que me he propuesto.

A mi esposa, quien con amor y su conocimiento me ayuda a alcanzar nuevas metas cada día, quien me ha dado fuerzas cuando más lo he necesitado y a quien amaré por siempre.

Espero que esta alegría que compartimos ahora sea una de muchas y que sea una pequeña retribución al infinito amor y aprecio que tengo por ustedes.

A mis amigos de la universidad, especialmente a Jeff, Cherk, Bryant, Subadra, Felipe y Mairo por estar presentes y hacer divertidos los tiempos oscuros de pandemia, con quienes tengo los mejores recuerdos de la carrera y quienes soportan mis preguntas y memes a cualquier hora del día.

A mi asesor el Dr. Pazos por su tiempo y dedicación para asesorar este trabajo, así como su buena disposición por compartir su conocimiento. A la Ing. Pérez por su dedicación para guiarme en el desarrollo de mi EPS, el cual sirvió de base a este trabajo y a Pablo Barrios, que aguanta siempre mis preguntas y quien me ayudó a aprender lo que ahora sé sobre hidrología.

A las personas que me han acompañado durante todo el trayecto de mi vida, estudios y vida laboral, a quienes no me alcanzarían las páginas para nombrar y quienes tienen mi aprecio de por vida.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS	VII
ÍNDICE DE CUADROS	XI
INTRODUCCIÓN	XIII
OBJETIVOS	XV
1. Descripción de las fuentes de información	1
1.1. Hidroestimadores	1
1.1.1. Multi-Source Weighted-Ensemble Precipitation	1
1.1.2. Climate Hazards Group InfraRed Precipitation with Station	1
1.1.3. Integrated Multi-satellitE Retrievals for GPM	2
1.2. Nivel de río	2
2. Conceptos de Redes Neuronales Artificiales	5
2.1. Redes Neuronales Artificiales	5
2.1.1. Consideraciones sobre la estructura de una RNA	6
2.1.2. Modelos de RNA utilizando hidroestimadores	7
2.2. Métricas de evaluación	8
2.2.1. Coeficiente de correlación de Pearson	8
2.2.2. Coeficiente de correlación de Spearman	9
2.2.3. Coeficiente de determinación R^2	9
2.2.4. Sesgo porcentual	10
2.2.5. Error Cuadrático Medio - MSE	11
2.2.6. Error Absoluto Medio - MAE	12
2.2.7. Error Absoluto Máximo - MaxAE	12
2.2.8. Función de pérdida de Huber	13
2.2.9. Eficiencia de Nash-Sutcliffe - NSE	13
2.2.10. Matriz de confusión	14

2.3. Entrenamiento de una RNA	16
2.3.1. Construcción de un modelo de RNA	17
2.3.2. Validación cruzada	17
3. Definición del problema	19
3.1. Área de estudio	19
3.1.1. Cuenca del río Acomé	21
3.2. Selección de parámetros de entrada	22
3.2.1. Variables de precipitación directa	23
3.2.2. Variables meteorológicas indirectas	28
3.2.2.1. Variables de resolución horaria	28
3.2.2.2. Variables de resolución mensual	29
3.2.3. Variables de nivel de río	30
3.2.4. Variables de entrada	31
4. Entrenamiento de modelos de RNA	33
4.1. Variaciones de geometría y ventana temporal	34
4.2. Variaciones de función de error	37
4.2.1. Evaluación de modelos	38
4.3. Variación de datos previos de nivel de río	43
5. Resultados finales	47
5.1. Cuenca del río Acomé	47
5.2. Cuenca del río Coyolate	51
5.2.1. Estación Puente Coyolate	51
5.2.2. Estación El Carrizal	55
5.3. Cuenca del río Ocosito	57
6. Discusión de resultados	61
6.1. Variación de parámetros	62
6.2. Pronósticos de nivel de río en la Vertiente del Océano Pacífico	64
6.2.1. Estación La Gomera	66
6.2.2. Estación Puente Coyolate	67
6.2.3. Estación El Carrizal	68
6.2.4. Estación Caballo Blanco	68
7. CONCLUSIONES	71

8. RECOMENDACIONES	75
BIBLIOGRAFÍA	77
ANEXOS	81

ÍNDICE DE FIGURAS

2.1. Modelo de una Neurona Artificial	5
2.2. Validación cruzada utilizando el método de K-pliegues realizando 10 particiones [27]	18
3.1. Vertientes y cuencas hidrográficas de Guatemala delimitadas por el Instituto Geográfico Nacional -IGN-	20
3.2. Estación “La Gomera” en la cuenca hidrográfica del río Acomé	22
3.3. Píxeles correspondientes a la cuenca del río Acomé para cada producto satelital	24
3.4. Coeficiente de correlación de Spearman para las variables del MSWEP	25
3.5. Coeficiente de correlación de Spearman para las variables de CHIRPS	25
3.6. Coeficiente de correlación de Spearman para las variables del IMERG	25
3.7. Correlación por píxel para distintos horarios del segundo día previo al día a pronosticar	27
3.8. Comportamiento de la correlación de las variables meteorológicas indirectas mensuales de precipitación	30
3.9. Comportamiento de la correlación de las variables mensuales (izquierda) y registros diarios (derecha) de nivel de río.	31
4.1. Rendimiento del error y tiempo de entrenamiento por época para distintas cantidades de capas ocultas utilizando 5 folios con datos del MSWEP y el IMERG	37
4.2. Rendimiento del mejor modelo para el IMERG tanto en entrenamiento como validación. El mejor modelo se obtuvo utilizando Error Absoluto Máximo - MAE	41
4.3. Rendimiento del mejor modelo para el MSWEP tanto en entrenamiento como validación. El mejor modelo se obtuvo utilizando Huber Loss, $\alpha = 0.1$	42

4.4.	Rendimiento de los modelos al quitar días previos de nivel de río como variable de entrada para cada producto satelital.	44
4.5.	Rendimiento de los modelos por producto satelital al no considerar datos previos de nivel de río.	44
4.6.	Rendimiento del error y tiempo de entrenamiento por época para distintas cantidades de capas ocultas utilizando 5 folios con datos del MSWEP y el IMERG.	45
5.1.	Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.	48
5.2.	Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.	49
5.3.	Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.	50
5.4.	Píxeles correspondientes a la cuenca del río Acomé para cada producto satelital	51
5.5.	Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.	52
5.6.	Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.	53
5.7.	Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.	54
5.8.	Píxeles correspondientes a la cuenca del río Coyolate para cada producto satelital	55
5.9.	Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.	55
5.10.	Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.	56
5.11.	Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.	56
5.12.	Píxeles correspondientes a la cuenca del río Ocosito por producto satelital	57

5.13. Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.	58
5.14. Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.	59
5.15. Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.	60
8.1. Coeficiente de correlación de Spearman para las variables de MSWEP para la estación “Puente Coyolate”	81
8.2. Coeficiente de correlación de Spearman para las variables del IMERG para la estación “Puente Coyolate”	81
8.3. Rendimiento de los modelos utilizando distinta función de error para los datos satelitales del IMERG durante el entrenamiento (azul) y el año de validación (rojo)	82
8.4. Rendimiento de los modelos para el IMERG utilizando distinta función de error para el período de entrenamiento (azul) y el año de validación (naranja)	83
8.5. Rendimiento de los modelos utilizando distinta función de error para los datos satelitales del MSWEP durante el entrenamiento (azul) y el año de validación (rojo)	84
8.6. Rendimiento de los modelos para el MSWEP utilizando distinta función de error para el período de entrenamiento (azul) y el año de validación (naranja)	85
8.7. Coeficiente de correlación de Spearman para las variables de MSWEP para la estación “Caballo Blanco”	86
8.8. Coeficiente de correlación de Spearman para las variables del IMERG para la estación “Caballo Blanco”	86

ÍNDICE DE CUADROS

2.1. Interpretación de la magnitud del coeficiente de correlación de Pearson según las sugerencias de Cohen [22].	8
2.2. Matriz de confusión para una clasificación binaria [29]	15
3.1. Resultados obtenidos por Martínez y Pérez [27] en “Uso de RNA para la predicción de nivel de agua en ríos en la vertiente del Pacífico de Guatemala”	19
3.2. Información disponible para la estación “La Gomera” según los datos registrados por el INSIVUMEH.	21
3.3. Cantidad de variables distintas de precipitación por producto	24
3.4. Variables meteorológicas directas relevantes	26
3.5. Cantidad de variables meteorológicas horarias indirectas distintas de precipitación por producto	29
3.6. Cantidad de variables meteorológicas mensuales indirectas distintas de precipitación por producto	29
3.7. Cantidad de variables de nivel de río relevantes	31
3.8. Vector de entrada para un modelo de RNA por producto satelital	31
4.1. Rendimiento de distintos modelos de RNA para datos del MSWEP al variar la ventana temporal de entrenamiento y el número de capas ocultas por iteración.	35
4.2. Rendimiento de distintos modelos de RNA para datos del IMERG al variar la ventana temporal de entrenamiento y el número de capas ocultas por iteración.	36
4.3. Combinación ideal de número de folios y capas ocultas, como modelo base para una RNA por producto satelital	36
4.4. Rendimiento de distintos modelos de RNA al variar la función de pérdida del modelo.	38

4.5.	Métricas de evaluación para cada modelo del IMERG y el MSWEP para los datos de entrenamiento (2012-2022) y validación (2023) . . .	39
4.6.	Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “LG-IMERG-MAE”. .	40
4.7.	Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “LG-MSWEP-HB01”. .	40
4.8.	Rendimiento de distintos modelos de RNA al variar la cantidad de días previos de nivel de agua en ríos para el modelo.	43
5.1.	Hiperparámetros finales para entrenar un modelo de RNA por producto satelital	47
5.2.	Rendimiento de los modelos según los parámetros definidos por producto satelital.	47
5.3.	Métricas de evaluación para los modelos finales para la estación “La Gomera” en la cuenca del río Acomé para los datos de entrenamiento (2013-2022) y validación (2023).	48
5.4.	Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “LG-IMERG-MAE-FINAL”.	48
5.5.	Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “LG-MSWEP-HB01”. .	48
5.6.	Rendimiento de los modelos según los parámetros definidos por producto satelital.	51
5.7.	Métricas de evaluación de los modelos para la estación “Puente Coyolate” en la cuenca del Río Coyolate para los datos de entrenamiento (2013-2022) y validación (2023).	52
5.8.	Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “PC-IMERG-MAE”. .	52
5.9.	Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “PC-MSWEP-HB01”. .	52

5.10. Rendimiento de los modelos según los parámetros definidos por producto satelital.	55
5.11. Rendimiento de los modelos por producto satelital.	57
5.12. Métricas de evaluación para los modelos finales para la estación “Caballo Blanco” en la cuenca del río Ocosito para los datos de entrenamiento (2013-2022) y validación (2023).	58
5.13. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “CB-IMERG-MAE”.	58
5.14. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “CB-MSWEP-HB01”.	58
6.1. Umbrales del coeficiente de correlación de Spearman definidos por cada estación hidrométrica considerada.	62
6.2. Cantidad de neuronas de la capa de entrada por estación y producto satelital.	62
6.3. Comparación de modelos de RNA utilizando estaciones meteorológicas y satelitales	65

INTRODUCCIÓN

Pronosticar el nivel de agua de un río durante eventos de lluvia fuerte es importante para la gestión de riesgo, ya que brinda un espacio de tiempo previo con el cual se puedan tomar acciones anticipatorias. Un pronóstico de nivel de agua en un río se puede realizar mediante distintos métodos como aquellos basados en modelos físicos o también métodos basados en datos, que mediante algoritmos de aprendizaje de máquina, modelan las relaciones entre las variables descriptivas del sistema hidrológico.

Distintas investigaciones han utilizado datos de precipitación y modelos de aprendizaje de máquina para pronosticar la relación que hay entre ambas variables, esto por la ventaja que tienen las Redes Neuronales Artificiales (RNA) para reproducir las relaciones desconocidas que existen entre algunas variables como la lluvia y el nivel o caudal del río [6, 20, 24, 25, 27].

Es importante considerar que la estructura de una RNA es dependiente del problema, por lo que esta generalización de que las RNA obtienen mejores rendimientos pueden o no aplicarse al caso a trabajar [15].

En trabajos previos realizados en la Vertiente del Océano Pacífico de Guatemala [27], se utilizaron datos de precipitación de estaciones meteorológicas y estaciones hidrométricas de nivel de río como vector de entrada para un modelo de RNA, sin embargo, en otros experimentos se ha observado que utilizar información rasterizada, presenta una mejora obteniendo predicciones más certeras que utilizando únicamente estaciones meteorológicas [6], por lo que, se toma en consideración esta información para observar si esta mejora también puede ser alcanzada con la información disponible para Guatemala.

Estos modelos pueden ajustarse bien a una escala climática, ya que ciertos errores se suavizan en períodos largos de tiempo [9], pero, al momento de trabajar en períodos cortos, estos errores no se pueden tolerar y deben validarse según el uso que se les dará a los datos. Además, también debe considerarse que cada uno de los modelos se basa en distintas metodologías para estimar la precipitación y que cada

uno ajusta el error de sus estimaciones utilizando distinta cantidad de pluviómetros o datos disponibles, lo que implica que un hidroestimador global tendrá distintos rendimientos en diferentes regiones según sus características geográficas y climáticas.

OBJETIVOS

General

Desarrollar un modelo para pronosticar el nivel de agua en ríos en la vertiente del Océano Pacífico de Guatemala, utilizando redes neuronales artificiales (RNA).

Específicos

1. Evaluar datos de precipitación líquida de distintos productos satelitales que se publican casi en tiempo real y determinar si son útiles para realizar pronósticos de nivel de río utilizando modelos de RNA.
2. Determinar una cuenca hidrográfica de la vertiente del Océano Pacífico como área de interés en la cual se implementen modelos de RNA con los que se pueda determinar una serie de hiper parámetros de referencia que sean útiles para realizar pronósticos de nivel de agua en ríos.
3. Desarrollar una metodología con la cual se disminuya hasta un número deseado la cantidad de predictores a utilizar como vector de entrada de los modelos de RNA, sin afectar significativamente la capacidad de predicción de los mismos.
4. Implementar al menos un modelo de RNA para distintas estaciones hidrométricas de la Vertiente del Océano Pacífico para observar el rendimiento de los mismos y determinar si la metodología es replicable de forma aceptable en la región.

1. Descripción de las fuentes de información

1.1. Hidroestimadores

1.1.1. Multi-Source Weighted-Ensemble Precipitation

El MSWEP [3] es un producto realizado por la organización GloH2O y presenta datos de precipitación a nivel global con una resolución espacial de $0.1^\circ/10\text{ km}$ y resolución temporal de 3 horas. El MSWEP presenta datos de precipitación desde el año 1979 hasta el presente casi en tiempo real. Dicha información es publicada las categorías “Past_nogauge”, “Past” y “NRT” (Near Real Time).

Las categorías de “Past_nogauge” y “Past” contienen los registros históricos de satélites y reanálisis desde 1979 hasta el 31 de diciembre de 2020. En el caso de la categoría “Past”, esta es la que considera los datos de estaciones meteorológicas disponibles. Por otro lado, la categoría “NRT” presenta los datos Casi en Tiempo Real (NRT, por sus siglas en inglés) y a Febrero de 2024, se ha observado que la información es publicada con un retraso de aproximadamente 6 a 7 horas entre la hora actual de publicación y el último registro del MSWEP.

1.1.2. Climate Hazards Group InfraRed Precipitation with Station

El CHIRPS [11] es un producto elaborado por científicos del Centro de Observación y Ciencia de los Recursos Terrestres (EROS) del USGS y es un producto diseñado para estimar lluvia en zonas con pocos o nulos registros de estaciones meteorológicas en tierra.

CHIRPS es un modelo que considera datos satelitales en grilla y datos interpolados de estaciones, con lo que se presenta una estimación de precipitación con una resolución espacial de $0.05^\circ/5\text{ km}$.

A febrero de 2024, CHIRPS cuenta con distintas resoluciones temporales sujetas a distintas regiones disponibles. Considerando que para realizar un pronóstico cercano

se necesita información cercana al tiempo real, se seleccionó la región global con datos preliminares diarios, pero se ha observado que dicha información se publica con 3 días de retraso entre el día actual de la publicación y el último registro de CHIRPS.

1.1.3. Integrated Multi-satellitE Retrievals for GPM

El IMERG [19] es un producto realizado por la NASA que utiliza información de múltiples satélites para estimar precipitación global con una resolución espacial de $0.1^\circ/10\text{ km}$ y resolución temporal de 30 minutos. El IMERG presenta datos de precipitación desde el año 2000 al considerar dentro de sus fuentes de información el programa The Tropical Rainfall Measuring Mission (TRMM, por sus siglas en inglés) y los nuevos datos obtenidos de la misión “Global Precipitation Measurement” (GPM, por sus siglas en inglés).

Los datos de precipitación del IMERG se presentan en las categorías de “Early Run”, “Late Run” y “Final Run” los cuales se publican con distinto retraso cada una. Los datos “Early Run” es el conjunto de datos de precipitación que se publica con un retraso de entre 3 a 4 horas y es el producto con menor latencia pero con menor calidad en las estimaciones presentadas.

Los datos “Late Run” tienen un retraso de 12 horas y mejora en la calidad de sus estimaciones al considerar aquellos datos (satelitales o de estaciones meteorológicas) que no estaban disponibles al momento de realizar el “Early Run” y los datos del “Final Run”, que presentan la mejor estimación de precipitación pero tienen un retraso de tres meses y medio, ya que considera datos mensuales registrados por estaciones meteorológicas para mejorar la estimación.

1.2. Nivel de río

El Instituto Nacional de Sismología, Vulcanología, Meteorología e Hidrología (INSIVUMEH), es la entidad encargada de monitorear y estudiar los fenómenos hidrometeorológicos en Guatemala y con su red de estaciones de Hidrológicas, mantiene un registro de datos diarios de nivel de río en más de 60 estaciones en distintos ríos de Guatemala. Estos datos cuentan con registros digitados desde el año 2003 en adelante, aunque la disponibilidad de datos varía para cada estación.

Actualmente, el INSIVUMEH cuenta con estaciones automáticas que miden el nivel de agua de forma continua y con estaciones “convencionales”, donde un encargado registra el nivel de agua en distintos horarios recurrentes diariamente. Actualmente

se registran 4 o más datos de nivel de agua diariamente, pero en su mayoría, los registros más antiguos solo presentan el nivel de agua a las 7 y a las 17 horas de cada día, por lo que se solo se considerarán estos horarios, con el fin de contar con la mayor cantidad de datos disponibles al momento de realizar un entrenamiento. Bajo el mismo criterio, se descarta la utilización de la red de estaciones automáticas del departamento de Investigación y Servicios Hídricos, ya que cuenta con menos de dos años de registros de nivel de agua.

2. Conceptos de Redes Neuronales Artificiales

2.1. Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) son un modelo matemático que imita a grandes rasgos el funcionamiento de las neuronas biológicas [15] y que se compone de unidades básicas llamadas “Neuronas” con vías de entrada simulando dendritas y vías de salida simulando axones. Otra consideración que imita el comportamiento de las neuronas biológicas, es la existencia de conexiones entre neuronas, cada una con distinta fuerza según el aprendizaje de la red, lo cual es representado en el modelo matemático como un “peso” que representa el grado de conexión de las neuronas conectadas por dicha vía. El modelo de una RNA puede ser descrito por la ecuación propuesta por Olabe [30]:

$$NET = X * W \quad (2.1)$$

Siendo “NET” la salida, “X” el vector de entrada y “W” el vector de pesos. Para generalizar la aplicación del método de RNA debe añadir además una función de activación “F” a la señal de salida de la red, ya que así se consigue representar una mayor variedad de funciones sean estas lineales o no. Considerando este añadido al modelo matemático, se obtiene una estructura final como se muestra en la Figura 2.1.

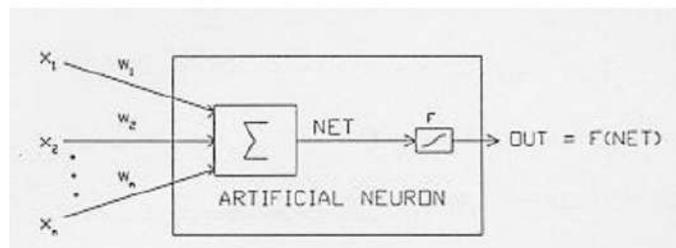


Figura 2.1. Modelo de una Neurona Artificial

Fuente: Redes Neuronales artificiales y sus aplicaciones [30].

Se ha demostrado con anterioridad que una RNA multicapa pre-alimentada es capaz de aproximar cualquier función lineal, no lineal o aproximaciones parciales de funciones, con cualquier grado de precisión deseada [18], pero se debe recalcar que el rendimiento de la RNA está sujeto al método de entrenamiento, la construcción de su estructura (como la cantidad de capas ocultas de la red), la cantidad y calidad de los datos disponibles, así como la relación que exista entre los datos de entrada y salida [37], etc. Por lo que de no alcanzarse el rendimiento adecuado, debe revisarse con detenimiento el proceso de entrenamiento ya que este es completamente dependiente del problema a resolver y no cuenta con una solución universal para cualquier problema de interés.

2.1.1. Consideraciones sobre la estructura de una RNA

Al momento de iniciar el proceso de construcción de la estructura de una RNA es importante definir previamente los criterios de rendimiento que se evaluarán sobre los resultados de la misma, ya que esto afectará significativamente las variaciones que deben realizarse para optimizar o no un parámetro seleccionado [25].

Holger R. Maier [25] presenta el comportamiento general de las RNA según distintos criterios de rendimiento. Por ejemplo, una red neuronal “pequeña” suele realizar una mejor generalización del pronóstico, requiere menos recursos y tiene una mejor rapidez de procesamiento. Por otro lado, las redes “grandes” tienen la ventaja de aprender en pocos ciclos de entrenamiento y pueden crear mejores zonas de decisión, así como evitar de mejor forma los mínimos locales en la hipersuperficie de error, pero requieren de una mayor cantidad de recursos y tiempo para poder desarrollar el entrenamiento.

Independientemente de las métricas de rendimiento seleccionadas, es importante considerar las siguientes características de las RNA:

- Una red neuronal sencilla, sin capas ocultas se comporta de forma similar a un modelo lineal [37].
- Se necesita al menos una capa oculta para que una RNA sea un aproximador universal de funciones continuas [15] [18]. Y en caso de ser una aproximación lineal, la estructura realizará un mapeo más suave [37].
- La geometría de la red es altamente dependiente del problema a resolver [12][25] y en general, para encontrar el modelo que mejor se ajuste a la respuesta esperada será mediante ensayo y error [5][12][20], ya que resultados previos

de otros modelos podrían no aplicarse al nuevo caso a trabajar [15].

- Las RNA no son capaces de extrapolar información mas allá del rango de datos proporcionados durante el entrenamiento [25], por lo que el dominio del problema debe ser igual o un subconjunto del dominio del entrenamiento y deben estar uniformemente distribuidos. [10]
- Una RNA es más propensa a sufrir de sobreajuste en los resultados al tener más parámetros por ajustar que un modelo estadístico [15] , por lo que deben considerarse métricas que identifiquen este comportamiento.

2.1.2. Modelos de RNA utilizando hidroestimadores

Considerando que el problema a resolver es realizar un pronóstico de nivel de agua de ríos, se debe considerar como punto de partida que este fenómeno forma parte del ciclo hidrológico, ya que como menciona Dignman [23], “Cualquier superficie natural expuesta puede considerarse un volumen de control sobre el que opera el ciclo hidrológico”. En este caso se considera la delimitación de una cuenca hidrográfica como la superficie expuesta donde ocurre el proceso de lluvia-escorrentía, siendo la lluvia (precipitación líquida) la entrada el sistema y el caudal del río la salida del mismo. Aunque hay otros elementos que forman parte del proceso lluvia-escorrentía como lo son la evaporación, la transpiración entre otros, se pueden considerar como despreciables si se comparan con el caudal durante una tormenta [7].

Dentro de las ventajas que presenta considerar únicamente datos de precipitación líquida como entrada para nuestro sistema hidrológico, es la variedad de fuentes de información disponibles con dichos datos. En el caso de Guatemala existen instituciones que cuentan con pluviómetros en tierra que registran precipitación sobre determinados puntos, así como misiones satelitales que registran una estimación de cuanta lluvia pudo haber precipitado sobre ciertas zonas del planeta para distintas resoluciones temporales.

Las RNA son muy utilizadas para modelar series temporales no lineales [40] y para aplicaciones hidrológicas se ha ido aumentando con el tiempo el uso de modelos de RNA por su ventaja para reproducir relaciones desconocidas entre el conjunto de datos de entrada y la salida deseada [5] y en estudios previos, se ha mostrado que los modelos de RNA multicapa se adaptan bien para realizar pronósticos de nivel de río con información de imágenes satelitales [6][28], ya que las mismas suelen tolerar información incompleta o con ruido y en algunos casos pueden presentar un mejor rendimiento que otro tipos de modelos en condiciones similares [28] [37].

Comparando el rendimiento de modelos que usan datos de lluvia de pluviómetros en tierra, contra modelos que usaron datos satelitales rasterizados, se observan resultados similares o en algunos casos, como los presentados por Chen Zeqiang [6], con un rendimiento superior.

2.2. Métricas de evaluación

2.2.1. Coeficiente de correlación de Pearson

Las dos variables deben ser de intervalo o de razón aunque no es necesario que ambas tengan el mismo nivel de medición.

El coeficiente de correlación de Pearson es una métrica que representa la fuerza con la que dos variables se vinculan y la dirección de la relación lineal de dichas variables. Fue presentada por Karl Pearson en 1885 en una ecuación definida de la siguiente forma:

$$R_{XY} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{(\sum(X - \bar{X})^2)(\sum(Y - \bar{Y})^2)}} \quad (2.2)$$

Coeficiente de correlación R de Pearson [22]

El coeficiente de Pearson es un número real, adimensional, con un rango definido de -1 a 1, donde la fuerza de la relación entre las variables X,Y es más fuerte conforme el valor tienda a los extremos antes citados [22] y la cual puede ser clasificada según el valor obtenido tomando como referencia la propuesta de Cohen [8], donde se toma el valor absoluto de R ya que la fuerza de la relación es independiente del signo.

Cuadro 2.1. Interpretación de la magnitud del coeficiente de correlación de Pearson según las sugerencias de Cohen [22].

Rango de valores de R	Interpretación
$0.00 \leq R < 0.10$	Correlación nula
$0.10 \leq R < 0.30$	Correlación débil
$0.30 \leq R < 0.50$	Correlación moderada
$0.50 \leq R < 1.00$	Correlación fuerte

Una consideración importante a realizar es que la determinación del mismo es simétrica, por lo que se obtendrá el mismo resultado si se cambia el orden de X

y Y ($R_{XY} = R_{YX}$), por lo que no debe cometerse el error de definir una variable como explicativa de la otra ya que no es posible determinarlo únicamente con este parámetro y existe la posibilidad de que sea una variable externa Z la que origine tanto a X como a Y o ser una relación puramente por el azar [22].

2.2.2. Coeficiente de correlación de Spearman

El coeficiente de correlación de Spearman es un coeficiente no paramétrico que parte del coeficiente de correlación de Pearson sustituyendo las variables X y Y por los rangos de dichas variables lo que define el coeficiente como la siguiente ecuación:

$$\gamma_S = 1 - \frac{\sum_{i=1}^n d_i^2}{n^3 - n} \quad (2.3)$$

Coeficiente de correlación γ_S de Spearman [33]

donde n es el número de observaciones y d_i es la diferencia entre los rangos de cada par de datos. La sustitución realizada permite que el coeficiente de Spearman pueda aplicarse aún cuando no se cumplan los siguientes supuestos de la correlación de Pearson:

- Debe existir una relación de tipo lineal entre las variables (X , Y).
- Para cada valor de X , hay una subpoblación de valores de Y normalmente distribuidas.
- Para cada valor de Y hay una subpoblación de valores X que están normalmente distribuidas

Ya que γ_S parte del coeficiente de Pearson, también se encuentra acotada en un rango de -1 a 1 , donde la fuerza de la relación entre variables es mayor entre más se acerque a los límites antes mencionados. Además, γ_S mide la tendencia de X y Y a relacionarse en forma monótona creciente o decreciente y no se encuentra restringido a descubrir sólo una asociación lineal entre las variables [33].

Algunas de las ventajas del coeficiente de Spearman es que los supuestos son menos estrictos, es más robusto a valores atípicos (outliers)

2.2.3. Coeficiente de determinación R^2

El coeficiente de determinación R^2 se utiliza como medida de bondad del ajuste entre dos o más variables que “recoge cómo en términos generales la recta de regre-

sión ajustada resume o describe los datos”[34] y puede ser utilizado para evaluar la capacidad de un modelo para predecir valores distintos de la media, que se asume como el modelo por defecto en caso no hubiera otro disponible [5]. Este coeficiente está definido mediante la siguiente ecuación:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.4)$$

Coeficiente de determinación R^2 [5]

donde y_i son las observaciones, \hat{y}_i el valor calculado por un modelo y \bar{y} el promedio de las observaciones. El rango del coeficiente R^2 está acotado entre 0 y 1, siendo un valor de 1 un ajuste lineal perfecto y el valor de 0 que no existe representatividad del modelo lineal. El denominador de la ecuación 2.4 representa la variación total de las observaciones respecto de su media muestral, mientras que el numerador es la variación residual o no explicada de los valores observados. Si los residuos son grandes, estos alejan la línea de regresión estimada de los valores observados, dejando a la ecuación como poco representativa. mientras que con valores pequeños, la representación será elevada [34].

Un valor de R^2 por arriba de 0.9 indica un rendimiento satisfactorio del modelo, los valores entre 0.8 a 0.9, indican un buen rendimiento y aquellos por debajo de 0.8 corresponderían a un rendimiento insatisfactorio [5].

Es recomendable utilizar el coeficiente R^2 como una medida para evaluar el grado de ajuste de un modelo respecto al conjunto de observaciones, pero debe complementarse con otras métricas de medición [34] ya que R^2 no es del todo confiable debido a la dispersión de sus sesgos [31] que pueden ser sustanciales en muestras pequeñas y que prácticamente desaparecen al tener más de 50 observaciones y disminuyen a medida que aumenta el conjunto de muestras.

2.2.4. Sesgo porcentual

El sesgo porcentual (PBias, por sus siglas en inglés), es una medida sobre la tendencia promedio de los datos pronosticados respecto a las observaciones, definido mediante la siguiente ecuación:

$$PBias = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{\sum_{i=1}^n y_i} * 100 \quad (2.5)$$

Sesgo Porcentual [6]

El cálculo del PBias determina la diferencia entre las observaciones y los datos simulados normalizados por el promedio de las observaciones. Un valor negativo del PBias indica que el conjunto de datos simulados es mayor que el conjunto de observaciones, por lo que se considera que el modelo tiende a subestimar las observaciones y en caso contrario, un valor positivo indicaría que el modelo subestima las observaciones [4]. El valor óptimo para el Pbias es 0 o cercano al mismo en ambas direcciones, ya que indicaría una correcta calibración del modelo y con un sesgo mínimo.

2.2.5. Error Cuadrático Medio - MSE

El Error Cuadrático Medio (MSE, por sus siglas en inglés) es una métrica estadística utilizada para evaluar la precisión de un modelo mediante la diferencia entre los datos pronosticados y las observaciones, el cual se define mediante la siguiente ecuación:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (2.6)$$

Error Cuadrático Medio (MSE, por sus siglas en inglés) [20]

El MSE es un valor que indica la capacidad de un modelo para reproducir la realidad y para modelos que pronostican una variable continua, es una métrica ideal para evaluar el rendimiento del modelo ya que por su definición, para datos normalmente distribuidos, minimizar el MSE equivale a minimizar la entropía cruzada, por lo que en términos probabilísticos, minimizar el MSE equivale a maximizar la verosimilitud de los datos [17] ya que el mejor modelo es aquel que minimiza la entropía cruzada entre las predicciones del modelo y los datos de entrenamiento.

Dentro de las ventajas que presenta el MSE respecto a otras métricas es que es fácil de calcular, su derivada parcial respecto a los pesos también es de fácil cálculo y en la interpretación física del mismo, al penalizar los errores grandes, se puede utilizar para identificar si un pronóstico es bueno para caudales o niveles altos [25] y debido a que los modelos suelen pronosticar mejor los niveles bajos que los altos, diversos trabajos lo consideran la métrica de evaluación correcta [20]. Se debe considerar que al comprimir toda la información del modelo en un solo número, el MSE proporciona poca información sobre que aspectos del rendimiento del modelo se comportan de una buena o mala manera [17].

Para poder utilizar el MSE como función de pérdida es necesario que la información siga una distribución normal para obtener resultados óptimos [25] pero en el caso de información que contiene datos atípicos esto no se cumple, por lo que es necesario considerar otras métricas que puedan ajustarse de una mejor forma.

Otra métrica bastante utilizada relacionada al MSE es la raíz cuadrada del mismo, conocida como RMSE, por sus siglas en inglés. Se debe aclarar que tomar la raíz del MSE no afecta a los rangos relativos de los modelos, pero tiene la ventaja de devolver una métrica con las mismas unidades que las observaciones y el valor pronosticado, por lo que brinda una interpretación física más sencilla que representa el error estándar para errores distribuidos normalmente [16].

2.2.6. Error Absoluto Medio - MAE

El Error Medio Absoluto (MAE, por sus siglas en inglés) es una medida utilizada para determinar la diferencia entre un valor pronosticado y las observaciones. Esta medida está definida mediante la siguiente ecuación:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.7)$$

EL MAE está expresado en las mismas unidades que los datos originales por lo que facilita la interpretación del valor obtenido como la magnitud de las diferencias entre los valores pronosticados y las observaciones, sin ser penalizado fuertemente por valores atípicos en los datos de entrenamiento y es por esta característica que se le puede considerar una pérdida “robusta” [13], aunque debe considerarse que su eficacia irá disminuyendo al aumentar el número de valores atípicos volviéndose la cantidad de los mismos comparable al resto de los datos. Se debe mencionar que aunque el RMSE es una métrica muy utilizada al momento de evaluar el rendimiento de un modelo, este no es consistente con el MAE y no puede tomarse como una medida verdadera o confiable del “error promedio” [41], por lo que de querer evaluarse este aspecto, debe considerarse que la mejor medida del error promedio es el MAE, la cual complementará el análisis de rendimiento de los modelos considerados.

2.2.7. Error Absoluto Máximo - MaxAE

El Error Máximo Absoluto es una medida utilizada para determinar la diferencia máxima entre un conjunto de valores pronosticados y observaciones, este valor retorna la diferencia más grande de todas las diferencias obtenidas y se presen-

ta en las mismas unidades que los valores originales. Esta medida se puede definir de la siguiente manera:

$$MaxAE = \max\{|y_i - \hat{y}_i| \mid 1 \leq i \leq n\} \quad (2.8)$$

Error Absoluto Máximo (MaxAE, por sus siglas en inglés)

2.2.8. Función de pérdida de Huber

La función de pérdida de Huber (Huber Loss) es una medida de la diferencia que existe entre un valor observado y uno pronosticado que se ajusta de buena manera a conjuntos de datos con valores atípicos y que suelen afectar en mayor medida a otras medidas definidas. La pérdida de Huber está definida mediante la siguiente ecuación:

$$\rho_\delta(z) = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2 & \text{si } |y_i - \hat{y}_i| \leq \delta, \\ \delta(|y_i - \hat{y}_i| - \frac{1}{2}\delta) & \text{si } |y_i - \hat{y}_i| > \delta \end{cases} \quad (2.9)$$

Pérdida de Huber [21]

donde δ es un hiperparámetro de ajuste que determina el umbral con el cual se realiza una penalización lineal o una penalización cuadrática a la diferencia obtenida. La pérdida de Huber se define partiendo de los pros y contras de las pérdidas absolutas (como el MAE) y cuadráticas (como el MSE), ya que, aunque las primeras son robustas y tienen a distorsionarse menos ante la presencia de valores atípicos (siempre y cuando no sean comparables en cantidad con el resto de los datos) no son diferenciables en todos sus puntos. Por otro lado, las pérdidas cuadradas al ser convexas, tienen una rapidez de aprendizaje superior debido a ser una función suave y diferenciable [13] y la pérdida de Huber mezcla el comportamiento de ambas, ya que es robusta ante datos atípicos, pero tiene la suavidad y diferenciable de una pérdida cuadrática [21] por lo que tendrán una convergencia rápida cerca de una pérdida despreciable [13] y permitirá que los parámetros de la RNA se ajusten sin dejarse influenciar fuertemente por los valores atípicos de los datos [21].

2.2.9. Eficiencia de Nash-Sutcliffe - NSE

La eficiencia de Nash-Sutcliffe es una medida utilizada para evaluar el rendimiento de modelos de pronóstico y junto con el MSE es ampliamente utilizado en la hidrología [14] y se define mediante la siguiente ecuación:

$$NSE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.10)$$

Eficiencia de Nash-Sutcliffe [42]

La cual puede relacionarse también con el MSE mediante la siguiente forma:

$$NSE = 1 - \frac{MSE}{\sum_{i=1}^n (y_i - \bar{y})^2 / N} \quad (2.11)$$

Eficiencia de Nash-Sutcliffe relacionada con el MSE [14]

El NSE busca mejorar algunas de las deficiencias del MSE como puede ser la variabilidad en su magnitud y que no es fácilmente comparable entre un modelo y otro, sobre todo cuando los mismos presentan su salida en distintos formatos. Esto lo consigue ya que el NSE puede interpretarse como la capacidad del modelo respecto a un modelo de referencia (la media de los valores observados) por lo que es una medida adimensional y que se encuentra definida en el intervalo $(-\infty, 1]$ [42].

Los valores del NSE que sean inferiores a cero implican que el modelo tiene un peor rendimiento en el MSE que la media de las observaciones o dicho de otra forma, que utilizar la media de las observaciones como pronóstico en cada una de iteraciones, tiene un mejor rendimiento del MSE que el modelo proporcionado. Por otro lado, si el valor del NSE es superior a cero, debe interpretarse que a medida que aumente y se acerque a 1, así aumenta la calidad del modelo [42].

2.2.10. Matriz de confusión

Una matriz de confusión es una métrica utilizada para analizar el rendimiento de modelos de clasificación comparando los resultados pronosticados contra los resultados reales. En el caso de considerar solo dos clases a clasificar (clasificación binaria), se cuenta con una matriz de 2×2 como se presenta a continuación:

En la matriz de confusión, los Verdaderos Positivos (TP, por sus siglas en inglés) y los Verdaderos Negativos (TN, por sus siglas en inglés) son los casos donde la realidad se corresponde correctamente con el pronóstico, por lo que se ha realizado la clasificación correctamente. Por otro lado, los Falsos Positivos (FP, por sus siglas en inglés) representarían las falsas alarmas emitidas (en el caso de estar clasificando un aumento en el nivel de río como instancia positiva) y los Falsos Negativos (FN, por sus siglas en inglés) representarían las alertas perdidas o no emitidas.

Cuadro 2.2. Matriz de confusión para una clasificación binaria [29]

	Observación Positiva	Observación Negativa
Estimado Positivo	Verdaderos Positivos (TP)	Falsos Positivos (FP)
Estimado Negativo	Falsos Negativos (FN)	Verdaderos Negativos (TN)

De la matriz de confusión se pueden obtener distintos parámetros que brindan información sobre el rendimiento del pronóstico, entre estos se encuentran las siguientes:

La exactitud es una cantidad que representa el porcentaje de datos que han sido clasificados correctamente por el modelo y está definida como:

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN}. \quad (2.12)$$

La precisión, representa la proporción de aciertos existentes de entre todas las instancias que fueron pronosticadas como positivas y entre más cercano sea el valor de la precisión a 1, indicaría que existe una menor proporción de falsas alarmas emitidas. La precisión está definida como:

$$Precision = \frac{TP}{FP + TP}. \quad (2.13)$$

La sensibilidad mide la proporción de instancias positivas pronosticadas respecto a todas las instancias positivas reales, lo cual indica si existe una mayor o menor cantidad de verdaderos positivos, si el valor de la sensibilidad es mayor o menor respectivamente. La sensibilidad se define como:

$$Sensibilidad = \frac{TP}{TP + FN}. \quad (2.14)$$

El valor predictivo negativo representa la proporción de instancias negativas de aciertos existentes que fueron pronosticadas como negativas y entre más cercano sea el valor del valor predictivo negativo a 1, indicaría que existe una menor proporción

de alertas perdidas. El NPV (por sus siglas en inglés), está definido como:

$$\text{Valor Predictivo Negativo} = \frac{TN}{TN + FN}. \quad (2.15)$$

La especificidad representa la proporción de instancias negativas pronosticadas respecto a todas las instancias negativas reales, lo cual indica si existe una mayor o menor cantidad de verdaderos negativos si el valor de la especificidad es mayor o menor respectivamente. La especificidad se define como:

$$\text{Especificidad} = \frac{TN}{TN + FP}. \quad (2.16)$$

2.3. Entrenamiento de una RNA

El proceso de “entrenamiento” de un modelo de RNA es realizar el ajuste de los pesos y sesgos de las conexiones de la red para que la misma sea capaz de resolver un problema deseado. Este proceso, dependiendo de como se realice puede clasificarse como supervisado o no supervisado.

Para el caso de realizar un pronóstico de nivel de agua en ríos se considera que el entrenamiento será supervisado, ya que se proveen los datos seleccionados de entrada para ser procesados por la RNA y el resultado de dicho proceso será comparado con el valor real para que así pueda darse el ajuste de los pesos de la red.

Ya que se asume que una RNA no tendrá ningún conocimiento a priori en su estructura previo a ser entrenada, todos los pesos de la red se inician de forma aleatoria [20] y una vez proporcionados los datos (si se utiliza el algoritmo de retro propagación) el error entre la salida resultante y la observación se propaga hacia las capas previas para ajustar los pesos de cada uno de los nodos de la red [26] [25]. El proceso deberá repetirse una y otra vez hasta que se utilicen todos los datos del conjunto de entrenamiento y en reiteradas ocasiones hasta que los pesos se definan correctamente y se alcance un valor de error lo suficientemente pequeño [25], [12], o cuando el usuario determine que el valor del error es aceptable, pero no existe una métrica que indique cuando es el tiempo óptimo para detener el proceso de entrenamiento y es peligroso debido a que se puede detener prematuramente antes de alcanzar el mínimo error deseado.

2.3.1. Construcción de un modelo de RNA

Para realizar la construcción de un modelo de RNA es importante definir los criterios por los que será evaluado el modelo ya que estos tendrán un impacto significativo en la arquitectura y en las técnicas de optimización de pesos a utilizar [25].

En el caso de utilizar el algoritmo de retro propagación, se requiere estimar la derivada parcial de la función de costo según cada nodo de la RNA. Los parámetros se ajustarán según el descenso del gradiente a un paso específico definido como “tasa de aprendizaje” [32]. Este tamaño de paso γ representa la magnitud del ajuste que se realizará en cada iteración. Además, el método de retro propagación, si el valor de γ es más pequeño, el aprendizaje de la RNA será más lento y podría quedarse atrapado el error en un mínimo local. Por otro lado, si el valor de γ es más grande, se puede evitar este problema, pero si el tamaño es muy grande, puede quedarse oscilando sin alcanzar un mínimo [25].

Según las recomendaciones de Ms. Sonali. B. Maind [26], para que una RNA pueda resolver un problema, el diseñador debe revisar distintos aspectos como lo son los datos de entrada y salida, el número de capas, el número de elementos por capas, las conexiones entre capas, las funciones de suma, transferencia y entrenamiento e inclusive los propios pesos de inicialización. Son esos cambios los que Maind define como el “arte” de la construcción de modelos de RNA y la razón por la que se puede considerar que “el entrenamiento más que una técnica, es un análisis consciente para asegurarse que la RNA funcione correctamente” [26].

El vector de pesos resultante de un entrenamiento satisfactorio representa el conocimiento adquirido sobre el problema [20], aunque la interpretación física de este conocimiento adquirido no es fácil de interpretar o no siempre es interpretable.

Cuando finalmente el sistema ha sido correctamente entrenado y no es necesario un aprendizaje adicional, los pesos se pueden “congelar” y no se deberá repetir el entrenamiento cada vez que se desee utilizar el conocimiento aprendido por la RNA para calcular el valor de salida con nuevos datos de entrada.

2.3.2. Validación cruzada

En el caso de una RNA utilizada para realizar pronósticos, se busca que una vez entrenada la red, el conocimiento adquirido por la misma sea útil para realizar pronósticos certeros con datos que no han sido utilizados durante el entrenamiento. Esta capacidad es lo que se conoce como la generalidad del modelo [25]. Un incon-

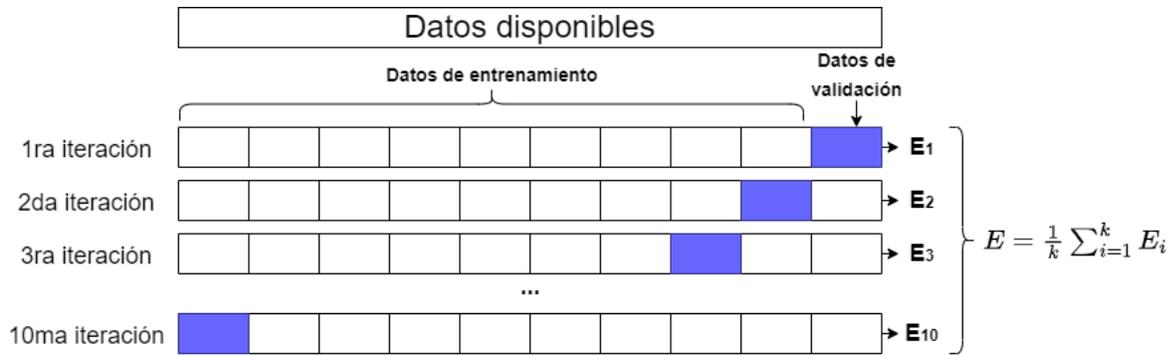


Figura 2.2. Validación cruzada utilizando el método de K-pliegues realizando 10 particiones [27]

veniente presente que es una RNA con múltiples capas ocultas puede ser capaz de memorizar datos [26]. Es por ello que deben definirse métricas robustas que evalúen el rendimiento de la RNA con datos que no hayan sido utilizados durante el entrenamiento y que evalúen correctamente la capacidad de generalizar del modelo.

Un método robusto y utilizado ampliamente para evaluar el rendimiento de un entrenamiento de una RNA es el método de validación cruzada. Este método lo que propone es dividir los datos disponibles una o varias veces para estimar el riesgo de cada algoritmo. Una parte de los datos es utilizado como muestra de entrenamiento y las partes restantes se utilizan para validar los resultados obtenidos por el entrenamiento, evitando así el sobreajuste del modelo [2].

Uno de los métodos de validación cruzada más conocidos es el de “K-pliegues” (K-fold o K-folding en inglés). Este método divide un conjunto de datos en k partes iguales y realiza distintas iteraciones sobre el mismo conjunto de datos pero utilizando k-1 partes para realizar el entrenamiento y la parte restante para validarlo.

El proceso se repite k veces, utilizando una parte distinta de validación para cada iteración y el promedio de los errores obtenidos, será el error esperado para el conjunto de entrenamiento completo, ya que los datos de validación nunca formaron parte del entrenamiento.

3. Definición del problema

3.1. Área de estudio

En el año de 1973, el Instituto Geográfico Nacional de Guatemala -IGN-, realizó una delimitación oficial de las 38 cuencas hidrográficas del país, que corresponden principalmente al punto de desembocadura de la escorrentía superficial que se recolecta en las distintas regiones del país [27]. Estas cuencas se agrupan en 3 grandes vertientes de la siguiente manera: 18 para la Vertiente del Océano Pacífico, 10 para la vertiente del Mar Caribe y 10 para la Vertiente del Golfo de México [38], como se puede observar en la Figura 3.1.

De estas tres vertientes, en el trabajo realizado por Martínez y Pérez [27], se determinó que la Vertiente del Océano Pacífico es la vertiente que cuenta con mayor cobertura de estaciones meteorológicas (más del 50% de la cobertura nacional) e hidrométricas (48% de la cobertura nacional) que proveen información suficiente para realizar un pronóstico de nivel de río utilizando RNA. A pesar de ser la región con mayor densidad de estaciones, en la Vertiente del Océano Pacífico únicamente se consiguió realizar modelos satisfactorios para tres estaciones hidrométricas en dos cuencas hidrográficas de las 18 existentes. El resumen de los resultados obtenidos por dicho trabajo se presenta en el Cuadro 3.1.

Cuadro 3.1. Resultados obtenidos por Martínez y Pérez [27] en “Uso de RNA para la predicción de nivel de agua en ríos en la vertiente del Pacífico de Guatemala”

Estación	Cuenca Hidrográfica	Cantidad de modelos	Mejor rendimiento (MSE)
La gomera	Río Acomé	3	0.0109
Puente Coyolate	Río Coyolate	2	0.0190
El Carrizal	Río Coyolate	2	0.0467

Vertientes y cuencas hidrográficas de Guatemala

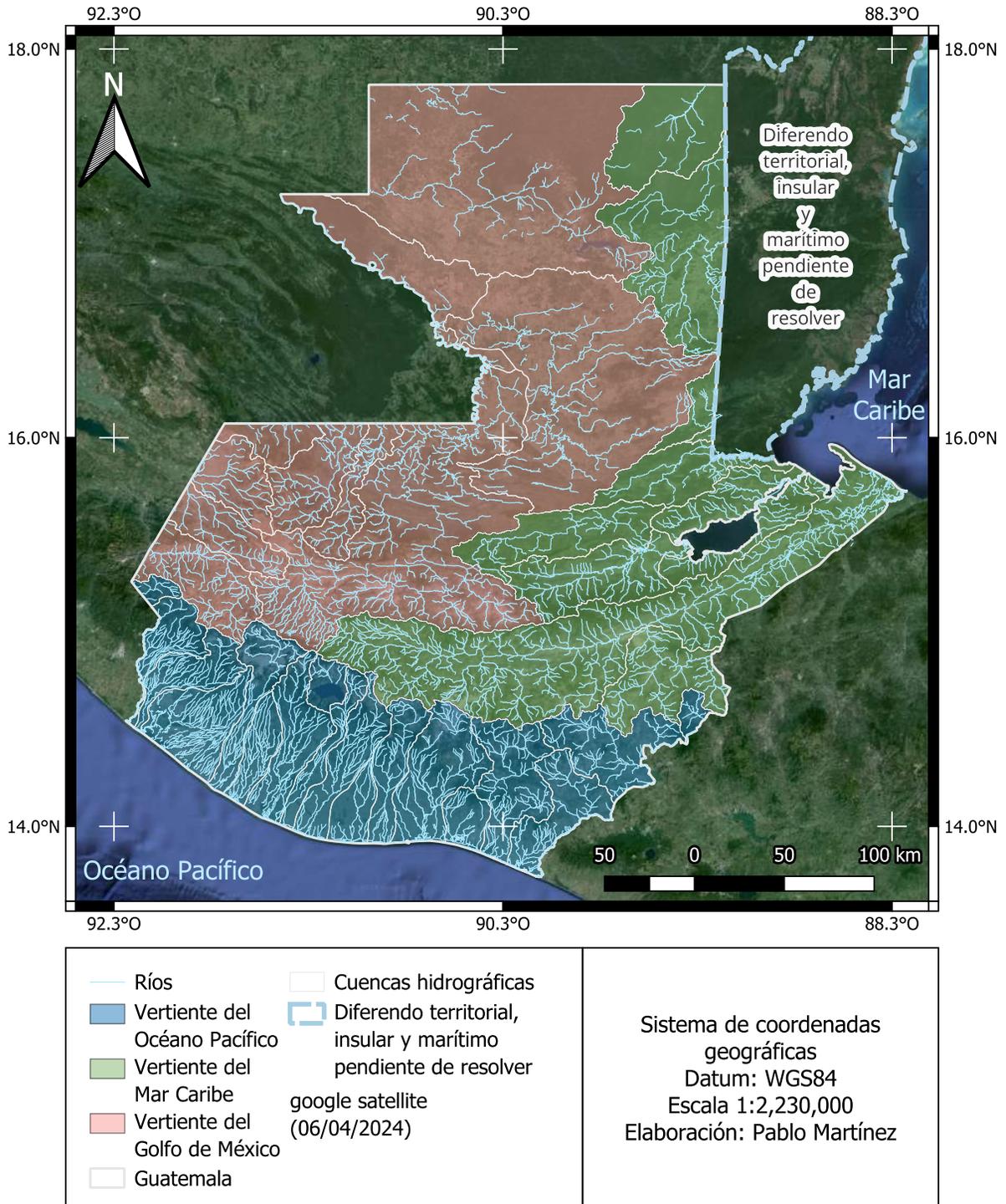


Figura 3.1. Vertientes y cuencas hidrográficas de Guatemala delimitadas por el Instituto Geográfico Nacional -IGN-

Debido a estas limitaciones en cuanto a la información disponible, es necesario contar con nuevas fuentes de información que permitan realizar pronósticos en ríos de la Vertiente del Océano Pacífico y que cubran más de tres estaciones en toda la región.

Tomando como base los resultados obtenidos por Martínez y Pérez [27], se define como primer acercamiento la estación de “La Gomera”, en la cuenca del río Acomé, que al ser la que mejor resultado obtuvo utilizando RNA para realizar pronósticos, el rendimiento de su mejor modelo será la marca a igualar o superar en cuanto a rendimiento para un modelo de RNA utilizando hidroestimadores satelitales.

3.1.1. Cuenca del río Acomé

La cuenca hidrográfica del río Acomé, está situada en jurisdicción del departamento de Escuintla [38]. Esta cuenca se extiende desde la zona conocida como pie de monte volcánico y la parte baja de la cuenca tiene como límite el Océano Pacífico [35], por lo que se encuentra en su totalidad en la Llanura Costera del Pacífico donde tiene altitudes entre los 0 - 625 msnm aproximadamente.

Los cauces de los ríos dentro de esta cuenca que desembocan al río Acomé se caracterizan por desplazarse en pendientes fuertes, por lo que suelen transportar gran cantidad de sedimentos, principalmente de materiales volcánicos [35].

Dentro de la cuenca del río Acomé se encuentra únicamente la estación hidrométrica de “La Gomera” que pertenece al INSIVUMEH. Esta estación cuenta registros digitados desde el año 2011 hasta la fecha. De los 12 años de registros, los primeros 9 tienen registros para las 07 y 17 horas, y a partir del año 2021 hasta la fecha, se cuenta con registros para las 7, 11, 14 y 17 horas. La ubicación de la estación se puede observar en la Figura 3.2 y el resumen de los datos disponibles para dicha estación posterior a la limpieza y procesamiento de la misma se presenta en el cuadro 3.2.

Cuadro 3.2. Información disponible para la estación “La Gomera” según los datos registrados por el INSIVUMEH.

Ventana temporal	Registros por horario			
	07:00	09:00	14:00	17:00
01/02/2011				
31/12/2023	4586	932	932	4586

Cuenca del río Acomé

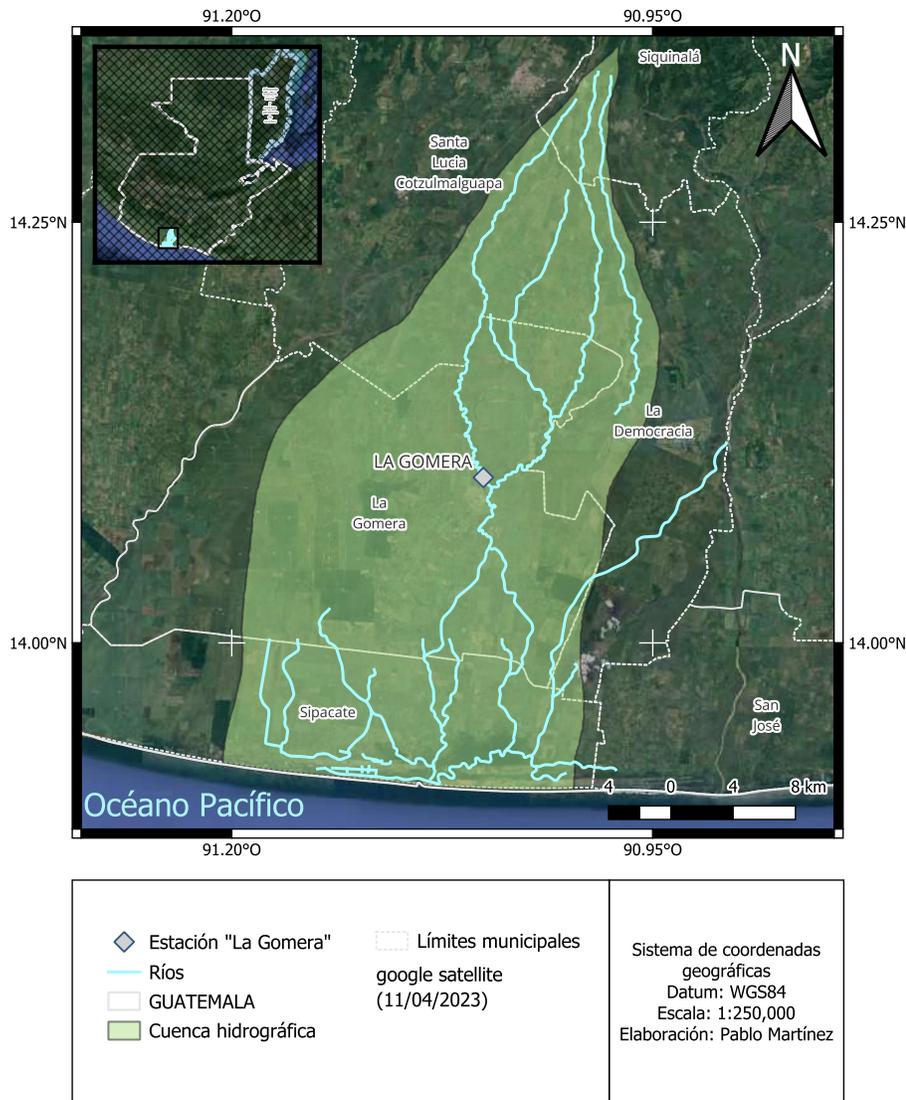


Figura 3.2. Estación "La Gomera" en la cuenca hidrográfica del río Acomé

3.2. Selección de parámetros de entrada

Generalmente se utiliza conocimiento a priori sobre el problema a resolver para realizar una selección de los parámetros de entrada de un modelo de RNA [25], esto permite definir el potencial de las variables que podrán considerarse según la relación más o menos significativa que tengan con la variable de salida e inclusive,

permite descartar en etapas tempranas, variables que no son relevantes para el problema y que causarían inconvenientes al momento de realizar el entrenamiento de una RNA sin aportar información significativa al modelo. El conocimiento sobre las variables relacionadas al problema a resolver se puede complementar con técnicas analíticas para determinar si la variable debería ser considerada o descartada. En el caso de un pronóstico de nivel de agua en ríos se deben considerar las variables del balance hídrico como lo son la “precipitación”, “evapotranspiración”, “infiltración”, entre otras. De estas, la precipitación líquida es una de las variables con mayor influencia en el nivel de agua de un río durante períodos normales, en términos operativos, es la variables que más información tiene disponible tanto en estaciones meteorológicas en la superficie terrestre, como en datos satelitales y, bajo casos puntuales, como los ocurridos durante períodos de lluvia fuerte, se pueden descartar algunas variables del ciclo hidrológico [5] dando mayor prioridad a la precipitación líquida.

Los procesos hidrológicos son un lento proceso acumulativo donde los cambios de nivel de agua están asociados con niveles de agua pasados [6] y tomando como referencia el trabajo de Chen Zeqiang [6], se considerará como variables propuestas, una parte de precipitación que afecta directamente el nivel de río, una parte meteorológica que afecta de forma indirecta y valores de nivel de río previos al nivel a pronosticar como variables de entrada.

3.2.1. Variables de precipitación directa

Los datos satelitales del MSWEP¹ [3], CHIRPS [11] e IMERG [19] son datos rasterizados con distinta resolución espacial, donde un píxel representa la precipitación aproximada en la zona cubierta por el píxel para un horario establecido. Los datos de los tres productos antes mencionados son a escala global, pero se considerarán únicamente aquellos píxeles que pertenecen a la delimitación de la cuenca del río Acomé, como se puede observar en la Figura 3.3.

Ya que cada producto cuenta con distinta resolución, cada uno tendrá distinta cantidad de píxeles disponibles y cada uno de ellos se considerará una variable distinta. Además, si se considera que cada horario tiene n píxeles disponibles por producto, si se tienen j horarios, se tendrán $n * j$ variables distintas por cada producto. Basado en estas consideraciones, se tiene una cantidad distinta de variables

¹Se debe considerar que en el MSWEP se presentan algunos valores de lluvia muy pequeños con ordenes de magnitud $1e-6$ o menores, por lo que todos los valores menores a 0.2 se consideraron como 0

iniciales propuestas para cada producto como se muestra en el cuadro 3.3.

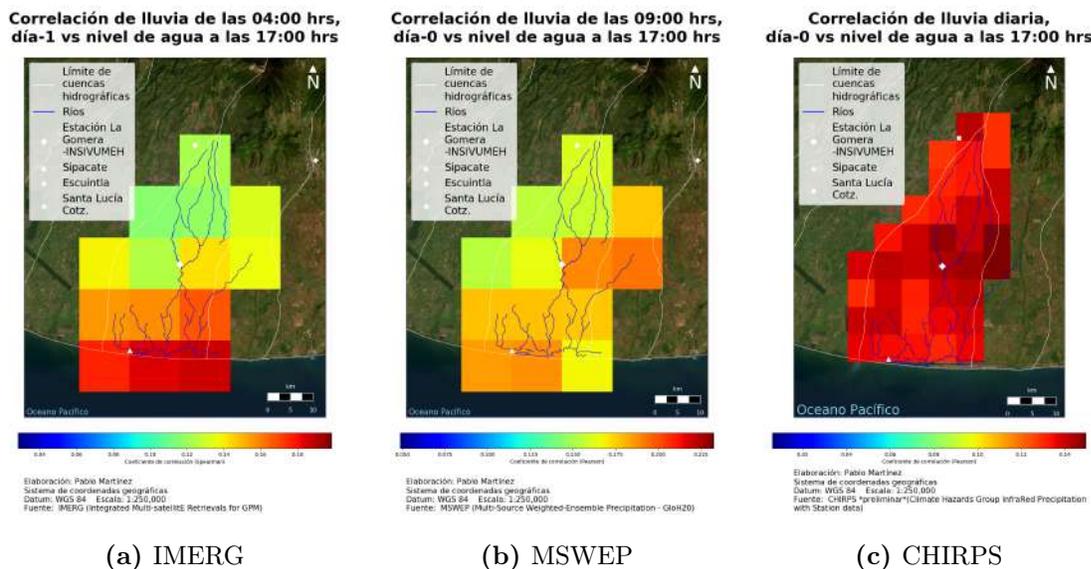


Figura 3.3. Píxeles correspondientes a la cuenca del río Acomé para cada producto satelital

Cuadro 3.3. Cantidad de variables distintas de precipitación por producto

Producto	Número de píxeles	Cantidad de días considerados	Cantidad de horarios por día	Variables totales iniciales
IMERG	14	8	48 día	4,704
MSWEP	14	8	8 día	896
CHIRPS	38	11	1 día	380

Para determinar si un determinado píxel para un determinado horario es una variable relevante para el modelo de RNA a construir, se calculó el coeficiente de Spearman para cada variable respecto al nivel de río diario a las 17:00 hrs. Los coeficientes resultantes para cada producto se pueden observar en las Figuras 3.4, 3.5 y 3.6. Con base a los valores obtenidos, se definió que para considerar una variable, esta debe tener un coeficiente de correlación en magnitud igual o superior a 0.15, ya que se busca mantener aquellas variables que tengan al menos una correlación débil (definida en el Cuadro 2.1) respecto al nivel de río a pronosticar.

De este análisis se descarta CHIRPS como fuente de datos ya que ninguna variable inicial supera en magnitud el umbral establecido de 0.15.

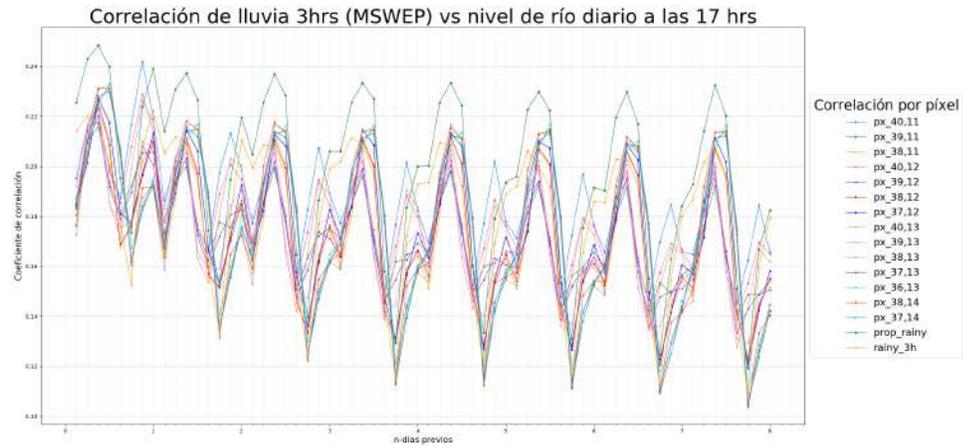


Figura 3.4. Coeficiente de correlación de Spearman para las variables del MSWEP

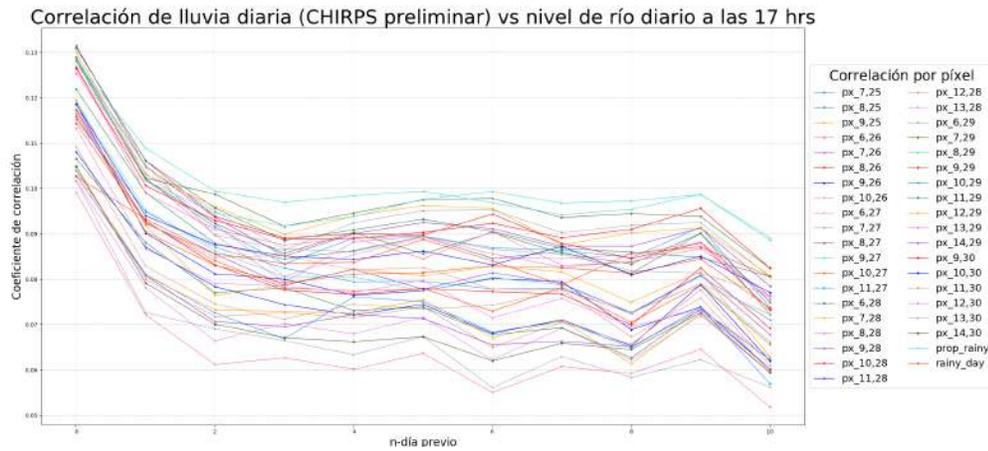


Figura 3.5. Coeficiente de correlación de Spearman para las variables de CHIRPS

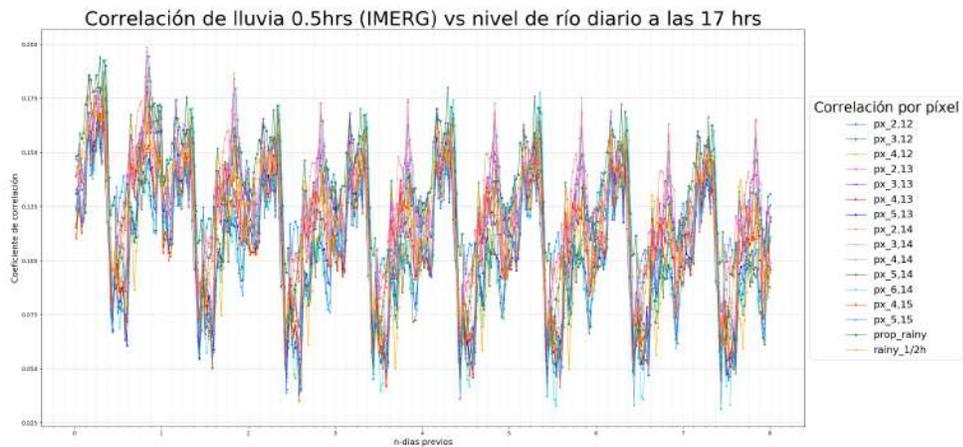


Figura 3.6. Coeficiente de correlación de Spearman para las variables del IMERG

Además, observando el comportamiento de los coeficientes de correlación para el MSWEP y el IMERG, se puede observar un comportamiento “recurrente” en un alto coeficiente de correlación en horarios similares a las 17:00 hrs en días previos al día a pronosticar, así como una disminución de correlación en horas de la mañana (Figuras 3.4 y 3.6). Es por ello que, partiendo de dicho comportamiento, se definen las siguientes consideraciones:

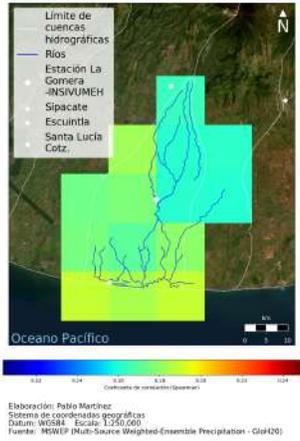
1. Para el MSWEP se considerarán únicamente aquellas variables que igualen o superen en magnitud el umbral de 0.15 de coeficiente de correlación y que cumplan con las siguientes condiciones:
 - Para el día actual de pronóstico y el día previo (D-1), se mantendrá la totalidad de variables que superen el umbral establecido.
 - Para el segundo y tercer día previo (D-2 y D-3 respectivamente), se considerarán únicamente los horarios entre las 9:00 y 18:00 horas, ya que son los horarios que presentan una correlación significativa en los píxeles de la parte alta de la cuenca (Figuras 3.7d a la 3.7g), ya que se da prioridad a aquellas zonas que se encuentren en igual o superior latitud respecto a la estación hidrométrica sobre la cual se realizará el pronóstico, según el sentido hidrológico [27].
 - Del cuarto día previo en adelante se descartarán todas las variables.
2. Para el IMERG, debido a tener en general valores más pequeños de coeficiente de correlación, se considerarán aquellas variables que igualen o superen en magnitud el umbral de 0.15 de coeficiente de correlación para el día actual de pronóstico y hasta el tercer día previo (D-0 hasta D-3, respectivamente).

De estas consideraciones se descarta una cantidad considerable de variables iniciales para el MSWEP y el IMERG, quedando únicamente aquellas que tienen una mayor relevancia y en cantidades como se muestra en el Cuadro 3.4.

Cuadro 3.4. Variables meteorológicas directas relevantes

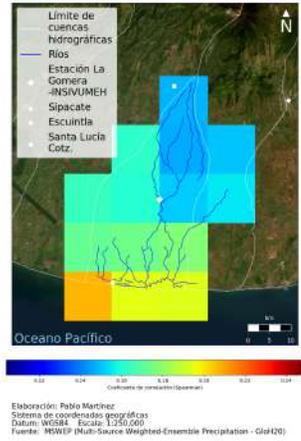
Producto	Variables relevantes
IMERG	280
MSWEP	261
CHIRPS	0

Correlación de lluvia de las 00:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



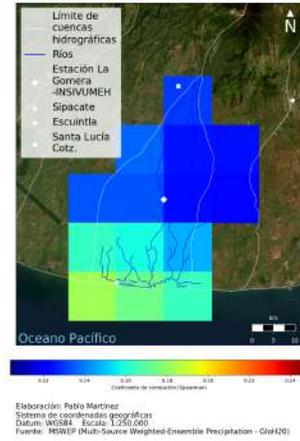
(a) 00 hrs D-2

Correlación de lluvia de las 03:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



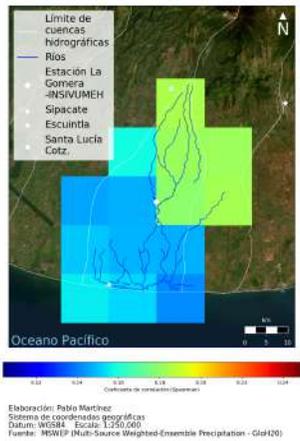
(b) 03 hrs D-2

Correlación de lluvia de las 06:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



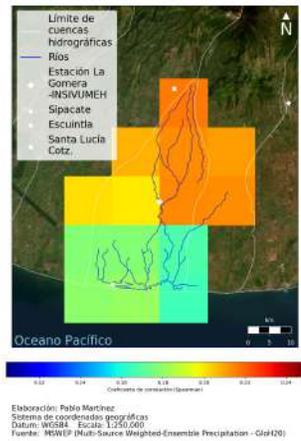
(c) 06 hrs D-2

Correlación de lluvia de las 09:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



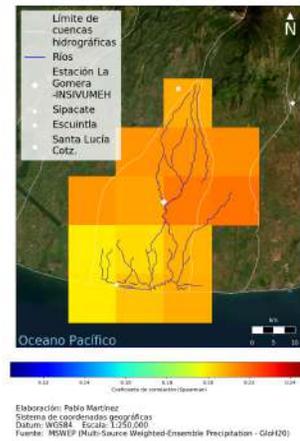
(d) 09 hrs D-2

Correlación de lluvia de las 12:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



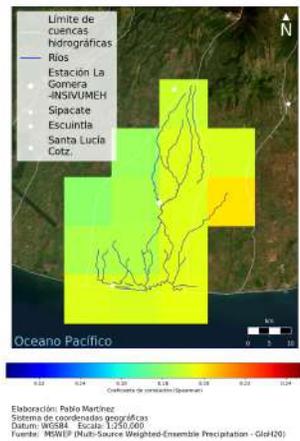
(e) 12 hrs D-2

Correlación de lluvia de las 15:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



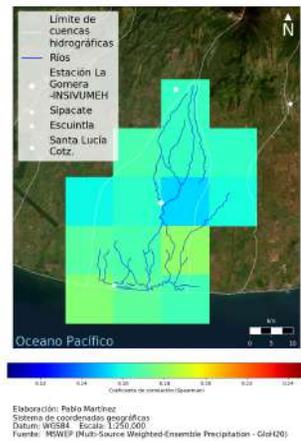
(f) 15 hrs D-2

Correlación de lluvia de las 18:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



(g) 18 hrs D-2

Correlación de lluvia de las 21:00 hrs, día-2 vs nivel de agua a las 17:00 hrs



(h) 21 hrs D-2

Figura 3.7. Correlación por píxel para distintos horarios del segundo día previo al día a pronosticar

3.2.2. Variables meteorológicas indirectas

Con el fin de aportar información extra a los modelos de RNA que complementen la relación directa que existe entre la lluvia y el nivel de agua en un río, se definieron dos tipos de variables que buscan proveer información extra sobre el día a pronosticar. El primer tipo de variables posee la misma resolución temporal que los datos horarios ya que se calcula a partir de los mismos y el segundo tipo tendrá una resolución mensual y se calcula mediante el apilado de información previa registrada.

3.2.2.1. Variables de resolución horaria

Para el caso de variables de resolución horaria, se crearon dos variables. La primera, denominada como “**prop_rainy**”, se calcula como la suma de todos los píxeles que registraron lluvia en el período de tiempo del producto satelital en la cuenca considerada. Para que se considere que un píxel registró lluvia, el valor de píxel debe superar el umbral de los 0.2mm y se toma como referencia mínima ya que es un valor que algunos pluviómetros en superficie tienen como valor mínimo de medida [36] [39]. En caso de no registrarse lluvia durante el período de tiempo considerado, el valor de la variable será asignado como 0 y su valor máximo será el número n de píxeles disponibles para el producto satelital utilizado. Esta variable se puede interpretar como la cantidad de área de la cuenca donde se registró lluvia en la resolución horaria considerada.

La segunda variable, denominada como “**rainy_resmin**”, se calcula como un valor 0 si no se registró lluvia en ninguno de los píxeles de la cuenca considerada, durante la resolución horaria de la información y como un 1 si al menos un píxel registro lluvia. Esta variable se puede interpretar como la presencia o ausencia de lluvia en la cuenca durante la resolución horaria considerada, sin importar que tanta área fue cubierta.

Considerando que cada producto satelital interpreta de distinta manera los eventos de lluvia según las resolución espacial y temporal del mismo, se obtiene una distinta cantidad de variables iniciales propuestas para cada producto. Además, realizando una discriminación similar a la realizada en la Sección 3.2.1, se consideran como relevantes solo aquellas variables que presenten un valor igual o superior en magnitud al umbral de 0.15 de coeficiente de correlación de Spearman, respecto al nivel de río diario a las 17:00 hrs. Basado en estas consideraciones, se tiene una cantidad distinta de variables meteorológicas de resolución horaria distintas para

cada producto como se muestra en el cuadro 3.5

Cuadro 3.5. Cantidad de variables meteorológicas horarias indirectas distintas de precipitación por producto

Producto	Variables horarias consideradas	Cantidad de días considerados	Cantidad de horarios por día	Variables totales iniciales	Variables relevantes
IMERG	2	8	48 día	768	60
MSWEP	2	8	8 día	128	38

3.2.2.2. Variables de resolución mensual

Para el caso de variables de resolución mensual, se define una ventana temporal de los primeros 25 meses calendario previos al mes actual del día donde se realizará el pronóstico diario y se crearon dos variables distintas. La primera de ellas es el promedio mensual de todos los datos satelitales registrados para el mes correspondiente sobre todos los píxeles de la cuenca considerada. La segunda variable, es el valor máximo registrado de lluvia para todo el mes correspondiente. De esta consideración, se obtienen inicialmente 50 variables para ambos productos satelitales, pero al realizar una discriminación similar a la realizada en la Sección 3.2.1, se consideran como relevantes solo aquellas variables que presenten un valor igual o superior en magnitud al umbral de 0.15 de coeficiente de correlación de Spearman, respecto al nivel de río diario a las 17:00 hrs. Basado en estas consideraciones, se tiene una cantidad distinta de variables meteorológicas mensuales indirectas propuestas para cada producto como se muestra en el cuadro 3.6.

Cuadro 3.6. Cantidad de variables meteorológicas mensuales indirectas distintas de precipitación por producto

Producto	Número de meses previos considerados	Variables mensuales utilizadas	Variables totales iniciales	Variables relevantes
IMERG	24	Media, Máximo	50	23
MSWEP	24	Media, Máximo	50	26

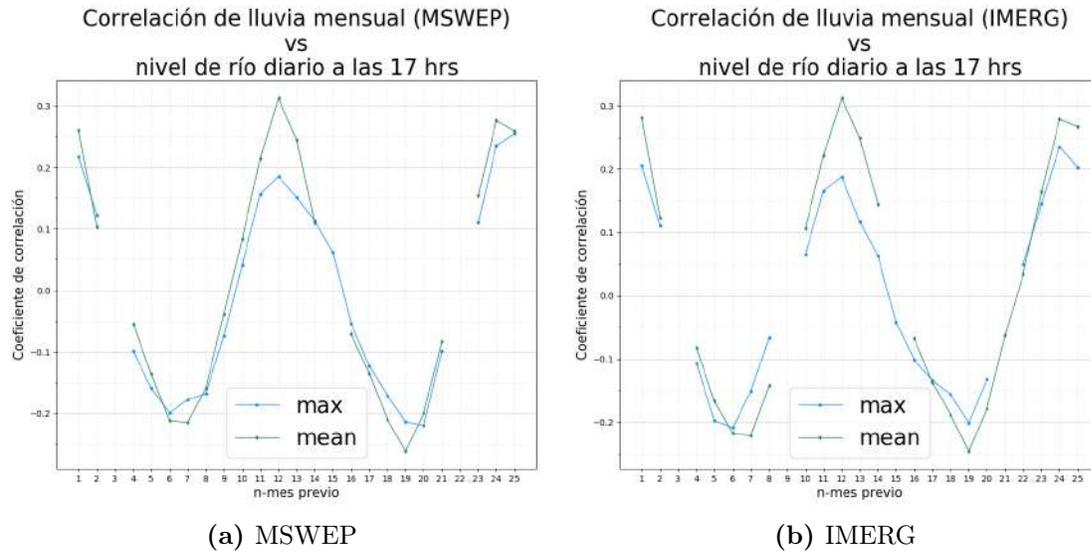


Figura 3.8. Comportamiento de la correlación de las variables meteorológicas indirectas mensuales de precipitación

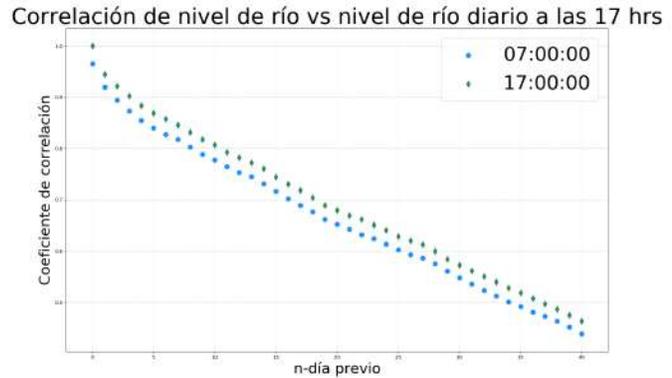
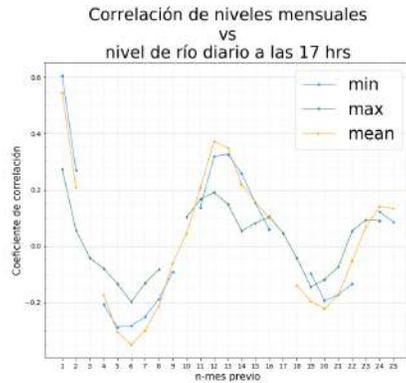
3.2.3. Variables de nivel de río

Para el caso de los datos de nivel de agua en el río registrados por el INSIVU-MEH, considerando que se deben tener en cuenta niveles de agua previos al nivel a pronosticar [6], se consideran de manera preliminar 40 días previos de nivel de río registrados en la estación a considerar. Posterior a ello se analiza el coeficiente de correlación de Spearman para cada variable de los registros de las 07:00 y 17:00 hrs y se considerarán como relevantes solo aquellas que superen una magnitud de 0.60 de coeficiente de correlación y para facilitar la implementación en caso de ser requerido, los registros posteriores al día previo (D-10), serán descartados en su totalidad.

Además, se consideran también variables mensuales que podrían brindar información sobre la época del año en la que se encuentra el pronóstico a realizar y brinde una referencia sobre el comportamiento esperado para dicha época. Para esto, se define una ventana temporal de los primeros 25 meses calendario previos al mes actual del día donde se realizará el pronóstico diario y se calculo el mínimo, media y máximo mensual para los 24 meses previos. De esta forma, realizando una discriminación similar a la realizada en la Sección 3.2.1, se considera como relevantes solo aquellas variables que presenten un valor igual o superior al umbral de 0.15 de coeficiente de correlación de Spearman, respecto al nivel de río diario a las 17:00 hrs. Basado en estas consideraciones se obtienen las variables relevantes presentadas en el cuadro 3.7.

Cuadro 3.7. Cantidad de variables de nivel de río relevantes

Estación hidrométrica	Número de meses previos considerados	Variables mensuales utilizadas	Registros diarios utilizados	Variables mensuales iniciales	Variables mensuales relevantes
La Gomera	24	Mínimo, Media, Máximo	21	72	32



(a) Correlación de las variables mensuales de nivel de río

(b) Correlación de los registros diarios de nivel de agua

Figura 3.9. Comportamiento de la correlación de las variables mensuales (izquierda) y registros diarios (derecha) de nivel de río.

3.2.4. Variables de entrada

Basado en la discriminación y selección de variables por fuente de información y considerando que las variables de nivel de río se deben considerar siempre, se obtiene un listado de variables a ser utilizadas como vector de entrada para un modelo de RNA. Se define un modelo para cada producto satelital con el cual se observará el rendimiento de cada uno por separado de forma inicial.

Realizando la unificación de todas las variables disponibles para cada producto satelital se obtienen los vectores de entrada por producto como se muestra en el cuadro 3.8

Cuadro 3.8. Vector de entrada para un modelo de RNA por producto satelital

PRODUCTO	RIO		METEOROLOGICAS			TOTAL
	DIARIO	MENSUAL	DIRECTAS	INDIRECTAS HORARIAS	INDIRECTAS MENSUALES	
IMERG	21	32	280	60	23	416
MSWEP	21	32	261	38	26	378

Los resultados de Nachimuthu Karunanithi [20] presentan que no fue necesario transformar o suavizar los datos de entrada y en el trabajo presentado por Holger Maier [25], la mayoría de artículos relacionados tampoco presentan una clara metodología para normalizar los datos a utilizar. Es por ello que más allá de la unificación de variables en una única base de datos, no se realizará otra modificación a los valores de entrada para cada producto satelital.

4. Entrenamiento de modelos de RNA

Como se discutió en la Sección 2.1.1, el esquema de una RNA es dependiente del problema y aunque se han presentado distintas aproximaciones en todo el mundo sobre como realizar un pronóstico de nivel de agua en ríos utilizando RNA [25], cada modelo está sujeto a consideraciones como la cantidad de estaciones meteorológicas o datos satelitales disponibles, la resolución temporal de los mismos, la morfología del río así como las condiciones climáticas de la zona, etc. Por lo que a continuación, se definirán distintas variaciones de parámetros que afectan en mayor o menor medida el rendimiento de un modelo de RNA para así caracterizar que combinación de parámetros podrían ser utilizados en la vertiente del Océano Pacífico de Guatemala para obtener un modelo con un rendimiento deseado.

Los parámetros que se variarán para el esquema de la RNA son los siguientes:

1. Geometría de la RNA y ventana temporal de los datos utilizados
2. Funciones de costo
3. Cantidad de datos previos de nivel de río

Antes de comenzar a ajustar los parámetros del modelo, es fundamental establecer un punto de partida con aquellos parámetros que no se estarán evaluando inicialmente. Para esto, se tomará en consideración el trabajo de Martínez y Pérez [27] en la Vertiente del Océano Pacífico, donde se utilizaron los siguientes parámetros para los modelos de RNA:

- Rango de épocas: Desde 10 hasta 400 a intervalos de 10 en 10
- Optimizador: Adam
- Tasa de aprendizaje del optimizador: 0.0001
- Función de costo: “mean_squared_error”
- Cantidad de neuronas por capa: Todas las capas ocultas tienen la misma cantidad de neuronas correspondiente a la longitud del vector de entrada n .

4.1. Variaciones de geometría y ventana temporal

El esquema del modelo de RNA se construirá considerando que se está pronosticando un único valor de nivel de agua correspondiente a las 17:00 hrs. para ello se fija una capa de salida con una única neurona, una capa de entrada con n neuronas, donde n es la longitud del vector de entrada según el producto satelital y siempre contará con al menos una capa oculta según las consideraciones descritas en la Sección 2.1.1. La variación de cantidad de capas ocultas se realizará aumentando desde 2 hasta 7 el número de capas ocultas de “ n ” neuronas cada una y se evaluará el rendimiento del modelo utilizando el método de validación cruzada de K-folding.

Considerando que el método de K-folding divide un conjunto de datos en K partes iguales, al dividir una ventana temporal de N años en K partes, se tendrá una ventana temporal por iteración de $N(K-1)/K$ años, donde al disminuir el número de folios, la ventana temporal por iteración irá disminuyendo, lo que representa una menor cantidad de datos para realizar el entrenamiento de la RNA. Bajo esta consideración, la variación de la ventana temporal se realizará variando el número de folios en 2, 3, 5 y 10, para observar el rendimiento del error en un entrenamiento al contar con una menor o mayor cantidad de datos.

Para la cuenca del río Acomé, al unir los datos de nivel de río con los datos satelitales (considerando que se eleminarán todos aquellos días donde no se realizó un registro de nivel de río o dato satelital), se tiene una ventana temporal de 3393 días (aproximadamente 9.3 años) para ambos productos. Para cada uno se separa el año 2023, ya que este servirá como validación final y no formará parte del proceso de k-folding, para así evaluar como se hubiera comportando un modelo en la realidad si este se hubiera implementado a principios del año seleccionado. Removiendo los 280 días disponibles para el 2023, se obtienen únicamente 3113 datos máximos disponibles para el entrenamiento (aproximadamente 8.53 años).

Una vez establecidas las bases de datos a utilizar para el entrenamiento se realizan las variaciones en el siguiente orden:

1. Número de folios [2-5]
2. Número de capas ocultas¹ [2-7]
3. Número de épocas²³ [10-400]

¹Para el caso del IMERG, se añaden 10 y 25 capas ocultas al análisis para observar el rendimiento del mismo con una mucho mayor cantidad de capas ocultas

²La iteración de épocas se realiza en intervalos de 10 en 10

³Para el caso del IMERG el rango va desde 10 - 450 para obtener mejores resultados

Los resultados obtenidos de las variaciones antes descritas se pueden observar en el cuadro 4.1 y 4.2.

Cuadro 4.1. Rendimiento de distintos modelos de RNA para datos del MSWEP al variar la ventana temporal de entrenamiento y el número de capas ocultas por iteración.

Número de folios	Años disponibles para entrenamiento	Cantidad de capas ocultas	MSE esperado mejor modelo	Error estándar (metros)	Mejor número de épocas	Tiempo de entrenamiento (segundos)
2	4.3 (1557 días)	2	0.0215	0.15	350	4734 (1.3 hrs)
		3	0.0159	0.13	370	3725 (1.0 hrs)
		4	0.0141	0.12	380	4131 (1.1 hrs)
		5	0.0135	0.12	370	4786 (1.3 hrs)
		6	0.0119	0.11	310	5376 (1.5 hrs)
		7	0.0114	0.11	370	6634 (1.8 hrs)
3	5.7 (2075 días)	2	0.0205	0.14	400	9554 (2.7 hrs)
		3	0.0131	0.11	370	7750 (2.2 hrs)
		4	0.0117	0.11	380	8596 (2.4 hrs)
		5	0.0113	0.11	390	9893 (2.7 hrs)
		6	0.0108	0.10	300	11065 (3.1 hrs)
5	6.8 (2490 días)	2	0.0188	0.14	370	18817 (5.2 hrs)
		3	0.0123	0.11	400	15401 (4.3 hrs)
		4	0.0107	0.10	370	17074 (4.7 hrs)
		5	0.0104	0.10	370	19361 (5.4 hrs)
		6	0.0099	0.10	370	22184 (6.2 hrs)
		7	0.0097	0.10	380	26320 (7.3 hrs)
10	7.7 (2802)	2	0.0161	0.13	380	34333 (9.4 hrs)
		3	0.0118	0.11	370	39973 (11.1 hrs)
		4	0.0109	0.10	330	46461 (12.9 hrs)
		5	0.0100	0.10	350	40175 (11.2 hrs)
		6	0.0095	0.10	350	39309 (10.9 hrs)
		7	0.0092	0.10	350	43605 (12.1 hrs)

De los resultados obtenidos al variar el número de folios y la cantidad de capas ocultas de la RNA se pueden observar las siguientes características:

- Al aumentar el número de folios aumenta el tiempo de entrenamiento para realizar una inspección en el rango de épocas consideradas.
- Al aumentar el número de capas ocultas aumenta el tiempo de entrenamiento (Figura 4.1d y 4.1b).
- Al aumentar el número de folios y capas ocultas se consigue un menor MSE, indicando un mejor rendimiento del mejor modelo seleccionado para una combinación específica de folios, capas ocultas y época.

Cuadro 4.2. Rendimiento de distintos modelos de RNA para datos del IMERG al variar la ventana temporal de entrenamiento y el número de capas ocultas por iteración.

Número de folios	Años disponibles para entrenamiento	Cantidad de capas ocultas	MSE esperado mejor modelo	Error estándar (metros)	Mejor número de épocas	Tiempo de entrenamiento (segundos)
5	6.8 (2490 días)	2	0.01867	0.14	440	18835 (5.2 hrs)
		3	0.01346	0.12	450	19271 (5.4 hrs)
		4	0.0124	0.11	370	24297 (6.7 hrs)
		5	0.01203	0.11	370	28020 (7.8 hrs)
		6	0.01169	0.11	380	27159 (7.5 hrs)
		7	0.01121	0.11	380	28668 (8.0 hrs)
		10	0.01082	0.10	430	37829 (10.5 hrs)
10	7.7 (2802 días)	25	0.01089	0.10	450	75564 (21.0 hrs)
		2	0.0172	0.13	440	47526 (13.2 hrs)
		3	0.01268	0.11	440	45206 (12.6 hrs)
		4	0.01135	0.11	440	58970 (16.4 hrs)
		5	0.01108	0.11	440	62721 (17.4 hrs)
		6	0.01059	0.10	350	72691 (20.2 hrs)
		7	0.01067	0.10	330	71520 (19.9 hrs)

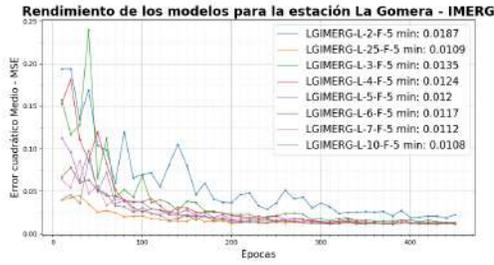
- El error estándar (RMSE) disminuye al tener una mayor cantidad de datos disponibles y un mayor número de capas ocultas, pero en un punto dicha disminución ya no es significativa, ya que su variación es menor que la magnitud de medición del nivel de agua en ríos ⁴ (Figura 4.1c y 4.1a).

Considerando que al aumentar el número de folios y capas ocultas mejora el rendimiento, pero también aumenta considerablemente el tiempo de entrenamiento, se selecciona una combinación ideal que permita realizar variaciones de funciones de error con una buena precisión pero que no extienda demasiado el tiempo de ejecución necesario para evaluar distintas funciones de error. Es por ello que se determina como ideal las siguientes combinaciones para cada producto satelital:

Cuadro 4.3. Combinación ideal de número de folios y capas ocultas, como modelo base para una RNA por producto satelital

Producto satelital	Número de folios	Cantidad de capas ocultas
IMERG	5	6
MSWEP	5	5

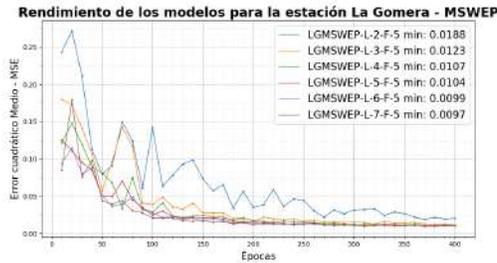
⁴La medición de nivel de agua se realiza en centímetros por lo que variaciones de milímetros no deben ser consideradas



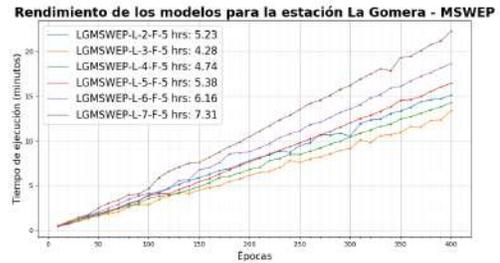
(a) Rendimiento del error por épocas - IMERG



(b) Tiempo de entrenamiento por épocas - IMERG



(c) Rendimiento del error por épocas - MSWEP



(d) Tiempo de entrenamiento por épocas - MSWEP

Figura 4.1. Rendimiento del error y tiempo de entrenamiento por época para distintas cantidades de capas ocultas utilizando 5 folios con datos del MSWEP y el IMERG

4.2. Variaciones de función de error

Como presenta Holger Maier [25], la función de error más comúnmente utilizada es el MSE debido a las ventajas que este presenta (discusión en la sección 2.2.5), pero al buscar obtener mejores resultados con los datos disponibles para Guatemala se evaluará si alguna función de error distinta como lo es el MAE, MaxAE o la Función de pérdida de Huber presentan un mejor rendimiento que el MSE. Para ello se utilizará la estructura recomendada para cada producto satelital definida en el cuadro 4.3.

Cabe resaltar que a diferencia de los resultados obtenidos en el cuadro 4.1 y 4.2, al utilizarse distintas funciones de error, el valor de la métrica obtenida para cada variación no es comparable con las otras por lo que de cada función de error deberá elegirse el mejor resultado para realizar un entrenamiento final y con el cual pueda realizarse el pronóstico para los datos de entrenamientos totales y los datos de validación (año 2023) con el cual, mediante distintas métricas de error se pueda analizar las fortalezas y debilidades de cada modelo para determinar que métrica es la que brinda un mejor rendimiento para el río seleccionado.

De las consideraciones previas y al realizar el entrenamiento de distintos modelos de RNA para cada uno de las funciones de pérdida consideradas se obtuvieron los resultados que se pueden observar en el cuadro 4.4.

Cuadro 4.4. Rendimiento de distintos modelos de RNA al variar la función de pérdida del modelo.

Producto satelital	Función de pérdida	Valor mínimo de la métrica obtenida	Época con mejor resultado	Nombre modelo
IMERG	Mean Squared Error	0.0117	380	LG-IMERG-MSE
	Mean Absolute Error	0.0374	390	LG-IMERG-MAE
	Max Absolute Error	0.3347	340	LG-IMERG-MAX
	Huber Loss 1.0	0.0045	310	LG-IMERG-HB1
	Huber Loss 0.5	0.0036	310	LG-IMERG-HB05
	Huber Loss 0.1	0.0018	350	LG-IMERG-HB01
MSWEP	Mean Squared Error	0.0107	370	LG-MSWEP-MSE
	Mean Absolute Error	0.0414	390	LG-MSWEP-MAE
	Max Absolute Error	0.3279	400	LG-MSWEP-MAX
	Huber Loss 1.0	0.0046	370	LG-MSWEP-HB1
	Huber Loss 0.5	0.0038	380	LG-MSWEP-HB05
	Huber Loss 0.1	0.0018	370	LG-MSWEP-HB01

4.2.1. Evaluación de modelos

El uso de una métrica de evaluación como lo es el MSE, describe de forma resumida en un único valor el rendimiento de distintos modelos al momento de su evaluación, pero no es capaz de brindar información relevante en distintos aspectos que también son de interés al momento de realizar un pronóstico. Es por ello que, considerando que un pronóstico de nivel de río puede ser útil para la gestión de riesgo en caso de inundaciones, se evaluará el error estándar (RMSE) del modelo como un tipo de incerteza sobre el valor pronosticado, el sesgo porcentual como un indicador de la tendencia a sobreestimar o subestimar del modelo, el NSE como indicador de si el modelo es útil o debe descartarse ya que es menos útil que el promedio de todos los datos, el MaxAE para conocer el máximo error obtenido hasta la fecha, la Sensibilidad y el Valor Predictivo Negativo, para saber que tan útil sería el modelo para emitir alertas de crecida o disminución de nivel de agua en el río y el Coeficiente

R^2 para observar que tan bien se ajustan los valores pronosticados respecto a los registros diarios, ya que de ser un pronóstico perfecto, el comportamiento de uno respecto al otro debería ser una relación lineal perfecta.

Es importante considerar que las distintas métricas son comparables únicamente entre si misma por lo que la comparación de cada métrica se realizará únicamente sobre la misma métrica para los distintos modelos realizados.

Para evaluar el rendimiento de los modelos presentados en el Cuadro 4.4, se realizó el entrenamiento de una RNA con todos los datos disponibles (2012-2022) con el número de épocas que presento el mejor resultado. Una vez obtenido cada modelo, se pasaron los datos de entrenamiento y validación por la RNA entrenada para obtener los pronósticos realizados diariamente para compararlo con los registros reales registrados. Los resultados para ambos productos satelitales se muestran en el Cuadro 4.5.

Cuadro 4.5. Métricas de evaluación para cada modelo del IMERG y el MSWEP para los datos de entrenamiento (2012-2022) y validación (2023)

		MODELO	PBias	MSE	RMSE	N	RMSE	MAE	MaxAE	NSE	R ²	P-val
IMERG	Entr.	LG-IMERG-MSE	15.1544	0.0067	0.08	81.6%	0.06	1.74	0.7801	0.9109	0	
		LG-IMERG-MAE	4.0375	0.0058	0.08	93.8%	0.03	2.28	0.8100	0.9017	0	
		LG-IMERG-MAX	-8.2502	0.0079	0.09	81.7%	0.06	2.26	0.7411	0.8672	0	
		LG-IMERG-HB1	3.1750	0.0054	0.07	88.0%	0.04	2.11	0.8206	0.9072	0	
		LG-IMERG-HB05	-3.4968	0.0058	0.08	89.0%	0.04	2.28	0.8088	0.9007	0	
	Valid.	LG-IMERG-HB01	2.0199	0.0058	0.08	92.2%	0.03	2.40	0.8077	0.8994	0	
		LG-IMERG-MSE	16.2962	0.0156	0.12	88.9%	0.06	1.18	-2.1561	0.4935	0	
		LG-IMERG-MAE	1.8327	0.0018	0.04	80.7%	0.03	0.24	0.6416	0.8712	0	
		LG-IMERG-MAX	-12.5447	0.0099	0.10	85.0%	0.06	0.62	-1.0047	0.5786	0	
		LG-IMERG-HB1	4.1502	0.0090	0.09	87.1%	0.05	0.90	-0.8218	0.6193	0	
MSWEP	Entr.	LG-IMERG-HB05	-3.4051	0.0078	0.09	88.9%	0.04	0.87	-0.5835	0.6175	0	
		LG-IMERG-HB01	2.3327	0.0037	0.06	85.0%	0.04	0.36	0.2495	0.7797	0	
		LG-MSWEP-MSE	9.2874	0.0068	0.08	84.9%	0.05	2.10	0.7750	0.8930	0	
		LG-MSWEP-MAE	4.8580	0.0067	0.08	94.2%	0.04	2.37	0.7787	0.8849	0	
		LG-MSWEP-MAX	-15.8617	0.0112	0.11	82.0%	0.07	2.29	0.6304	0.8235	0	
	Valid.	LG-MSWEP-HB1	4.6614	0.0068	0.08	86.5%	0.04	2.21	0.7762	0.8920	0	
		LG-MSWEP-HB05	12.8683	0.0067	0.08	87.3%	0.05	2.31	0.7796	0.8999	0	
		LG-MSWEP-HB01	6.7520	0.0064	0.08	91.6%	0.03	2.32	0.7877	0.8938	0	
		LG-MSWEP-MSE	29.2105	0.0050	0.07	70.0%	0.06	0.26	-0.0029	0.7853	0	
		LG-MSWEP-MAE	6.6343	0.0021	0.05	64.6%	0.04	0.20	0.5801	0.7890	0	
Valid.	LG-MSWEP-MAX	-3.7708	0.0062	0.08	70.4%	0.06	0.43	-0.2474	0.6848	0		
	LG-MSWEP-HB1	13.0524	0.0026	0.05	78.6%	0.04	0.26	0.4726	0.8399	0		
	LG-MSWEP-HB05	26.5460	0.0036	0.06	75.4%	0.05	0.33	0.2686	0.8530	0		
	LG-MSWEP-HB01	10.9474	0.0017	0.04	79.6%	0.03	0.24	0.6628	0.9009	0		

Al obtener las métricas de evaluación para los modelos a evaluar, se definen las siguientes consideraciones para descartar modelos y seleccionar el modelo con mejor rendimiento para cada producto satelital:

1. Buscando consistencia entre validación y entrenamiento, si en alguna de las dos ventanas el NSE es negativo, el modelo queda descartado automáticamente.

- Se considera que el resto de métricas son igualmente relevante y se asignará una puntuación de 3, 2 y 1 puntos al primer, segundo y tercer lugar respectivamente para cada métrica evaluada.

Bajo dichas consideraciones, se establece que para el IMERG, el mejor rendimiento se obtiene utilizando el MAE y para el MSWEP se obtiene utilizando Hubber Loss con $\alpha = 0.1$. Para los mejores modelos resultantes se calcula la información sobre el rendimiento de los mismos para la emisión de alertas por crecida o disminución de nivel de agua en el río. Para esto se construye una condición de aumento o disminución de nivel de río comparando el día actual con el día previo. Si el nivel de río del día actual es mayor que el día anterior, se clasifica como un aumento de nivel de río. Si el nivel de río es igual o menor al día anterior, se clasifica como una disminución de nivel de río, por lo que al únicamente tener dos clasificaciones posibles, es posible construir una clasificación binaria para obtener una matriz de confusión para cada conjunto de datos. Esto se puede observar para los modelos finales de cada producto satelital en el Cuadro 4.6 y 5.5.

	P	N
T	772	851
F	276	1214

(a) Entrenamiento (2012-2022)

	P	N
T	50	76
F	31	123

(b) Validación (2023)

	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	64 %	48 %	74 %	59 %	81 %
Valid.	62 %	40 %	62 %	62 %	80 %

(c) Rendimiento del modelo para la emisión de alertas

Cuadro 4.6. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “LG-IMERG-MAE”.

	P	N
T	716	878
F	332	1187

(a) Entrenamiento (2012-2022)

	P	N
T	49	75
F	32	124

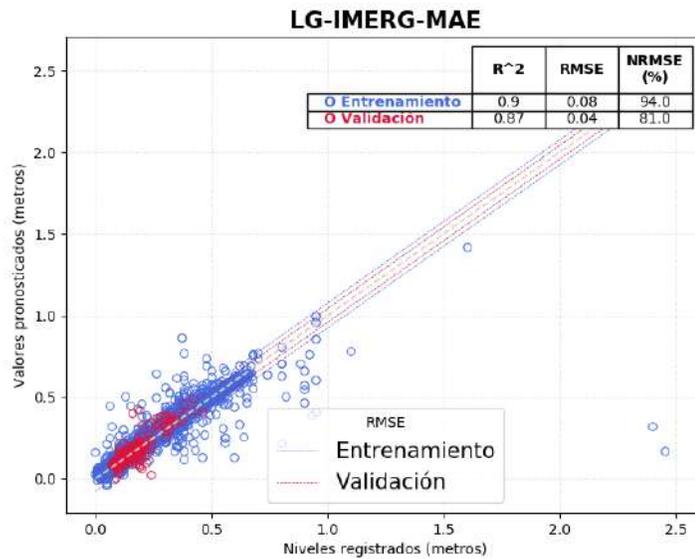
(b) Validación (2023)

	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	61 %	45 %	68 %	57 %	78 %
Valid.	62 %	40 %	60 %	62 %	79 %

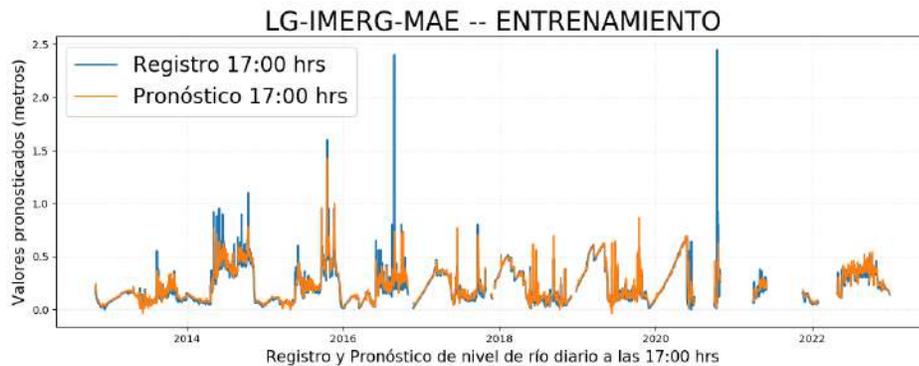
(c) Rendimiento del modelo para la emisión de alertas

Cuadro 4.7. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “LG-MSWEP-HB01”.

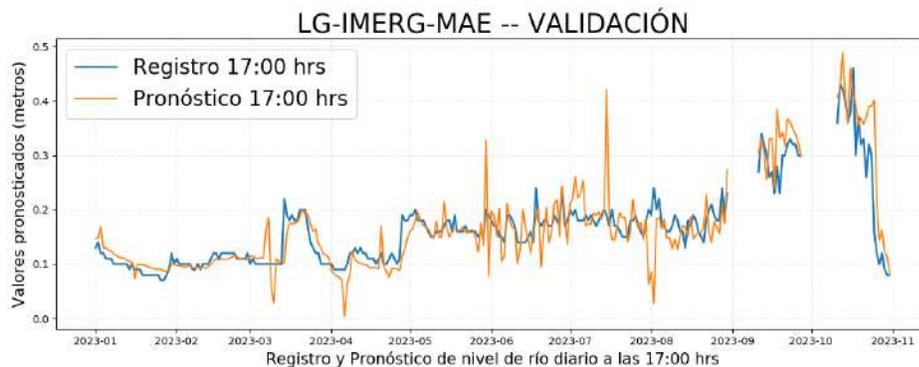
Además, al momento de graficar el comportamiento de cada conjunto de datos (Figura 4.2 y 4.3), se puede observar que el mejor modelo para cada producto satelital tiene un comportamiento bastante cercano al esperado y son buenos candidatos para realizar las variación de cantidad de datos previos de nivel de río.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos

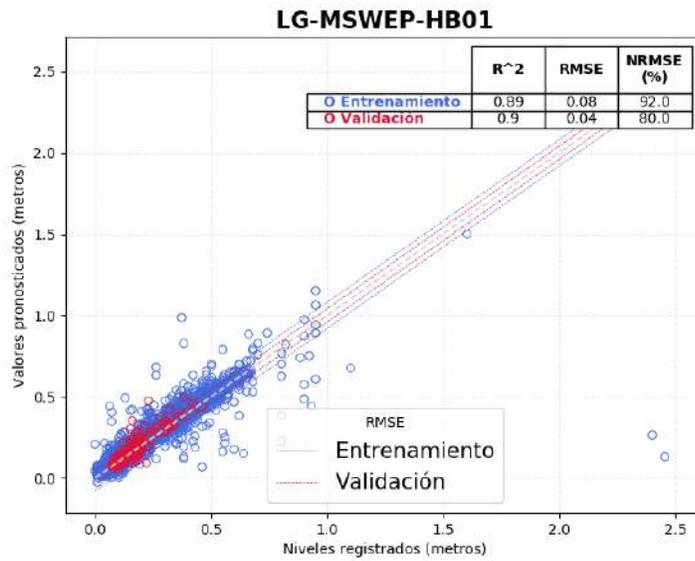


(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2012-2022)

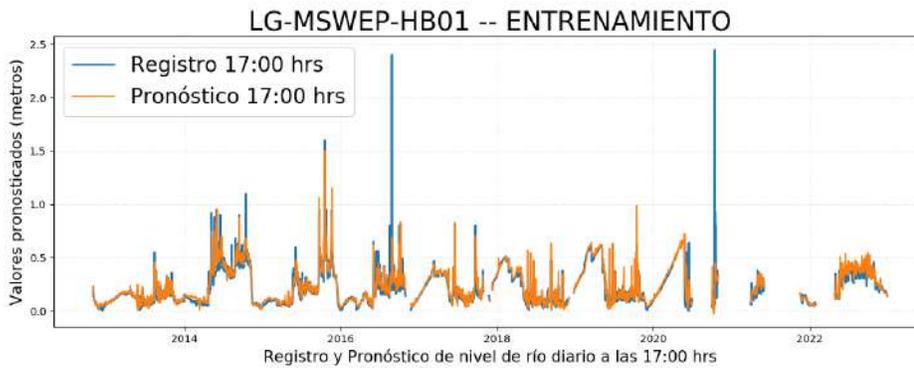


(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

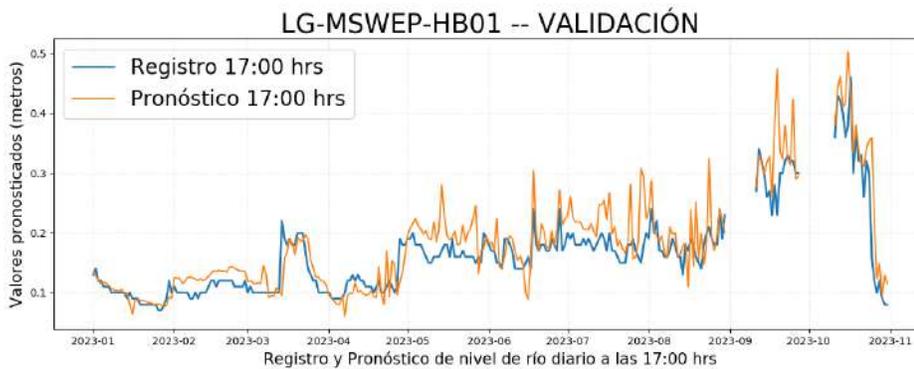
Figura 4.2. Rendimiento del mejor modelo para el IMERG tanto en entrenamiento como validación. El mejor modelo se obtuvo utilizando Error Absoluto Máximo - MAE



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos



(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2012-2022)



(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

Figura 4.3. Rendimiento del mejor modelo para el MSWEP tanto en entrenamiento como validación. El mejor modelo se obtuvo utilizando Hubber Loss, $\alpha = 0.1$

4.3. Variación de datos previos de nivel de río

En el trabajo presentado por M. Campolo [5], se muestra la importancia de considerar datos previos de nivel de río. Al añadir la información previa del nivel de agua se obtiene mejores ajustes que un modelo sin ninguna información de nivel de río previa. A pesar de esta mejora considerable, se debe considerar también que el ámbito operativo del registro de datos y como en Guatemala al ser datos diarios y al estar considerando 10 días previos de información, si por alguna eventualidad un dato no es registrado, los diez días siguientes no podría realizarse un pronóstico ya que las RNA no pueden trabajar con datos faltantes⁵. Es por ello que se analizará la repercusión de remover datos previos de nivel de río de los mejores modelos por producto satelital, para determinar si es posible contar con pronósticos de rendimiento aceptable con un menor requerimiento de datos previos de nivel de agua para que así se tenga una menor vulnerabilidad antes la pérdida de datos.

De las consideraciones previas, se realizan las variaciones de los modelos “LG-IMERG-MAE” y “LG-MSWEP-HB01” removiendo días previos como se puede observar en el cuadro 4.8.

Cuadro 4.8. Rendimiento de distintos modelos de RNA al variar la cantidad de días previos de nivel de agua en ríos para el modelo.

Producto satelital	Días previos	Valor mínimo de la métrica obtenida	Época con mejor resultado	Nombre modelo	Tiempo de entrenamiento (hrs)
IMERG	10	0.03740	390	LG-IMERG-MAE	6.37
	9	0.03952	390	LG-IMERG-MAE-D09	4.29
	7	0.03877	370	LG-IMERG-MAE-D07	4.81
	5	0.03813	380	LG-IMERG-MAE-D05	4.76
	3	0.03945	350	LG-IMERG-MAE-D03	4.81
	1	0.04162	360	LG-IMERG-MAE-D01	4.82
	0	0.08635	400	LG-IMERG-MAE-D00	4.59
MSWEP	10	0.00182	370	LG-MSWEP-HB01	5.89
	9	0.00185	390	LG-MSWEP-HB01-D09	4.59
	7	0.00184	380	LG-MSWEP-HB01-D07	4.94
	5	0.00185	380	LG-MSWEP-HB01-D05	5.10
	3	0.00186	390	LG-MSWEP-HB01-D03	5.17
	1	0.00208	380	LG-MSWEP-HB01-D01	5.29
	0	0.00476	400	LG-MSWEP-HB01-D00	4.79

⁵Se pueden utilizar métodos para completar datos faltantes, pero la repercusión de dicho cambio no se analizará en este trabajo

Al reducir la cantidad de días previos como dato de entrada de todos los modelos para cada producto satelital, no se presentan variaciones significativas si se compara el R^2 y el RMSE de cada modelo, teniendo una diferencia notable hasta que se descarta toda la información previa de nivel de río. Dicho comportamiento se presenta en ambos productos satelitales y se observa en la Figura 4.4.

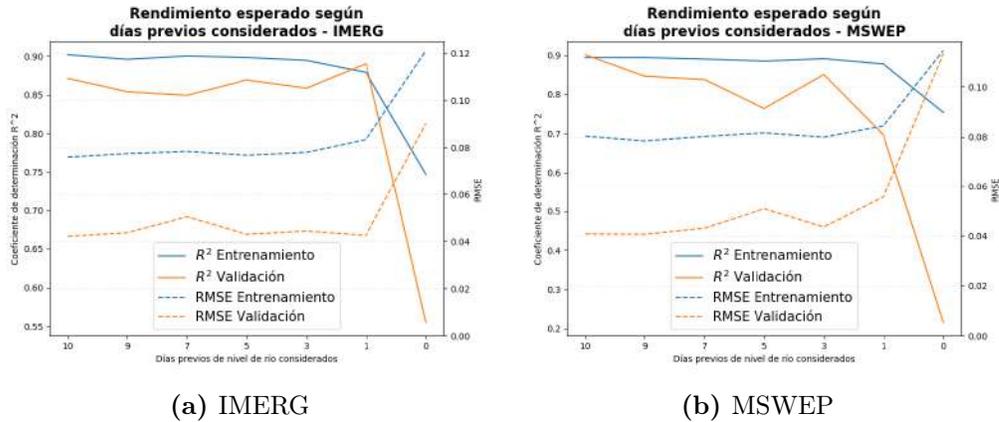


Figura 4.4. Rendimiento de los modelos al quitar días previos de nivel de río como variable de entrada para cada producto satelital.

El cambio, aunque podría considerarse mínimo, repercute de forma significativa al rendimiento y comportamiento del pronóstico como se puede observar en la Figura 4.5. Además, tomando en consideración el criterio de positividad del NSE, los dos modelos que no cuentan con información previa de nivel de río quedan descartados como modelos viables ya que ambos presentan un valor negativo de NSE con los datos de validación. Este comportamiento se puede observar en la Figura 4.6b y 4.6d, donde se observa que los datos pronosticados no corresponden en buena manera con los registros diarios realizados y no tienen una correspondencia lineal como si lo tienen los modelos con información previa de nivel de río (4.2a, 4.6a, 4.3a y 4.6c, para el IMERG y el MSWEP respectivamente).

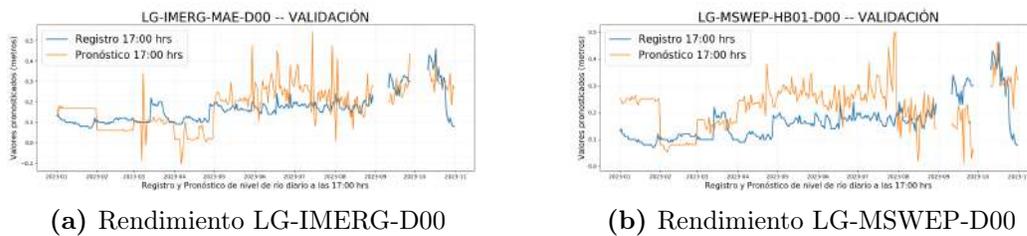
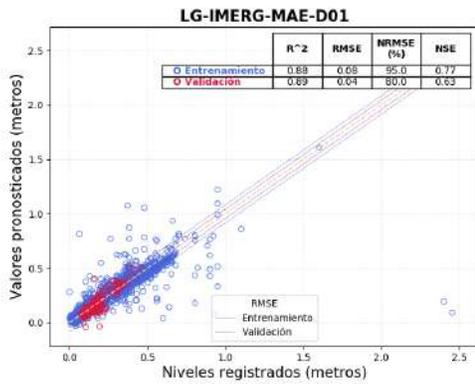
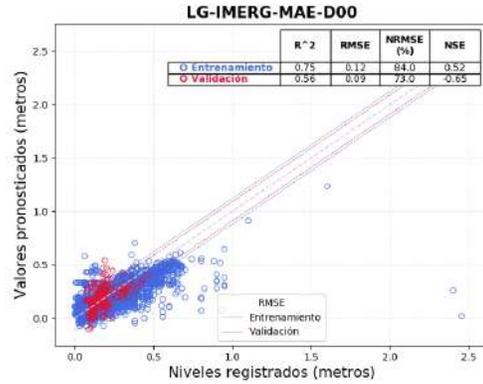


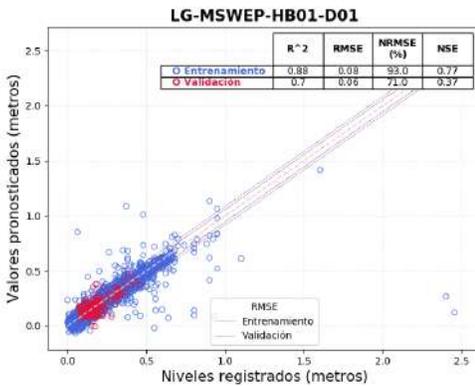
Figura 4.5. Rendimiento de los modelos por producto satelital al no considerar datos previos de nivel de río.



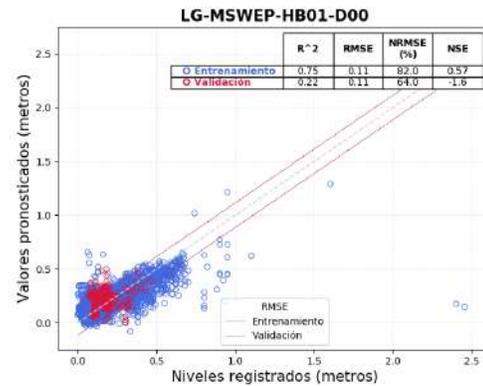
(a) IMERG - 1 día previos de nivel de río



(b) IMERG - sin datos previos de nivel de río



(c) MSWEP - 1 día previos de nivel de río



(d) MSWEP - sin datos previos de nivel de río

Figura 4.6. Rendimiento del error y tiempo de entrenamiento por época para distintas cantidades de capas ocultas utilizando 5 folios con datos del MSWEP y el IMERG.

5. Resultados finales

Tomando en consideración el análisis realizado en los capítulos previos a éste, se observó cómo distintas variaciones afectan el rendimiento de una RNA para realizar el pronóstico de nivel de río diario. Como etapa final de esta investigación, se definen los mejores parámetros encontrados para obtener un pronóstico con un rendimiento aceptable por producto satelital, con el cual se entrenará una RNA como modelo final para la estación “La Gomera” en el río Acomé, y con parámetros similares, se replicará el modelo en dos ríos más de la Vertiente del Océano Pacífico para observar si la metodología es replicable en otras zonas del país.

Los parámetros definidos como base para un modelo final de RNA se definen en el cuadro 5.1:

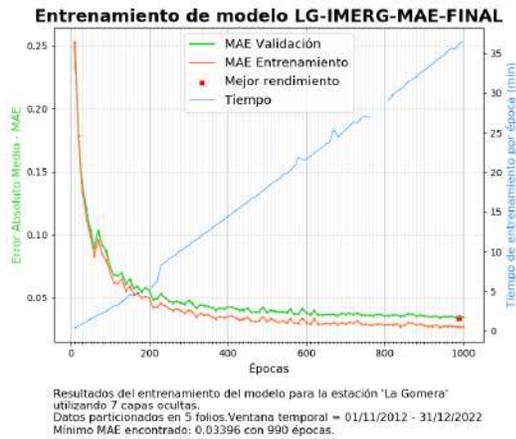
Cuadro 5.1. Hiperparámetros finales para entrenar un modelo de RNA por producto satelital

Producto satelital	Capas ocultas	Función de error	Número de folios	Días previos de nivel de río	Optimizador	Épocas
IMERG	7	Mean Absolute Error	5	5	Adam (0.0001)	10-1000
MSWEP	6	Hubber Loss (0.1)				

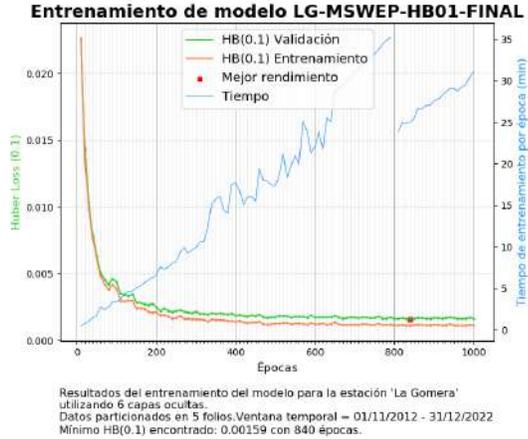
5.1. Cuenca del río Acomé

Cuadro 5.2. Rendimiento de los modelos según los parámetros definidos por producto satelital.

Producto satelital	Función de pérdida	Valor mínimo de la métrica obtenida	Época con mejor resultado	Nombre modelo
IMERG	Mean Absolute Error	0.0339	990	LG-IMERG-MAE-FINAL
MSWEP	Huber Loss 0.1	0.0016	840	LG-MSWEP-HB01-FINAL



(a) Entrenamiento para el IMERG



(b) Entrenamiento para el MSWEP

Figura 5.1. Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.

Cuadro 5.3. Métricas de evaluación para los modelos finales para la estación “La Gomera” en la cuenca del río Acomé para los datos de entrenamiento (2013-2022) y validación (2023).

		MODELO	PBias	MSE	RMSE	N	RMSE	MAE	MaxAE	NSE	R ²	P-val
IMERG	Entr.	LG-IMERG-MAE-FINAL	-0.5403	0.0059	0.0771	0.9570	0.0240	2.3599	0.8042	0.8976	0	
	Valid.	LG-IMERG-MAE-FINAL	-3.3225	0.0012	0.0346	0.8179	0.0221	0.2110	0.7575	0.9048	0	
MSWEP	Entr.	LG-MSWEP-HB01-FINAL	-3.3107	0.0058	0.0762	0.9531	0.0292	2.3317	0.8089	0.9022	0	
	Valid.	LG-MSWEP-HB01-FINAL	0.5086	0.009	0.0307	0.7964	0.0221	0.1837	0.8097	0.9044	0	

	P	N
T	819	1242
F	823	229

(a) Entrenamiento (2013-2022)

	P	N
T	51	136
F	63	30

(b) Validación (2023)

	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	66 %	50 %	78 %	60 %	84 %
Valid.	67 %	45 %	63 %	68 %	82 %

(c) Rendimiento del modelo para la emisión de alertas

Cuadro 5.4. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “LG-IMERG-MAE-FINAL”.

	P	N
T	768	1185
F	880	280

(a) Entrenamiento (2013-2022)

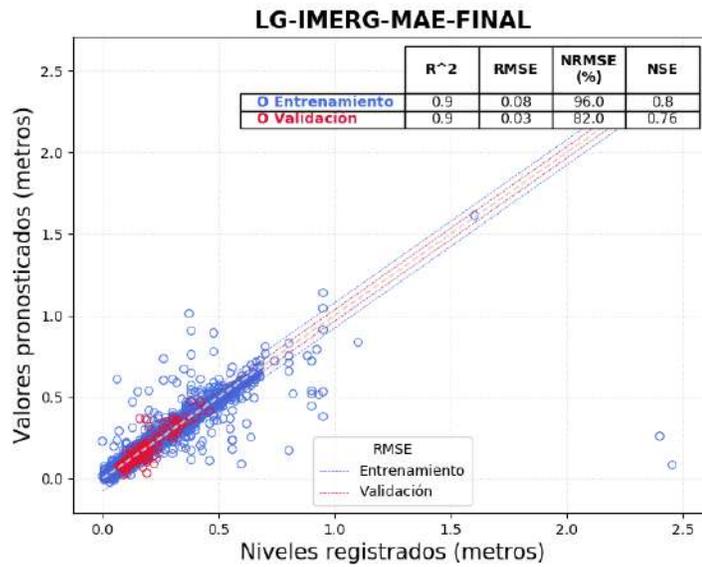
	P	N
T	48	130
F	69	33

(b) Validación (2023)

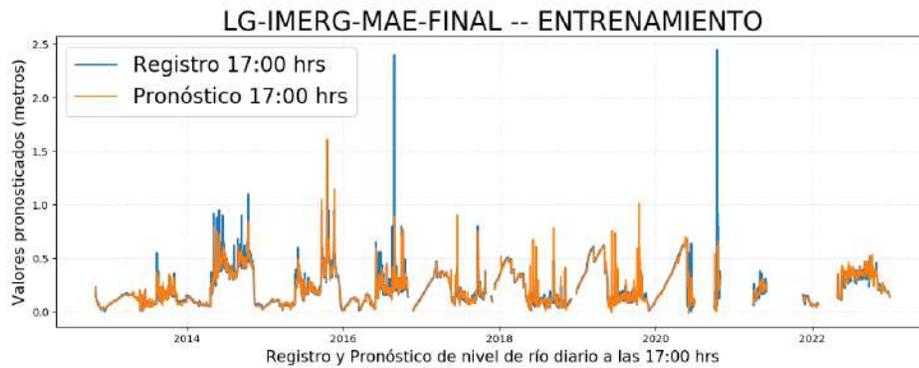
	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	63 %	47 %	73 %	57 %	81 %
Valid.	64 %	41 %	59 %	65 %	80 %

(c) Rendimiento del modelo para la emisión de alertas

Cuadro 5.5. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “LG-MSWEP-HB01”.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos

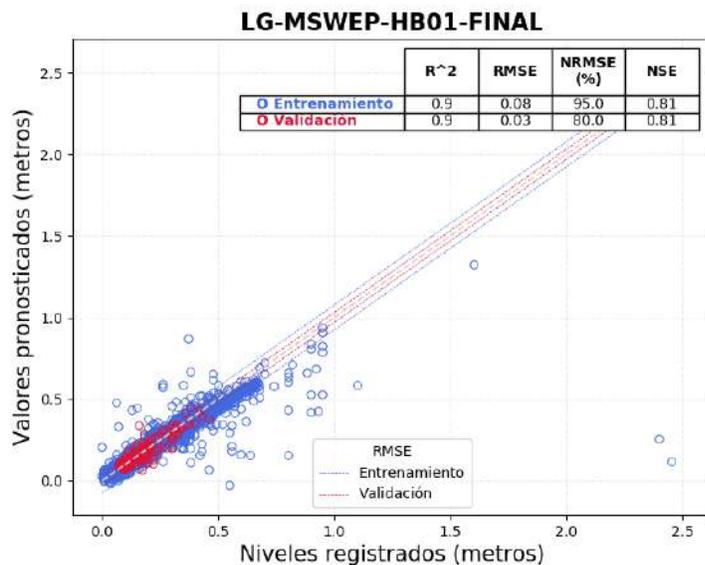


(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2013-2022)

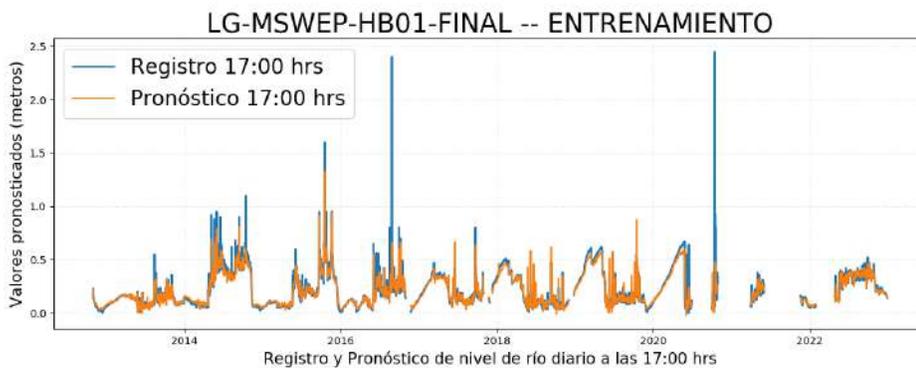


(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

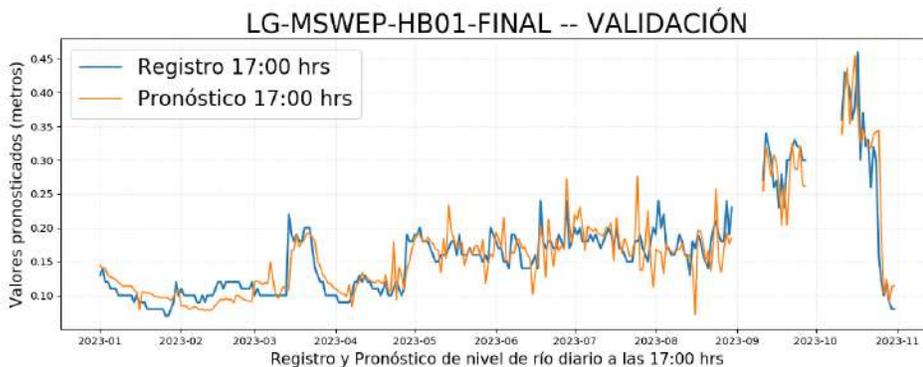
Figura 5.2. Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos



(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2013-2022)



(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

Figura 5.3. Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.

5.2. Cuenca del río Coyolate

En el trabajo realizado por Martínez y Pérez [27], se determinó que la cuenca del río Coyolate es la segunda cuenca hidrográfica donde se pueden realizar pronósticos de nivel de río según la distribución de estaciones disponibles. El mejor pronóstico realizado en esta cuenca es la estación “Puente Coyolate”, por lo que para realizar un pronóstico utilizando la información de ambos productos satelitales se construye la base de datos según los píxeles pertenecientes a la cuenca (Figura 5.4).

5.2.1. Estación Puente Coyolate

De la información disponible se seleccionaron aquellos píxeles que tienen un coeficiente de correlación de Spearman mayor o igual a 0.25, esto según los resultados obtenidos que se pueden observar en la Figura 8.1 y 8.2. Posterior a ello, al determinar las variables a ser consideradas, se construye la base de datos por producto satelital a ser utilizada y se realiza un entrenamiento para cada producto obteniendo los resultados siguientes:

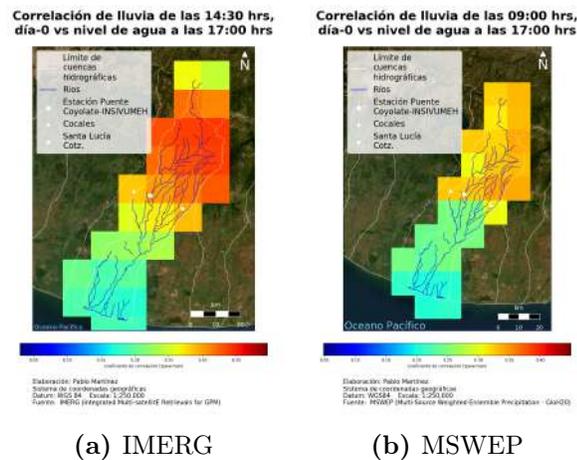
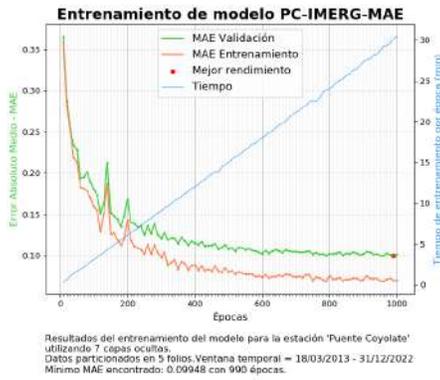


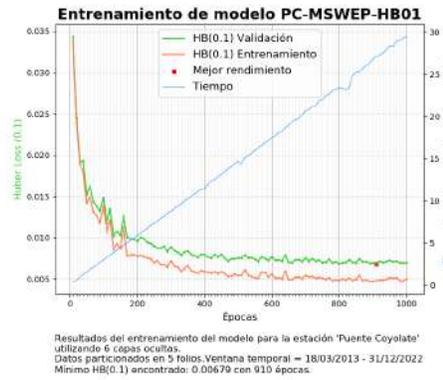
Figura 5.4. Píxeles correspondientes a la cuenca del río Acomé para cada producto satelital

Cuadro 5.6. Rendimiento de los modelos según los parámetros definidos por producto satelital.

Producto satelital	Función de pérdida	Valor mínimo de la métrica obtenida	Época con mejor resultado	Nombre modelo
IMERG	Mean Absolute Error	0.0994	990	PC-IMERG-MAE
MSWEP	Huber Loss 0.1	0.0067	910	PC-MSWEP-HB01



(a) Entrenamiento para el IMERG



(b) Entrenamiento para el MSWEP

Figura 5.5. Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.

Cuadro 5.7. Métricas de evaluación de los modelos para la estación “Puente Coyolate” en la cuenca del Río Coyolate para los datos de entrenamiento (2013-2022) y validación (2023).

		MODELO	PBias	MSE	RMSE	N_RMSE	MAE	MaxAE	NSE	R ²	P-val
IMERG	Entr.	PC-IMERG-MAE	-4.2307	0.0282	0.1679	0.9142	0.0664	2.0105	0.7783	0.8893	0
	Valid.	PC-IMERG-MAE	-4.2314	0.0260	0.1611	0.8992	0.0825	1.0350	0.5750	0.7667	0
MSWEP	Entr.	PC-MSWEP-HB01	-8.3214	0.0355	0.1883	0.9262	0.0766	2.0649	0.7211	0.8711	0
	Valid.	PC-MSWEP-HB01	-11.9159	0.0222	0.1490	0.8975	0.0813	1.2311	0.6329	0.8292	0

	P	N
T	441	1161
F	515	213

(a) Entrenamiento (2013-2022)

	P	N
T	51	183
F	82	41

(b) Validación (2023)

	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	69%	46%	67%	69%	84%
Valid.	66%	38%	55%	69%	82%

(c) Rendimiento del modelo para la emisión de alertas

Cuadro 5.8. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “PC-IMERG-MAE”.

	P	N
T	435	1073
F	603	219

(a) Entrenamiento (2013-2022)

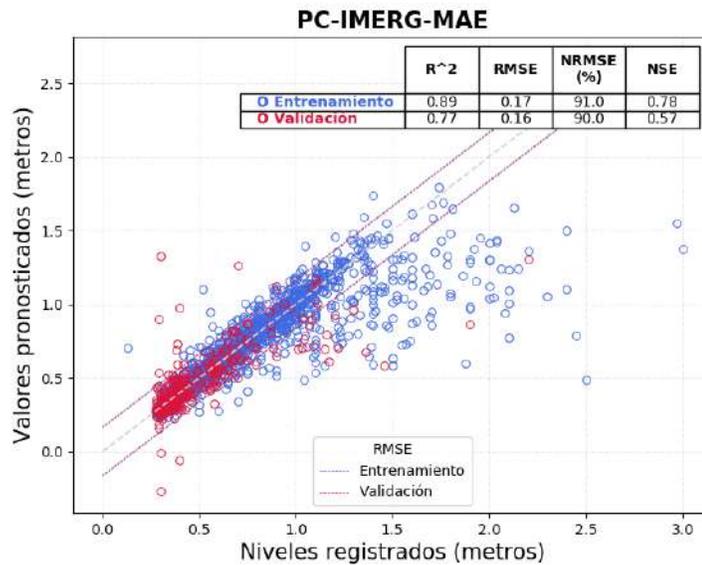
	P	N
T	56	162
F	105	38

(b) Validación (2023)

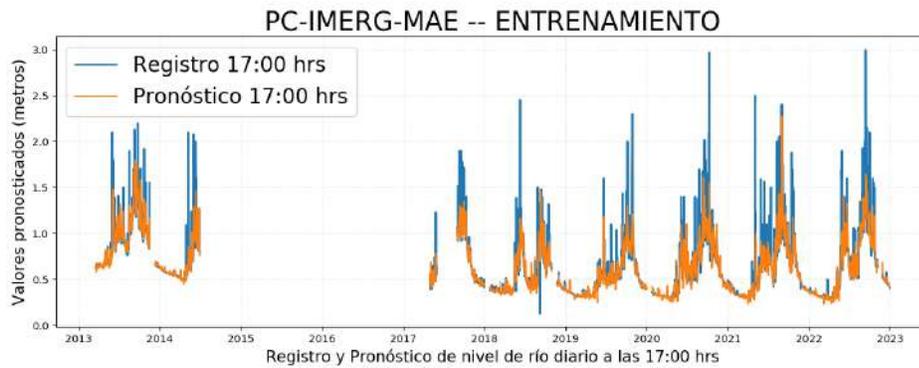
	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	65%	42%	67%	64%	83%
Valid.	60%	35%	60%	61%	81%

(c) Rendimiento del modelo para la emisión de alertas

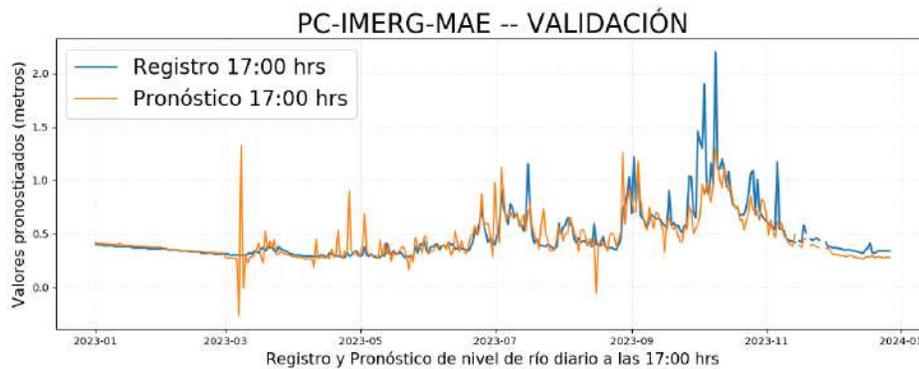
Cuadro 5.9. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “PC-MSWEP-HB01”.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos

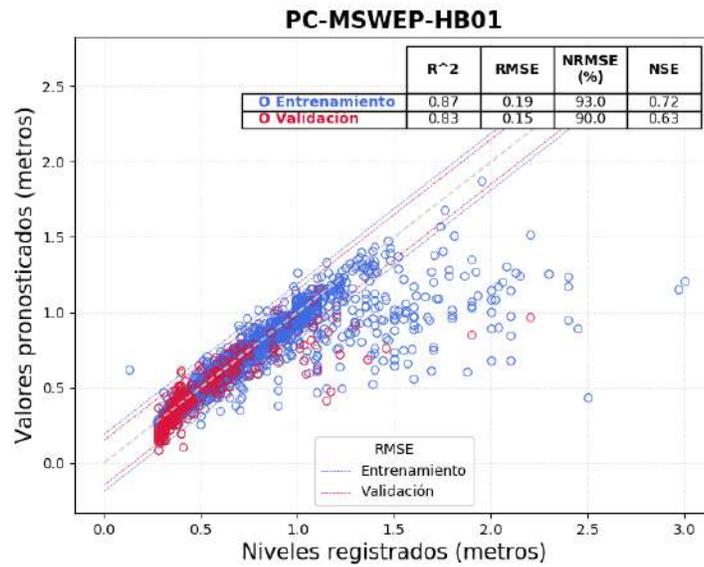


(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2013-2022)

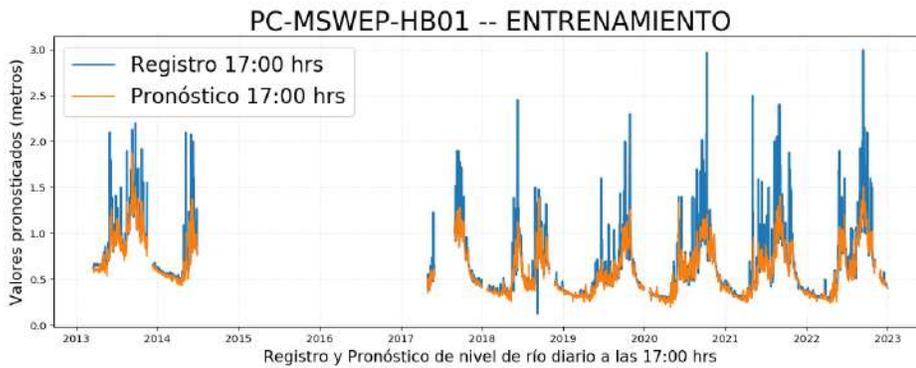


(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

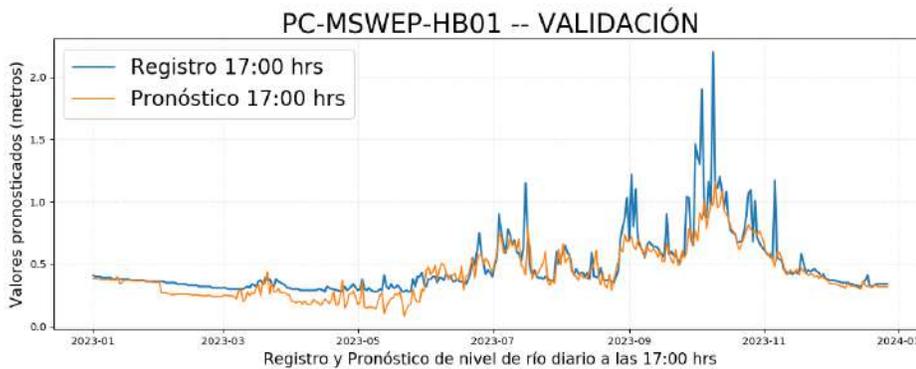
Figura 5.6. Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos



(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2013-2022)



(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

Figura 5.7. Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.

5.2.2. Estación El Carrizal

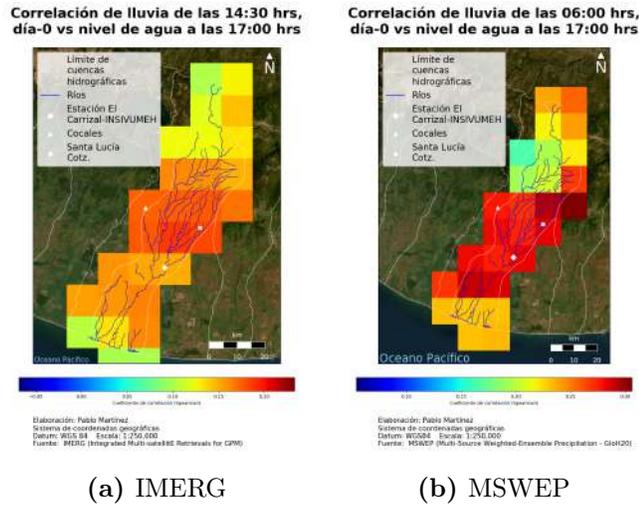


Figura 5.8. Píxeles correspondientes a la cuenca del río Coyoate para cada producto satelital

Cuadro 5.10. Rendimiento de los modelos según los parámetros definidos por producto satelital.

Producto satelital	Función de pérdida	Valor mínimo de la métrica obtenida	Época con mejor resultado	Nombre modelo
IMERG	Mean Absolute Error	0.1177	690	EC-IMERG-MAE
MSWEP	Huber Loss 0.1	0.0086	860	EC-MSWEP-HB01

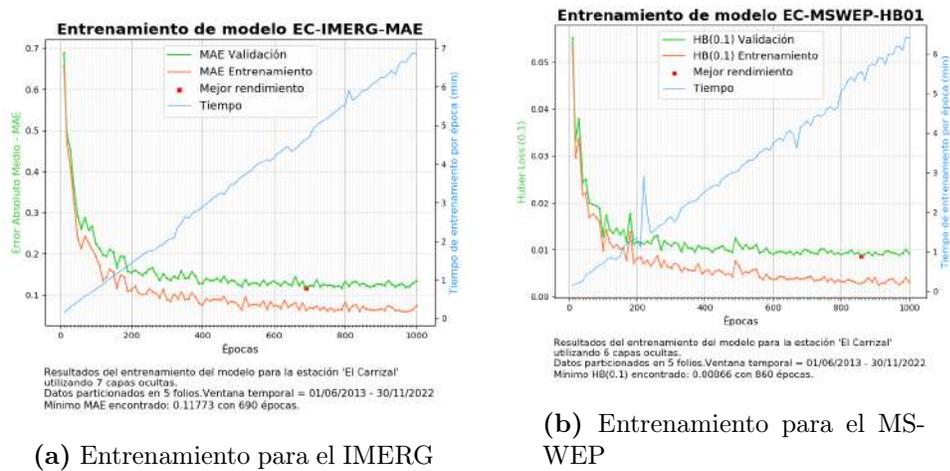
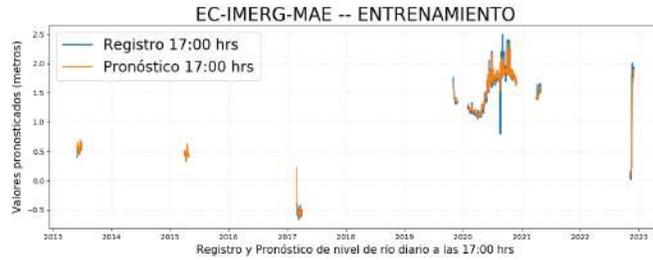
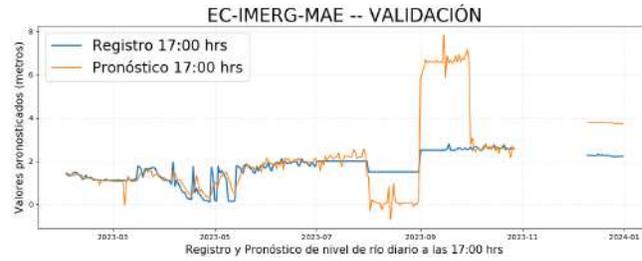


Figura 5.9. Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.

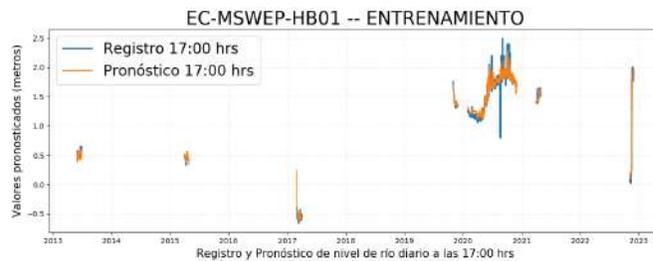


(a) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2012-2022)



(b) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

Figura 5.10. Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.



(a) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2012-2022)



(b) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

Figura 5.11. Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.

5.3. Cuenca del río Ocosito

En el trabajo de Martínez y Pérez [27], la cuenca del río Ocosito no presenta un pronóstico debido a que las estaciones meteorológicas disponibles se encuentran en latitudes más bajas que las estaciones hidrométricas del INSIVUMEH, por lo que se descartó como una cuenca viable debido a la falta de información disponible. Es por ello que se añade a este trabajo como una revisión si los productos satelitales pueden cubrir la información faltante y volver viable la realización de un pronóstico en más cuencas.

Dentro de la cuenca se encuentra disponible información de la estación “El Asintal” y “Caballo Blanco” siendo esta última la seleccionada para realizar un pronóstico. De la información disponible por producto satelital, se construye la base de datos según los píxeles pertenecientes a la cuenca (Figura 5.12). De la información disponible para esta cuenca se seleccionaron aquellos píxeles que tienen un coeficiente de correlación de Spearman mayor o igual a 0.35 y 0.45, para el IMERG y el MSWEP respectivamente, esto según los resultados obtenidos (Anexos 8.7 y 8.8). Es con esta selección de variables que se construye la base de datos para la estación y con la cual se realiza un entrenamiento por producto satelital obteniendo los resultados siguientes:

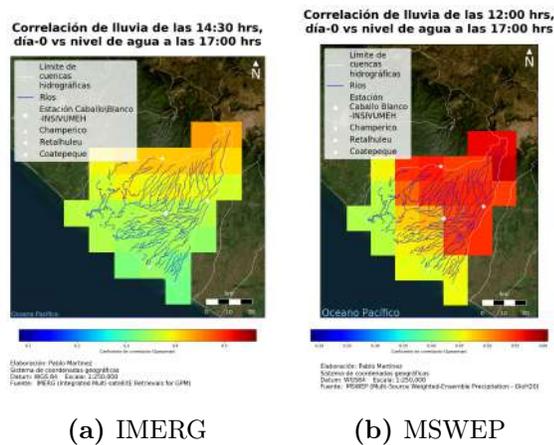


Figura 5.12. Píxeles correspondientes a la cuenca del río Ocosito por producto satelital

Cuadro 5.11. Rendimiento de los modelos por producto satelital.

Producto satelital	Función de pérdida	Valor mínimo de la métrica obtenida	Época con mejor resultado	Nombre modelo
IMERG	Mean Absolute Error	0.1385	520	CB-IMERG-MAE
MSWEP	Huber Loss 0.1	0.0061	870	CB-MSWEP-HB01

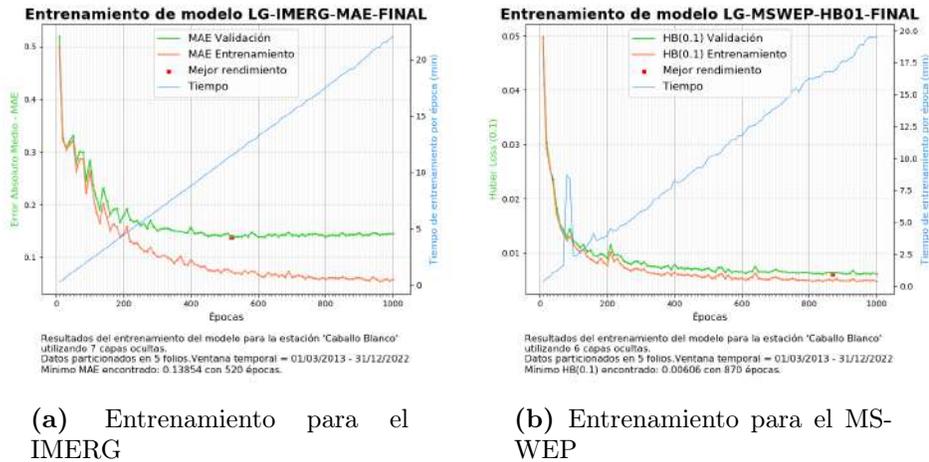


Figura 5.13. Rendimiento del error para los datos de entrenamiento, validación y tiempo de entrenamiento por época para los modelos por producto satelital.

Cuadro 5.12. Métricas de evaluación para los modelos finales para la estación “Caballo Blanco” en la cuenca del río Ocosito para los datos de entrenamiento (2013-2022) y validación (2023).

		MODELO	PBias	MSE	RMSE	N_RMSE	MAE	MaxAE	NSE	R ²	P-val
IMERG	Entr.	CB-IMERG-MAE	-6.6167	0.0421	0.2053	0.9557	0.0685	3.7281	0.7541	0.8814	0
	Valid.	CB-IMERG-MAE	0.3978	0.0521	0.2282	0.8533	0.1296	1.6834	0.6241	0.7922	0
MSWEP	Entr.	CB-MSWEP-HB01	-3.7318	0.0539	0.2322	0.9578	0.0749	3.6756	0.6854	0.8356	0
	Valid.	CB-MSWEP-HB01	4.7684	0.0352	0.1877	0.8507	0.1119	1.8940	0.7455	0.8672	0

	P	N
T	441	829
F	366	261

(a) Entrenamiento (2013-2022)

	P	N
T	62	136
F	77	59

(b) Validación (2023)

	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	67%	55%	63%	69%	76%
Valid.	59%	45%	51%	64%	70%

(c) Rendimiento del modelo para la emisión de alertas

Cuadro 5.13. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el IMERG utilizando el modelo “CB-IMERG-MAE”.

	P	N
T	455	820
F	375	247

(a) Entrenamiento (2013-2022)

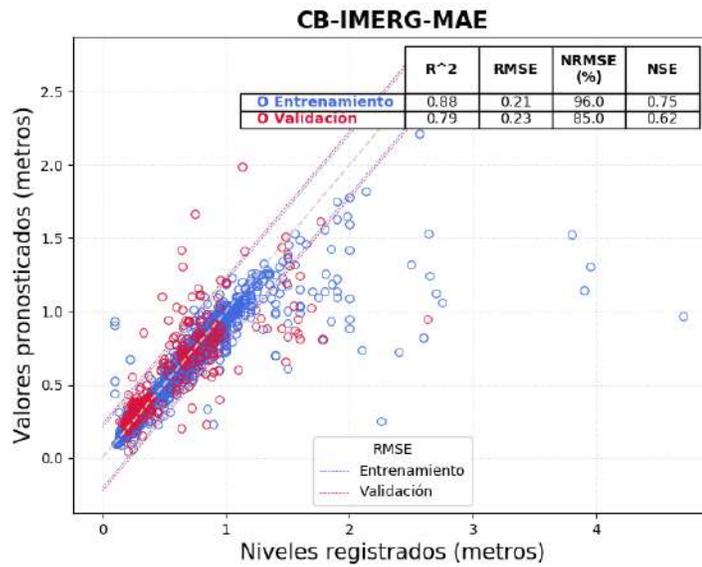
	P	N
T	75	135
F	79	46

(b) Validación (2023)

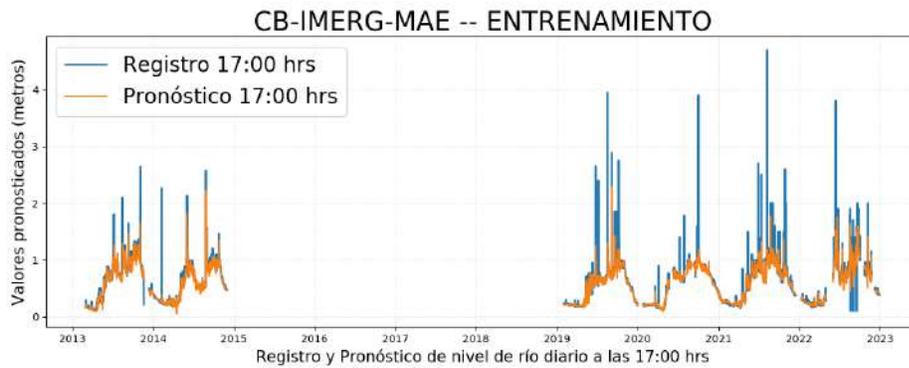
	Exac.	Prec.	Sensib.	Espec.	NValP.
Entr.	67%	55%	65%	69%	77%
Valid.	63%	49%	62%	63%	75%

(c) Rendimiento del modelo para la emisión de alertas

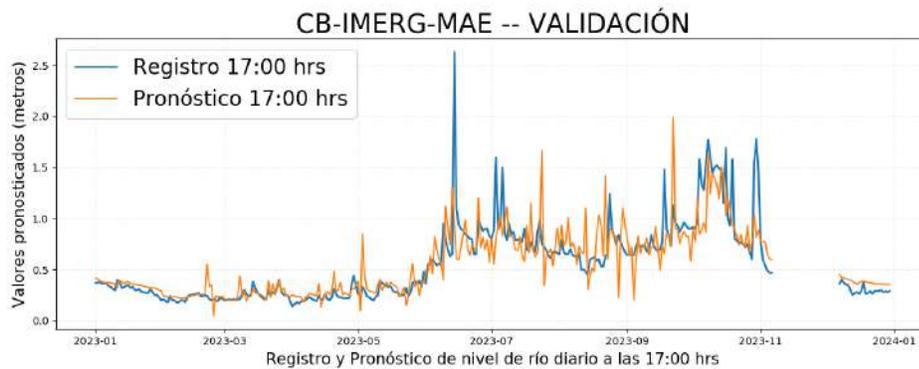
Cuadro 5.14. Matriz de confusión para cada conjunto de datos y valores resultantes de Exactitud, Precisión, Sensibilidad, Especificidad y Valor Predictivo Negativo para el MSWEP utilizando el modelo “CB-MSWEP-HB01”.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos

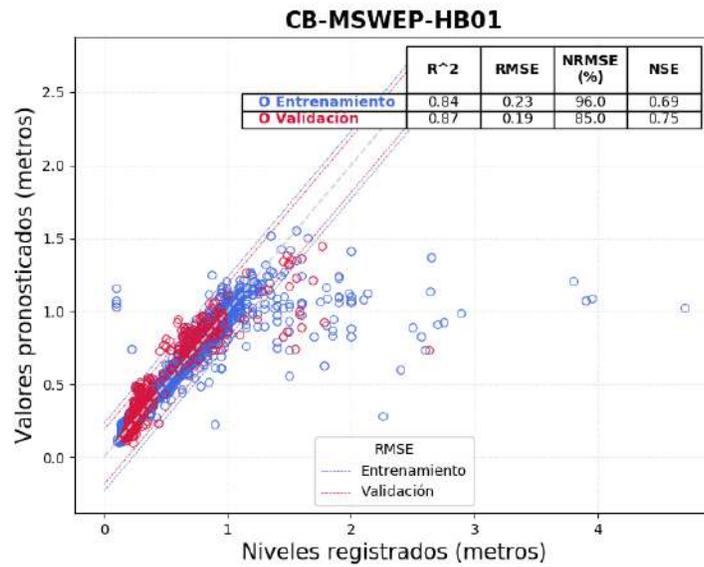


(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2013-2022)



(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

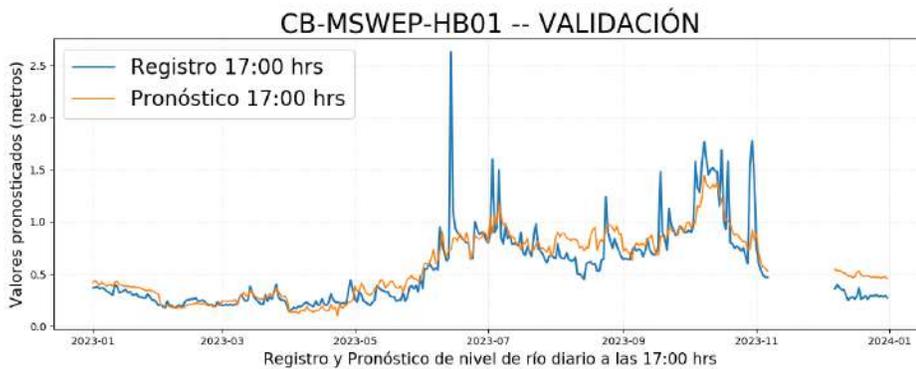
Figura 5.14. Rendimiento del modelo final para el IMERG tanto en entrenamiento como validación.



(a) Pronóstico diario vs Registro diario de nivel de agua en ríos



(b) Pronóstico diario y registros diarios de nivel de agua en el río para el período de entrenamiento (2013-2022)



(c) Pronóstico diario y registros diarios de nivel de agua en el río para el año de validación (2023)

Figura 5.15. Rendimiento del modelo final para el MSWEP tanto en entrenamiento como validación.

6. Discusión de resultados

Muchas instituciones internacionales ponen a disposición del público en general información sobre precipitación líquida a nivel global casi en tiempo real. Esta información es muy valiosa para países como Guatemala debido a la vulnerabilidad de la región ante tormentas e inundaciones [1]. A pesar de existir una gran variedad de productos satelitales, no todos ellos son útiles para realizar un pronóstico inmediato, ya que el retraso que tienen entre la hora de registro y su hora de publicación es muy lejano debido al procesamiento de los datos originales. Es por ello que de las fuentes disponibles se seleccionó como principales candidatos al CHIRPS, el IMERG y el MSWEP, pero de estos, debido al retraso de publicación y la calidad de los datos preliminares, se descartó la información diaria de CHIRPS (Cuadro: 3.4), quedando únicamente los productos del IMERG y el MSWEP. Partiendo del trabajo de Martínez y Pérez [27], se delimitó la zona de interés a la Vertiente del Océano Pacífico y se tomó como línea base los resultados obtenidos para las tres estaciones consideradas en dicho trabajo. Considerando que la estación que tiene un mejor rendimiento es la cuenca del río Acomé, en la estación de “La Gomera”, se definió la metodología para realizar la construcción de variables que puedan servir como vector de entrada para una Red Neuronal Artificial (RNA) en torno a dicha estación.

Generalmente, para realizar un entrenamiento de un modelo de RNA se utiliza conocimiento a priori sobre el problema a resolver para definir las variables de entrada [25] y considerando en este caso el ciclo hidrológico, así como la información satelital disponible considerada, se utilizarán 5 tipos de variables que son:

- Variables de nivel de río diarias.
- Variables mensuales de nivel de río.
- Variables de lluvia “directa” a resolución mínima.
- Variables de lluvia “indirecta” a resolución mínima.

- Variables mensuales de lluvia.

Buscando evitar agregar error al modelo debido a variables con información poco relevante para el modelo, se seleccionan solo aquellas que tengan un coeficiente de correlación de Spearman igual o superior a un umbral establecido según los resultados obtenidos por cada estación a pronosticar. Los umbrales definidos para cada estación se presentan en el cuadro 6.1. El descarte de variables repercute en la cantidad de variables disponibles como vector de entrada y por lo tanto en la cantidad de neuronas de la capa de entrada de la RNA, teniendo una cantidad de variables distinta por producto satelital como se presenta en el cuadro 6.2.

Cuadro 6.1. Umbrales del coeficiente de correlación de Spearman definidos por cada estación hidrométrica considerada.

Estación	Cuenca	Umbral Coef. Spearman			
		Río	IMERG	MSWEP	Var. Mensuales
La Gomera	Acomé	0.6	0.15	0.15	0.15
Puente Coyolate	Coyolate	0.6	0.25	0.25	0.15
El Carrizal	Coyolate	0.6	0.2	0.15	0.15
Caballo Blanco	Ocosito	0.6	0.45	0.35	0.2

Cuadro 6.2. Cantidad de neuronas de la capa de entrada por estación y producto satelital.

Estación	Cuenca	Cantidad de neuronas	
		IMERG	MSWEP
La Gomera	Acomé	406	368
Puente Coyolate	Coyolate	469	478
El Carrizal	Coyolate	316	308
Caballo Blanco	Ocosito	821	344

6.1. Variación de parámetros

Considerando que el diseño de una RNA es dependiente del problema a resolver [12] [25], para determinar los parámetros ideales suficientes para obtener un rendimiento aceptable de una RNA que pronostique niveles de agua en río en la vertiente del océano pacífico

Uno de los parámetros que se pueden variar al realizar un entrenamiento de una RNA es la cantidad de capas ocultas y neuronas por capa que tiene la red. Para

este caso, cada capa tiene la misma cantidad de neuronas que la capa de entrada de la red.

Del cuadro 4.1 y 4.2, se puede observar que al aumentar el número de capas ocultas, se mejora el rendimiento del MSE, pero si se transforma el valor al RMSE, este no presenta variaciones significativas que equiparen el tiempo de entrenamiento necesario para obtener dicho rendimiento. Por otro lado, al realizar la variación en la cantidad de folios por el método de validación cruzada, se observa que a mayor cantidad de folios (al tener mas datos disponibles para entrenamiento) se necesita una menor cantidad de capas ocultas para obtener un mejor rendimiento del MSE. El aumentar la cantidad de folios también repercute en el tiempo necesario de entrenamiento ya que se utilizan más datos por cada entrenamiento y una mayor cantidad de repeticiones para estimar el rendimiento del error del modelo. De ambos cuadros se puede observar que con al menos 5 folios se puede obtener un rendimiento aceptable por modelo sin aumentar de manera drástica el tiempo de entrenamiento.

Aunque el MSE y el RMSE se utilizan como una métrica para evaluar el rendimiento de una RNA debido a la interpretabilidad de este último como el error estándar del modelo, no es la única función de error que puede utilizarse para realizar el entrenamiento de una RNA. Las funciones de error consideradas para este trabajo fueron:

- Error Cuadrático Medio - MSE
- Error Absoluto Medio - MAE
- Error Absoluto Máximo - MaxAE
- Función de pérdida de Huber - Huber Loss $\rho_{\delta}(a)$
 - $\delta = 1.0$, $\rho_{1.0}(a)$
 - $\delta = 0.5$, $\rho_{0.5}(a)$
 - $\delta = 0.1$, $\rho_{0.1}(a)$

Los resultados de las variaciones al sustituir la función de error, habiendo fijado el número de capas ocultas y folios presentados en el cuadro 4.5, muestran que se puede obtener un mejor rendimiento de MSE con funciones de error distintas al MSE y además, al evaluar las distintas métricas de error propuestas, se determina que las funciones de error que permiten realizar un buen pronóstico de nivel de río son el MAE para la información del IMERG y la pérdida de Huber $\rho_{0.1}(a)$ para la información del MSWEP.

Como se muestra en el trabajo de Campolo [5], los datos previos de nivel de río son importantes para obtener mejores resultados en el pronóstico a realizar, pero considerando que en sistemas de registro de datos como el utilizado por el INSI-VUMEH, depender de 10 días de registro para realizar el pronóstico del día actual, podría ser perjudicial, ya que de no contarse con un registro se estarían perdiendo diez días de pronóstico a menos que se caracterice y utilice una metodología para rellenar datos faltantes sin que tenga repercusiones en el error de la RNA utilizada. Es por ello que se realizaron variaciones quitando días previos de nivel de río para observar si era posible prescindir de datos previos de nivel de río y así robustecer la operatividad de un modelo a costa de perder información siempre y cuando esto no imposibilite la realización de un pronóstico. Posterior a realizar las variaciones en cada producto satelital se obtienen los resultados presentados en el cuadro 4.8, de los cuales se puede observar que con al menos un día previo de nivel de río se mantiene un rendimiento similar en la métrica de la función de error utilizada. Es si no hasta que se remueven todos los datos previos de nivel de río que se presenta un aumento en el valor mínimo de la métrica obtenida al realizar el entrenamiento. Este aumento podría considerarse despreciable, pero al realizar un análisis más profundo con las otras métricas de evaluación consideras, se puede observar como el rendimiento es inadmisibile ya que el NSE es menor a cero en la validación y además, el modelo difiere en gran medida con el comportamiento real y no sería un pronóstico útil para sistemas de alerta temprana (Figuras 4.4 y 4.5). Es por ello que como un punto intermedio entre contar con información valiosa para el modelo y no ser tan susceptible a la pérdida de datos, se fija que para los modelos finales se debería contar con al menos 5 días previos de nivel de río.

6.2. Pronósticos de nivel de río en la Vertiente del Océano Pacífico

Con los parámetros fijados como estructura básica de una RNA se realizan dos modelos de pronóstico para distintas estaciones hidrométricas en la vertiente del Océano Pacífico. Cada una de ellas permite evaluar distintos aspectos del rendimiento esperado sobre un pronóstico de este tipo.

En el trabajo de Martínez y Pérez [27] se realizaron modelos para tres estaciones hidrométricas en la vertiente del Océano Pacífico de Guatemala. Para cada uno de ellos utilizó únicamente el MSE y el RMSE como métrica de evaluación y tuvieron

como datos de validación los meses de enero a mayo de 2023. Es por ello que, para realizar una comparación con los datos obtenidos en este trabajo solo se considerarán las mismas dos métricas de evaluación para comparar si el rendimiento de los modelos obtenidos en este trabajo son equiparables al trabajo previo realizado. Los resultados obtenidos se muestran en el cuadro 6.3.

Cuadro 6.3. Comparación de modelos de RNA utilizando estaciones meteorológicas y satelitales

		Rendimiento de modelos		
Estación	Cuenca	Trabajo Previo	IMERG	MSWEP
La Gomera	Acomé	0.0005	0.0006	0.0004
Puente Coyolate	Coyolate	0.0021	0.0152	0.0075
El Carrizal	Coyolate	0.0181	0.1243	0.1487
Caballo Blanco	Ocosito	-	0.0066	0.0046

Del cuadro anterior se puede observar primeramente que para la estación “La Gomera” el comportamiento es bastante similar, incluso superior para el modelo que utiliza los datos del MSWEP. En el caso de la estación Puente Coyolate y El carrizal, el rendimiento es peor que los obtenidos en el trabajo de Martínez y Pérez pero esto puede deberse a la cantidad de datos disponibles para el entrenamiento, ya que debido a las variables meteorológicas indirectas que consideran hasta 24 meses previos al día a pronosticar, se presenta más datos faltantes intermedios en el período de (2013-2022) para ambas estaciones (Figura 5.6b y 5.10a). A pesar de esto, el pronóstico de la estación “Puente Coyolate” si es apta para ser utilizada como sistema de alerta temprana y ambas estaciones podrían mejorar su rendimiento, al considerarse una ventana menor en las variables meteorológicas indirectas que podría permitir una mayor cantidad de datos de entrenamiento de la RNA, ya que a diferencia del trabajo de Martínez y Pérez [27], el comportamiento del entrenamiento si tiene una tendencia a converger al aumentar la cantidad de épocas durante el entrenamiento como se puede observar en las Figuras 5.5 y 5.9.

Para el caso de la estación Caballo Blanco no existe un valor de comparación ya que esta fue una estación descartada como viable para su utilización ya que no contaba con datos disponibles para realizar un pronóstico. Esto cambia al considerar datos satelitales ya que estos permiten realizar un pronóstico con un rendimiento aceptable.

Los resultados obtenidos para las cuatro estaciones consideradas en este trabajo muestran que la realización de un pronóstico en distintas estaciones de la

vertiente del Océano Pacífico es posible ya que la cobertura satelital es más grande y complementa en buena manera los registros de las estaciones hidrométricas del INSIVUMEH. Es importante considerar que una estructura general de RNA para la región no es lo más conveniente, ya que como se ha observado en los resultados previos, si permite realizar pronósticos con un rendimiento aceptable y capaces de formar parte de un sistema de alerta temprana, pero no asegura que sea el mejor rendimiento posible de la RNA ya que ciertos comportamientos en los datos, como la cantidad de datos faltantes durante el período de entrenamiento podrían corregirse al considerar ventanas más pequeñas de variables meteorológicas indirectas, así como las variables de río previas.

Como revisión final se analiza el comportamiento de cada par de modelos para cada una de las cuatro estaciones disponibles en este trabajo y se analiza su capacidad operativa para formar parte de un sistema de alerta temprana, por lo que para ello si se considera la ventana completa de datos de validación y las demás métricas de evaluación propuestas en este trabajo.

6.2.1. Estación La Gomera

De los dos modelos disponibles para la estación “La Gomera”, por las métricas de evaluación, se tiene que el mejor rendimiento lo tiene el modelo “LG-MSWEP-HB01-FINAL”, ambos son buenos modelos pero el modelo antes mencionado se comporta de mejor manera durante el año de validación, por lo que podría decirse que si se hubiera utilizado durante el año 2023 como un pronóstico operativo, hubiera sido la mejor opción disponible. Este modelo se caracteriza por tener un error estándar de 7 y 3 cm para entrenamiento y validación respectivamente, lo cual es un error aceptable sobre todo si se compara con la desviación estándar de cada conjunto de datos, ya que estos valores son menores a los 17 y 7 cm de desviación estándar, donde además, el 95 % y 77 % de los datos tienen un error menor al error estándar presentado previamente.

Por otro lado, la eficiencia de Nash–Sutcliffe nos indica que tanto para entrenamiento como para validación, el resultado del pronóstico es bastante bueno al presentarse un valor cercano a 1. Además, este comportamiento se mantiene tanto en entrenamiento y validación por lo que el modelo no está sobreajustado y se realiza un buen pronóstico aún con datos no vistos durante el entrenamiento y que podría ser útil posterior a su entrenamiento e implementación.

Aunque el valor pronosticado es bastante cercano a los niveles registrados, si

se considera su utilización como un pronóstico para emitir alertas por crecidas o disminuciones del nivel de río, se debe considerar que estas tendrían una exactitud del 64 %, pero considerando que de las alertas por aumento de nivel de río el 59 % hubieran sido falsas alarmas y para cuando se pronosticara que el nivel de río se mantendría igual o descendería su nivel, se hubiera acertado el 80 % de las veces, se recomienda una calibración a esta clasificación que permita obtener mejores resultados. Si dicha calibración no fuera posible de realizar, entonces hubiera sido recomendable utilizar el modelo “LG-IMERG-MAE-FINAL” ya que tiene un mejor rendimiento en cuanto a alertas por crecidas de nivel de río, aunque de igual forma, este modelo presenta más del 50 % de alertas emitidas como falsas alarmas.

6.2.2. Estación Puente Coyolate

De los dos modelos disponibles para la estación “Puente Coyolate” se tiene que ambos tienen un rendimiento bastante parecido y se decide que el mejor modelo es “PC-MSWEP-HB01” debido a que es el que mejor NSE presenta durante la validación y es la métrica con mayor peso al momento de seleccionar un modelo como se mencionó en la sección 4.2.1. Este modelo se caracteriza por tener un error estándar de 18 y 15 cm para entrenamiento y validación respectivamente, lo cual es un error aceptable al compararse con los 36 y 25 cm de desviación estándar que presentan los datos, donde además, el 93 % y 90 % de los datos son menores al error estándar presentado previamente.

Respecto al NSE, tanto para entrenamiento como para validación, el valor de eficiencia se mantiene superior a 0.6 por lo que sigue siendo un buen pronóstico, aunque si debe considerarse que durante la validación presenta una disminución respecto a los datos de entrenamiento. Este comportamiento podría deberse a los datos faltantes en la ventana de entrenamiento y que podrían corregirse estableciendo umbrales y ventanas que ajusten mejor a los datos de la estación durante la etapa de construcción de variables.

Este comportamiento repercute también al utilizar el modelo para emitir alertas por crecida o disminución de nivel de río ya que se tiene una exactitud de emisión de alertas del 60 % pero con un 65 % de falsas alarmas al emitir un pronóstico por crecida y un acierto del 80 % al indicar que el nivel de río va a permanecer igual o a descender.

6.2.3. Estación El Carrizal

En la estación “El Carrizal”, debido a que existen discontinuidades recurrentes entre el 2013 y 2022, las variables mensuales de río únicamente pueden calcularse para una menor cantidad de meses y reduce en forma considerable los datos de entrenamiento disponibles, por lo que para obtener un rendimiento correcto debería cambiarse la metodología para la creación de estas variables para que se pueda tener más datos disponibles para el entrenamiento pero sin perder la información de dichas variables. Además, para las gráficas del comportamiento durante el año 2023 se observan inconsistencias en los datos registrados proporcionados (Figura 5.11b y 5.11b), por lo que además, para esta estación en específico se recomienda un tratamiento de datos más profundo, así como un cambio a la metodología de creación de variables para que el mismo pueda ser utilizado como un pronóstico para alerta temprana. Debido a las inconsistencias presentadas, no se presentan los resultados de las métricas de evaluación obtenidas para los modelos de esta estación ya que ambos son modelos deficientes, pero se dejan registrados ya que representan casos que podrían presentarse si se desea replicar la metodología en el resto de estaciones hidrométricas de la vertiente del Océano Pacífico.

6.2.4. Estación Caballo Blanco

De los dos modelos disponibles para la estación “Caballo Blanco”, se selecciona como mejor modelo al “CB-MSWEP-HB01” ya que obtiene un mejor rendimiento durante la validación que durante el entrenamiento, el cual es el comportamiento deseado. En las otras métricas de evaluación tiene un comportamiento similar al otro modelo disponible y es por ello que se selecciona como mejor candidato para esta estación. Este modelo se caracteriza por tener un error estándar de 23 y 18 cm para entrenamiento y validación respectivamente, lo cual es un error aceptable al compararse con los 41 y 37 cm de desviación estándar que presentan los datos de nivel de río registrados por la estación, donde además, el 96 % y 85 % de los datos pronosticados son menores al error estándar presentado previamente. Respecto al NSE, tanto para entrenamiento y validación el valor de eficiencia se mantiene superior a 0.6, por lo que sigue siendo un buen pronóstico pero es recomendable refinar el entrenamiento para que el valor sea más cercano a 1. Este comportamiento, al igual que con la estación Puente Coyolate, podría deberse a los datos faltantes en la ventana de entrenamiento y que podrían corregirse estableciendo umbrales y ventanas que se ajusten de mejor manera a los datos de la estación durante la etapa

de construcción de variables.

En caso de utilizarse el pronóstico para emitir alertas por crecida de río, el modelo presenta una exactitud del 63 % pero con un 51 % de falsas alarmas emitidas por aumento de nivel y un acierto del 75 % al indicar que el nivel de río va a permanecer igual o descenderá su nivel.

7. CONCLUSIONES

- De los datos de misiones satelitales disponibles hasta el momento, se encontró que las mejores opciones en cuanto a calidad, cantidad y accesibilidad de los datos, son los productos CHIRPS[3], MSWEP [11] e IMERG [19]. De estas tres opciones, se deben seleccionar los productos preliminares (denominados “Early”) ya que el resto tarda mas de un mes en publicarse.
- La resolución espacial y temporal varía para cada producto satelital, por lo que deben definirse 2 criterios para delimitar la información. El primero es seleccionar únicamente los píxeles que pertenezcan a la cuenca hidrográfica a considerar, el segundo es considerar el tiempo entre registro y publicación de los datos para considerar cual es el último dato disponible a la hora en la que se realizará el pronóstico. Es debido a este segundo criterio que se descarta la información de CHIRPS, ya que al publicarse tres días después de ocurrida la lluvia, presenta baja o nula correlación con los datos de nivel de río a pronosticar.
- De las 18 cuencas hidrográficas de la vertiente del Océano Pacífico, se determinó que la cuenca del río Acomé es apropiada como cuenca de interés, debido a la cantidad y calidad de los datos de nivel de río disponibles para la estación “La Gomera”. Además, dicha estación cuenta con información de otros modelos de RNA con los que ya se realizó pronósticos de nivel de río (Martínez y Pérez [27]), por lo que se pueden comparar los pronósticos realizados utilizando datos de estaciones meteorológicas y los datos satelitales considerados en este trabajo.
- Una buena metodología para descartar información poco relevante para el modelo es considerar los siguientes criterios:
 1. Seleccionar únicamente píxeles que pertenezcan a la cuenca hidrográfica de interés.

2. Seleccionar aquellos píxeles cuyo coeficiente de correlación de Spearman sea mayor a un umbral establecido, siempre y cuando tenga al menos una correlación débil con el horario de nivel de río a pronosticar.
 3. Considerar únicamente horarios donde los píxeles con mayor correlación se ubiquen en la parte alta de la cuenca hidrográfica de interés.
- Para la correcta realización de un pronóstico es necesario contar con al menos un día previo de nivel de río. Puede considerarse una “N” cantidad de días previos, sin embargo, deberá ajustarse según la robustez del sistema de registro, en caso contrario, no se podrá realizar un pronóstico durante “N” días, a menos que se defina una metodología para rellenar datos faltantes sin que afecte significativamente al valor a pronosticar.
 - Se encontró que para la cuenca del río Acomé, es más conveniente utilizar el MAE y la pérdida de Huber con $\gamma = 0.1$ sobre el MSE como función de error. Estas funciones de error permiten obtener buenos resultados en el resto de estaciones en las que se realizaron pronósticos, aunque no se puede asegurar que en esos casos hayan obtenido un mejor rendimiento que utilizar el MSE.
 - La metodología propuesta para construir la geometría de la RNA, así como las funciones de error seleccionadas para cada producto satelital, proveen una buena base para realizar pronósticos en distintas estaciones hidrométricas de la vertiente del Océano Pacífico. Con ellas se pueden obtener buenos resultados en caso de contar con poco tiempo para realizar las variaciones consideradas en este trabajo, pero es recomendable analizar a profundidad el comportamiento de los datos para obtener un mejor rendimiento que se adapte a las características y necesidades de la estación a pronosticar.
 - La metodología propuesta para construir una geometría de RNA puede utilizarse tanto para pronosticar valores numéricos de nivel de río, así como para la emisión de alertas por aumento o disminución de nivel de agua. En caso de ser utilizado para emitir alerta por crecidas, deben calibrarse los umbrales de emisión de alerta para aumentar la exactitud y precisión del pronóstico emitido.
 - Si se almacenan los valores obtenidos del pronóstico diario realizado y se calcula el promedio mensual de nivel de agua en la estación considerada, se obtiene un resultado todavía más cercano al comportamiento real del río, es por ello

que la metodología propuesta podría considerarse para otro tipo de análisis hidrológico, así como para la reconstrucción de datos históricos de nivel de río no registrados.

8. RECOMENDACIONES

Considerando los resultados obtenidos y en busca de poder contar con pronósticos precisos y robustos ante la pérdida de datos, se recomiendan las siguientes acciones que permitan el correcto funcionamiento de la metodología propuesta:

- Si se desea realizar pronósticos de nivel de agua en distintas estaciones hidrométricas de la vertiente del Océano Pacífico de Guatemala, así como para el resto del país u otras regiones, debe considerarse un análisis profundo de los datos disponibles ya que la calidad de los mismos repercutirá en gran manera en los resultados a obtener.
- La información de días previos de nivel de río, así como las variables mensuales construidas, deben ajustarse a la cantidad de datos disponibles y a la presencia de datos faltantes durante el período de entrenamiento. En caso de no presentarse períodos continuos de información, la información perdida será proporcional a la cantidad de meses previos considerados.
- La geometría y parámetros de una RNA son dependientes del problema y para poder obtener mejores resultados en la estación hidrométrica considerada, se recomienda replicar las variaciones realizadas en este trabajo, y de ser posible, variar los otros parámetros que se fijaron al inicio del análisis desarrollado.
- Se recomienda a las instituciones competentes que refuercen los mecanismos de registro y validación, así como el acceso a los datos de nivel de río para una futura implementación de pronóstico diario, ya que actualmente solo se puede realizar con los datos proporcionados bajo solicitud.
- Para la realización de este trabajo se utilizaron los productos preliminares de cada misión satelital los cuales tienen una menor calidad que los productos finales, pero que son útiles al publicarse casi en tiempo real. Por su parte, dichos datos de mejor calidad pueden utilizarse para reconstruir datos históricos de

nivel de río no registrados, utilizando una metodología de pronóstico similar a la propuesta en este trabajo.

BIBLIOGRAFÍA

- [1] Em-dat, the international disaster database, 2023.
- [2] S. Arlot and A. Celisse. A survey of cross-validation procedures for model selection. 2010.
- [3] H. E. Beck, E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R. F. Adler. Mswep v2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society*, 100(3):473–500, Mar. 2019.
- [4] B. Berhanu, Y. Seleshi, S. S. Demisse, and A. M. Melesse. Bias correction and characterization of climate forecast system re-analysis daily precipitation in ethiopia using fuzzy overlay. *Meteorological Applications*, 23(2):230–243, Mar. 2016.
- [5] M. Campolo, P. Andreussi, and A. Soldati. River flood forecasting with a neural network model. *Water Resources Research*, 35(4):1191–1197, Apr. 1999.
- [6] Z. Chen, X. Lin, C. Xiong, and N. Chen. Modeling the relationship of precipitation and water level using grid precipitation products with a neural network model. *Remote Sensing*, 12(7), 2020.
- [7] V. Chow, D. Maidment, and L. Mays. Applied hydrology, tata mcgraw-hill education (india) private limited, 1988.
- [8] J. Cohen. *Statistical power analysis for the behavioral sciences*. Routledge, 2013.
- [9] E. E. Ebert. *Methods for Verifying Satellite Precipitation Estimates*, pages 345–356. Springer Netherlands, Dordrecht, 2007.
- [10] I. Flood and N. Kartam. Neural networks in civil engineering. i: Principles and understanding. *Journal of computing in civil engineering*, 8(2):131–148, 1994.

- [11] C. Funk, P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, and J. Michaelsen. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2(1), Dec. 2015.
- [12] M. Ghiassi, H. Saidane, and D. Zimbra. A dynamic artificial neural network model for forecasting time series events. *International Journal of Forecasting*, 21(2):341–362, Apr. 2005.
- [13] K. Gokcesu and H. Gokcesu. Generalized huber loss for robust learning and its efficient minimization for a robust statistics. *arXiv preprint arXiv:2108.12627*, 2021.
- [14] H. V. Gupta and H. Kling. On typical range, sensitivity, and normalization of mean squared error and nash-sutcliffe efficiency type metrics. *Water Resources Research*, 47(10), Oct. 2011.
- [15] T. Hill, L. Marquez, M. O’Connor, and W. Remus. Artificial neural network models for forecasting and decision making. *International journal of forecasting*, 10(1):5–15, 1994.
- [16] T. O. Hodson. Root-mean-square error (rmse) or mean absolute error (mae): when to use them or not. *Geoscientific Model Development*, 15(14):5481–5487, July 2022.
- [17] T. O. Hodson, T. M. Over, and S. S. Foks. Mean squared error, deconstructed. *Journal of Advances in Modeling Earth Systems*, 13(12), Dec. 2021.
- [18] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, Jan. 1989.
- [19] G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, E. F. Stocker, and D. B. Wolff. The Integrated Multi-satellitE Retrievals for GPM (IMERG): A Brief Overview. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5):2944–2950, 2017.
- [20] N. Karunanithi, W. J. Grenney, D. Whitley, and K. Bovee. Neural networks for river flow prediction. *Journal of Computing in Civil Engineering*, 8(2):201–220, 1994.

- [21] K. Kumar. Machine learning in parameter estimation of nonlinear systems. *arXiv preprint arXiv:2308.12393*, 2023.
- [22] J. D. H. Lalinde, F. E. Castro, J. E. Rodríguez, J. G. C. Rangel, C. A. T. Sierra, M. K. A. Torrado, S. M. C. Sierra, and V. J. B. Pirela. Sobre el uso adecuado del coeficiente de correlación de pearson: definición, propiedades y suposiciones. *Archivos venezolanos de Farmacología y Terapéutica*, 37(5):587–595, 2018.
- [23] D. S. Lawrence. Physical hydrology, waveland press, 2015.
- [24] P. Leahy, G. Kiely, and G. Corcoran. Structural optimisation and input selection of an artificial neural network for river level prediction. *Journal of Hydrology*, 355(1):192–201, 2008.
- [25] H. R. Maier and G. C. Dandy. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling Software*, 15(1):101–124, 2000.
- [26] S. B. Maind, P. Wankar, et al. Research paper on basic of artificial neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(1):96–100, 2014.
- [27] P. Martínez and K. Pérez. Uso de rna para la predicción de nivel de agua en ríos en la vertiente del pacífico de guatemala. Oct. 2023.
- [28] T. Masters. *Practical neural network recipes in C++*. Morgan Kaufmann, 1993.
- [29] K. P. Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [30] X. B. Olabe. Redes neuronales artificiales y sus aplicaciones. *Publicaciones de la Escuela de Ingenieros*, 1998.
- [31] R. P. Palma. Análisis crítico del coeficiente de determinación (r^2), como indicador de la calidad de modelos lineales y no lineales. *Matemática*, 20(2), 2022.
- [32] B. Pan, K. Hsu, A. AghaKouchak, and S. Sorooshian. Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 55(3):2301–2321, Mar. 2019.
- [33] L. F. Restrepo and J. González. De pearson a spearman. *Revista Colombiana de Ciencias Pecuarias*, 20(2):183–192, 2007.

- [34] E. M. Rodríguez. Errores frecuentes en la interpretación del coeficiente de determinación lineal. *Anuario jurídico y económico escurialense*, (38):315–331, 2005.
- [35] M. R. Salguero. Estudio hidrogeológico con fines de riego, de la cuenca del río acomé, escuintla, guatemala. Tesis, 2002.
- [36] SUTRON. Tipping bucket rain gauge 5600-0425, 2023.
- [37] Z. Tang, C. de Almeida, and P. A. Fishwick. Time series forecasting using neural networks vs. box- jenkins methodology. *SIMULATION*, 57(5):303–310, Nov. 1991.
- [38] UPGGR. *Mapa de Cuencas Hidrográficas a Escala 1:50,000. República de Guatemala, Método de Pfafstetter*, volume 1. MAGA, 2009.
- [39] Vaisala. Rain gauge qmr101 and qmr101m, 2021.
- [40] J. D. Velásquez, F. A. Villa, and R. C. Souza. Time series forecasting using cascade correlation networks. *Ingeniería e Investigación*, 30(1):157–162, 2010.
- [41] C. Willmott and K. Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, 30:79–82, 2005.
- [42] M. Zeybek. Nash-sutcliffe efficiency approach for quality improvement. *J. Appl. Math. Comput*, 2:496–503, 2018.

ANEXOS

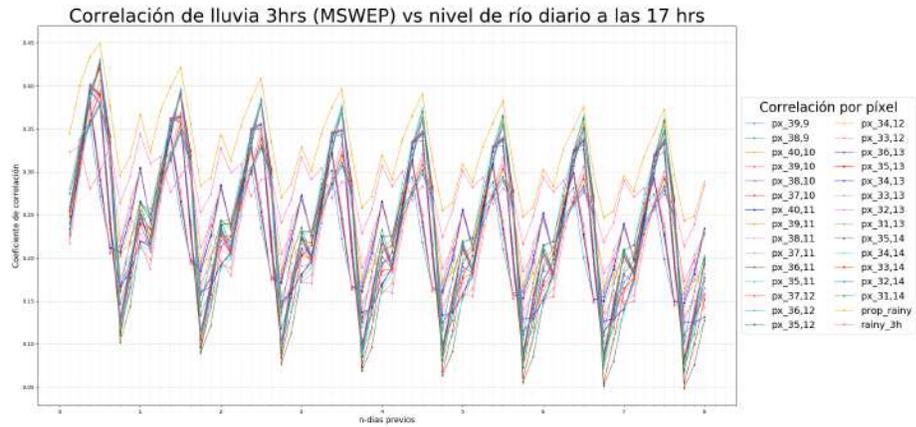


Figura 8.1. Coeficiente de correlación de Spearman para las variables de MSWEP para la estación “Puente Coyolate”

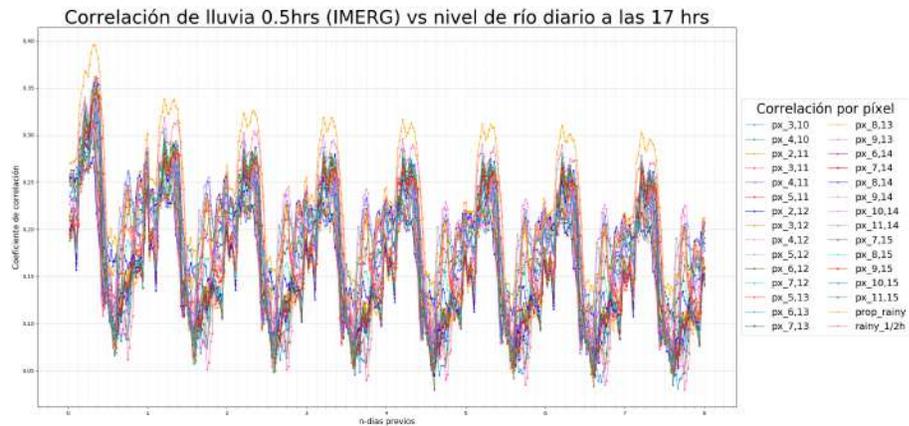
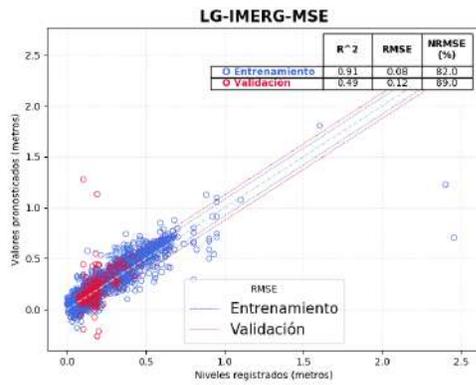
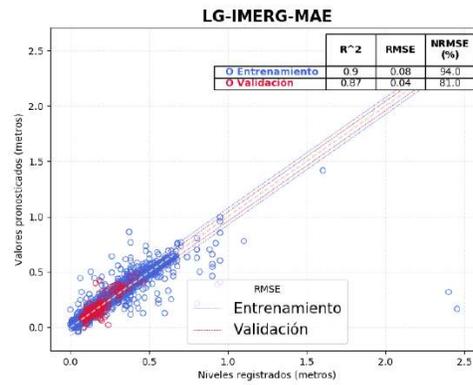


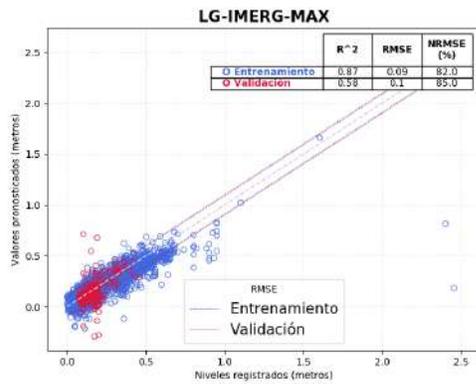
Figura 8.2. Coeficiente de correlación de Spearman para las variables del IMERG para la estación “Puente Coyolate”



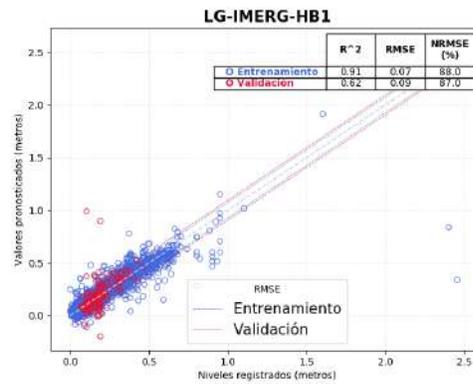
(a) Rendimiento utilizando Error Cuadrático Medio - MSE



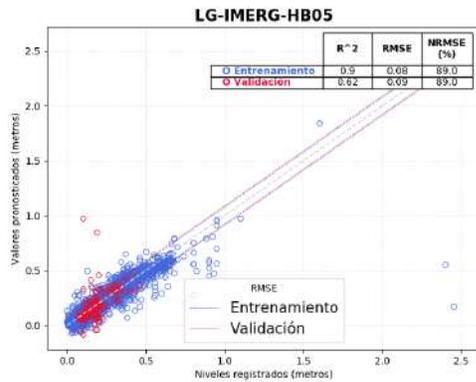
(b) Rendimiento utilizando Error Absoluto Medio - MAE



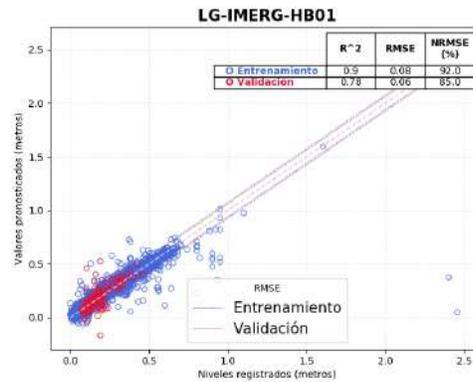
(c) Rendimiento utilizando Error Absoluto Máximo - MaxAE



(d) Rendimiento utilizando Huber Loss, $\alpha = 1.0$



(e) Rendimiento utilizando Huber Loss, $\alpha = 0.5$



(f) Rendimiento utilizando Huber Loss, $\alpha = 0.1$

Figura 8.3. Rendimiento de los modelos utilizando distinta función de error para los datos satelitales del IMERG durante el entrenamiento (azul) y el año de validación (rojo)

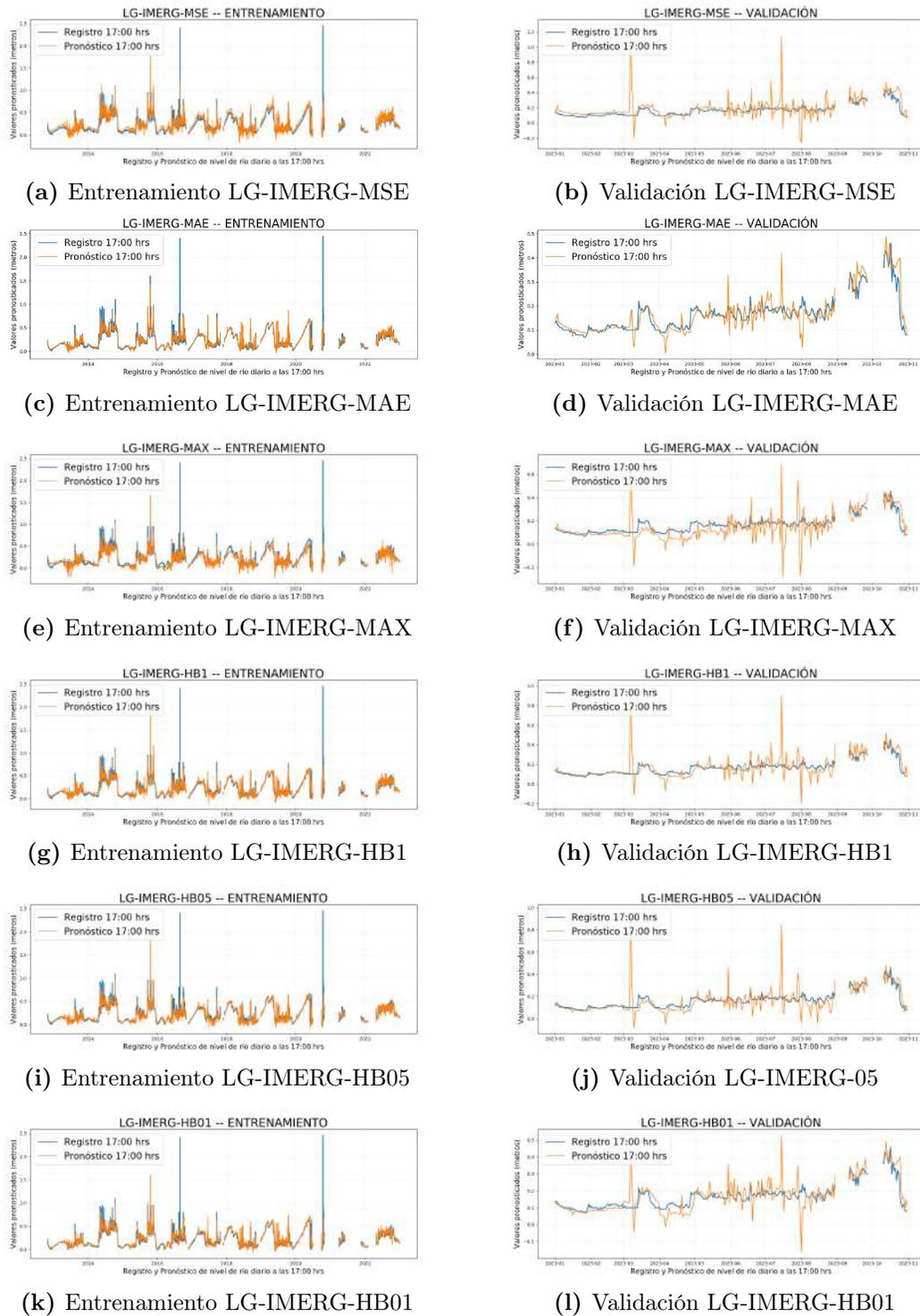
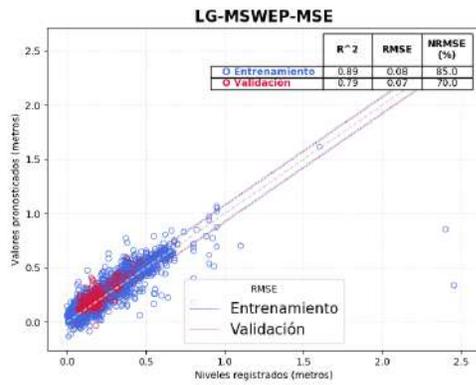
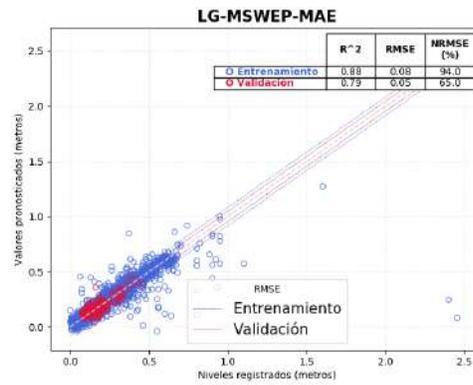


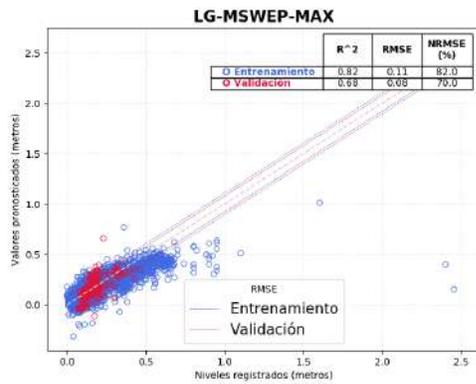
Figura 8.4. Rendimiento de los modelos para el IMERG utilizando distinta función de error para el período de entrenamiento (azul) y el año de validación (naranja)



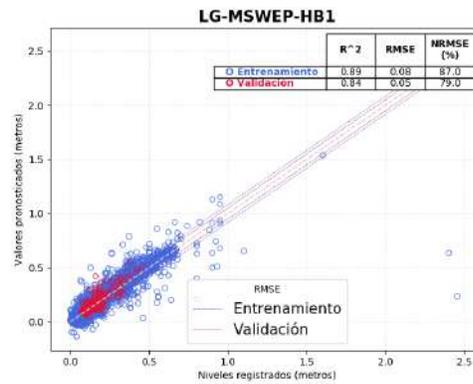
(a) Rendimiento utilizando Error Cuadrático Medio - MSE



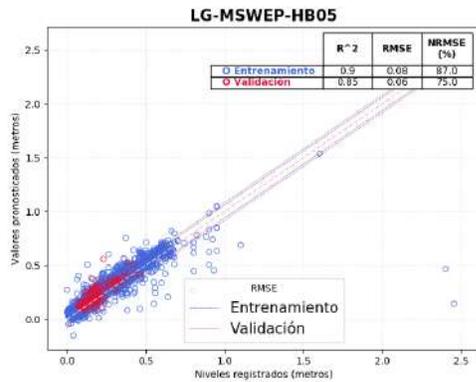
(b) Rendimiento utilizando Error Absoluto Medio - MAE



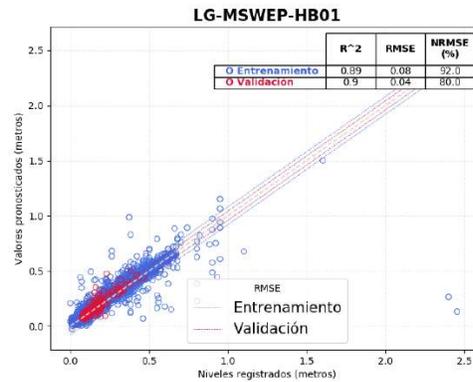
(c) Rendimiento utilizando Error Absoluto Máximo - MaxAE



(d) Rendimiento utilizando Hubber Loss, $\alpha = 1.0$



(e) Rendimiento utilizando Hubber Loss, $\alpha = 0.5$

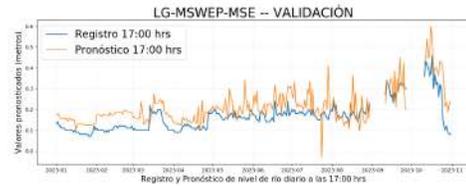


(f) Rendimiento utilizando Hubber Loss, $\alpha = 0.1$

Figura 8.5. Rendimiento de los modelos utilizando distinta función de error para los datos satelitales del MSWEP durante el entrenamiento (azul) y el año de validación (rojo)



(a) Entrenamiento LG-MSWEP-MSE



(b) Validación LG-MSWEP-MSE



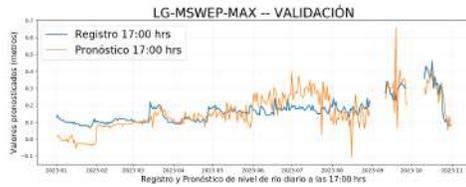
(c) Entrenamiento LG-MSWEP-MAE



(d) Validación LG-MSWEP-MAE



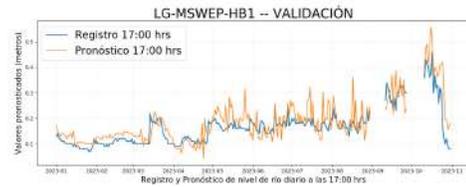
(e) Entrenamiento LG-MSWEP-MAX



(f) Validación LG-MSWEP-MAX



(g) Entrenamiento LG-MSWEP-HB1



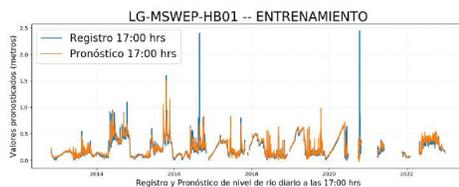
(h) Validación LG-MSWEP-HB1



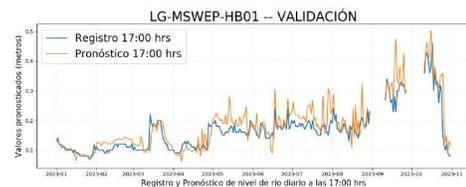
(i) Entrenamiento LG-MSWEP-HB05



(j) Validación LG-MSWEP-05



(k) Entrenamiento LG-MSWEP-HB01



(l) Validación LG-MSWEP-HB01

Figura 8.6. Rendimiento de los modelos para el MSWEP utilizando distinta función de error para el período de entrenamiento (azul) y el año de validación (naranja)

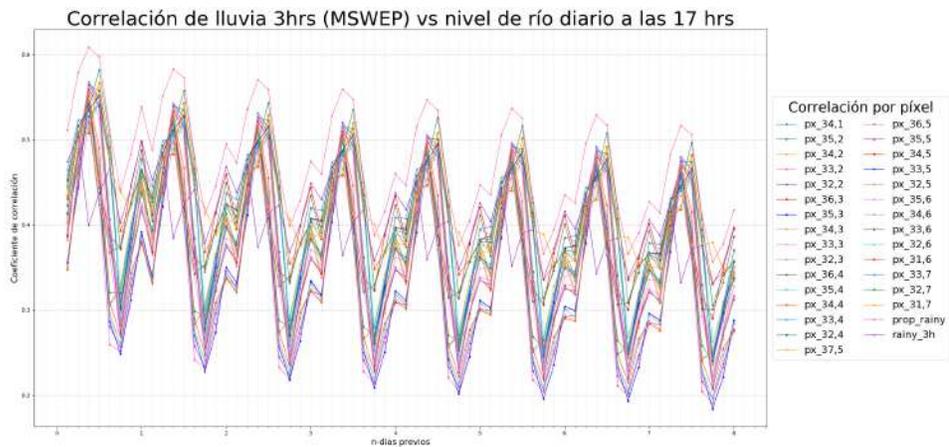


Figura 8.7. Coeficiente de correlación de Spearman para las variables de MSWEP para la estación “Caballo Blanco”

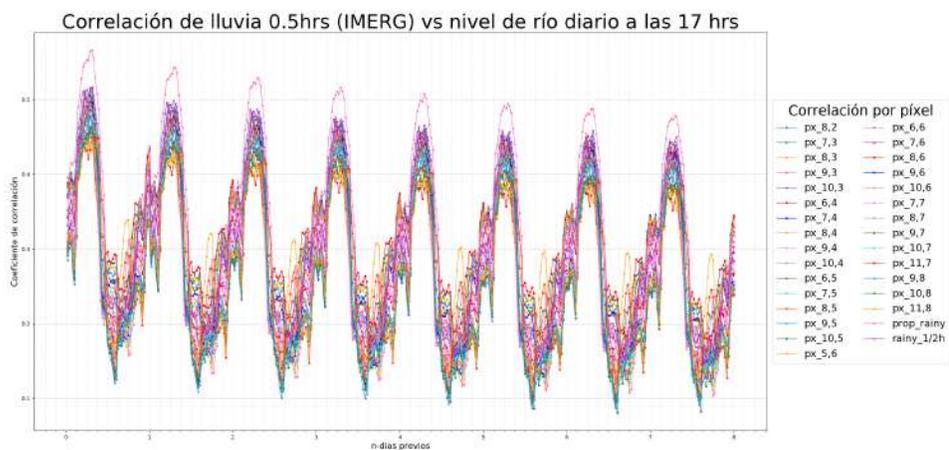


Figura 8.8. Coeficiente de correlación de Spearman para las variables del IMERG para la estación “Caballo Blanco”