



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

**DISEÑO E IMPLEMENTACIÓN DE MODELOS DE INTERVALOS  
NEURONALES PARA SISTEMA DE GESTIÓN DE AGUA-ENERGÍA DE  
CULTIVOS**

MEMORIA PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL ELÉCTRICO

LUCAS FELIPE ROJAS VICTORIANO

PROFESORA GUÍA:  
DORIS SÁEZ HUEICHAPAN

MIEMBROS DE LA COMISIÓN:  
CONSTANZA AHUMADA SANHUEZA  
LINDA DANIELE

Este trabajo ha sido parcialmente financiado por:

ANID/FONDECYT 1220507

SERC-Chile ANID/FONDAP/15110019

Instituto Sistemas Complejos de Ingeniería (ISCI) ANID PIA/PUENTE AFB220003

SANTIAGO DE CHILE

2023

# DISEÑO E IMPLEMENTACIÓN DE MODELOS DE INTERVALOS NEURONALES PARA SISTEMA DE GESTIÓN DE AGUA-ENERGÍA DE CULTIVOS

Esta memoria enfrenta el problema asociado al impacto de las variables climatológicas en un sistema de predicción de variables climáticas, para su uso en un sistema de control de agua y energía en una comunidad lafkenche, teniendo en cuenta la cuantificación de la incertidumbre. El trabajo se divide en tres partes: búsqueda de fuentes de datos, desarrollo de modelos de predicción basado en modelos de redes neuronales y creación de modelos de intervalos de predicción basados en modelos de redes neuronales. Para encontrar las fuentes de datos, se realiza una comparación entre los datos obtenidos de una estación meteorológica *in situ* ubicada en la comunidad y datos externos. Posteriormente, se desarrollan modelos de predicción para predecir los valores futuros de las variables climáticas necesarias para el sistema de control. Se optimiza la arquitectura de los modelos para obtener los mejores resultados posibles. Finalmente, se utilizan estas arquitecturas óptimas para desarrollar modelos de intervalos y lograr la cuantificación de la incertidumbre. Los resultados obtenidos se validan comparándolos con los resultados de otro modelo implementado por otro integrante del proyecto, entrenado con los mismos datos.

# Agradecimientos

Quisiera agradecer a mi familia y mis amigos que me han acompañado en algún momento de este largo camino. Es importante destacar la importancia del apoyo emocional y psicológico que ellos me han brindado a lo largo de mi carrera. La dedicación y el esfuerzo requiere de una gran cantidad de tiempo y energía, y es a menudo un camino lleno de altibajos. Su apoyo fue fundamental para mantener la motivación y la confianza en uno mismo durante los momentos difíciles.

Asimismo, le agradezco a todos los profesores y académicos que me dieron sus sabios consejos, proporcionándome una guía invaluable y una retroalimentación constructiva que ayudó en la mejora de mi trabajo. En particular quisiera agradecerle a la Prof. Dra. Doris Saéz, por haberme brindando apoyo y orientación en cada etapa de la memoria. Sus consejos y su experiencia en el campo han sido de gran ayuda para avanzar en la memoria y completarla con éxito. De igual forma me quiero referir a las Prof. Dra. Constanza Ahumada y Prof. Dra. Linda Danielle, que tuvieron la disposición a ayudarme continuamente durante el proceso mediante todas las reuniones realizadas.

Finalmente, quiero agradecer al *Solar Energy Research Center* SERC-Chile ANID/FONDAP/15110019, al Instituto Sistemas Complejos de Ingeniería (ISCI) ANID PIA/PUENTE AFB220003, y al proyecto ANID/CONICYT/FONDECYT 1220507 "*Distributed Predictive Control Strategies based on Evolving Prediction Intervals for Energy-Water Microgrids*" por su apoyo financiero para el desarrollo de esta memoria y los trabajos de investigación relacionados.

# Tabla de Contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Formulación del problema . . . . .	1
1.2. Hipótesis . . . . .	4
1.3. Objetivos . . . . .	4
1.3.1. Objetivos generales . . . . .	4
1.3.2. Objetivos específicos . . . . .	4
1.3.3. Estructura memoria . . . . .	4
1.3.4. Publicaciones generadas durante esta memoria . . . . .	5
<b>2. Marco Teórico</b>	<b>6</b>
2.1. Modelos de regresión . . . . .	6
2.1.1. Redes Neuronales . . . . .	6
2.1.1.1. Modelo Perceptron Multicapa . . . . .	7
2.1.2. Redes Recurrentes . . . . .	9
2.1.2.1. Modelo <i>Long Short Term Memory</i> (LSTM) . . . . .	9
2.2. Intervalos de predicción . . . . .	11
2.2.1. Métricas de intervalos de predicción . . . . .	11
2.3. Clasificación de métodos de intervalos de predicción . . . . .	12
2.3.1. Métodos secuenciales . . . . .	13
2.3.1.1. Método Delta . . . . .	13
2.3.1.2. Método Bayesiano . . . . .	13
2.3.1.3. Mean-Variance Estimation Method (MVEM) . . . . .	14
2.3.1.4. Método Bootstrap . . . . .	15
2.3.1.5. Método de la Covarianza . . . . .	16
2.3.2. Métodos directos . . . . .	18
2.3.2.1. Lower Upper Bound Estimation (LUBE) . . . . .	18
2.3.2.2. Quality Driven (QD) . . . . .	19
2.3.2.3. Joint Supervision . . . . .	19
2.3.3. Intervalos de predicción basados en modelos de <i>Deep Learning</i> . . . . .	20
2.3.3.1. Cuantización de la incertidumbre mediante <i>Monte Carlo Dropout</i> (MCD) . . . . .	20
2.4. Discusión . . . . .	21
<b>3. Estado del Arte</b>	<b>22</b>
<b>4. Establecimiento de fuente de datos</b>	<b>27</b>
4.1. Datos de la comunidad . . . . .	27
4.2. Búsqueda de fuentes de datos . . . . .	28

4.3.	Selección de datos . . . . .	32
4.3.1.	Precipitaciones . . . . .	32
4.3.2.	Temperatura, radiación solar y velocidad de viento . . . . .	35
4.3.3.	Discusión . . . . .	37
<b>5.</b>	<b>Diseño e implementación de modelos de predicción</b>	<b>38</b>
5.1.	Modelo de predicción de perceptrón multicapa . . . . .	38
5.1.1.	Selección de datos . . . . .	38
5.1.2.	Selección de modelos . . . . .	39
5.1.3.	Metodología utilizada para el entrenamiento del perceptrón multicapa	39
5.1.3.1.	Separación de conjuntos . . . . .	40
5.1.3.2.	Preprocesamiento . . . . .	40
5.1.3.3.	Búsqueda del número de neuronas óptimo . . . . .	40
5.1.3.4.	Optimización del número de regresores . . . . .	40
5.1.4.	Resultados búsqueda inicial de número de neuronas en la red perceptrón multicapa . . . . .	41
5.1.5.	Optimización del número de regresores . . . . .	42
5.1.6.	Optimización final del número de neuronas . . . . .	43
5.2.	Entrenamiento LSTM . . . . .	43
5.3.	Modelo de referencia y validación de resultados . . . . .	44
5.4.	Resultados de predicción . . . . .	44
5.5.	Discusión . . . . .	53
<b>6.</b>	<b>Diseño e implementación de intervalos de predicción</b>	<b>55</b>
6.1.	Modelos de intervalos de predicción neuronales . . . . .	55
6.1.1.	Preprocesamiento . . . . .	55
6.1.2.	<i>JointSupervision</i> (JS) . . . . .	55
6.1.2.1.	Entrenamiento . . . . .	56
6.1.2.2.	Arquitectura . . . . .	56
6.1.3.	<i>Quality Driven</i> (QD) . . . . .	56
6.1.3.1.	Entrenamiento . . . . .	56
6.1.3.2.	Arquitectura del modelo . . . . .	57
6.2.	Implementación de modelos de intervalos neuronales . . . . .	57
6.2.1.	Comparación y validación de resultados . . . . .	57
6.2.2.	Resultados de predicción de intervalos . . . . .	57
6.3.	Discusión . . . . .	68
<b>7.</b>	<b>Conclusiones y trabajo futuro</b>	<b>70</b>
7.1.	Conclusiones . . . . .	70
7.2.	Trabajo Futuro . . . . .	71
	<b>Bibliografía</b>	<b>73</b>
	<b>Anexos</b>	<b>80</b>
A.	Resultados de entrenamiento perceptrón multicapa . . . . .	80
B.	Análisis de sensibilidad . . . . .	83
C.	Resultado de intervalos . . . . .	86

# Índice de Tablas

4.1.	Cantidad de datos disponibles de las distintas fuentes de datos para la radiación solar agregadas de manera diaria. . . . .	31
4.2.	Cantidad de datos disponibles de las distintas fuentes de datos para la temperatura del aire agregadas de manera diaria. . . . .	31
4.3.	Cantidad de datos disponibles de las distintas fuentes de datos para la velocidad de viento agregadas de manera diaria. . . . .	32
4.4.	Cantidad de datos disponibles de las distintas fuentes de datos para la precipitación diaria. . . . .	32
4.5.	Error entre datos de radiación solar agregadas de manera diaria de la comunidad y las demás fuentes de datos (en color azul los errores de menor magnitud). . .	36
4.6.	Error entre datos de temperatura agregadas de manera diaria de la comunidad y las demás fuentes de datos (en color azul los errores de menor magnitud). . .	36
4.7.	Error entre datos de velocidad de viento agregadas de manera diaria de la comunidad y las demás fuentes de datos (en color azul los errores de menor magnitud). . .	37
5.1.	Estadísticas de las fuentes de datos seleccionadas para usarlas en la implementación de los modelos de predicción. . . . .	39
5.2.	Número de neuronas en la capa oculta seleccionadas en la optimización inicial de hiperparámetros, con los regresores originales. . . . .	41
5.3.	Regresores seleccionados en el análisis de sensibilidad para cada uno de los conjuntos (m: denota la cantidad de minutos, h: la cantidad de horas, d: la cantidad de días). . . . .	42
5.4.	Número de neuronas en la capa oculta seleccionadas en la optimización final de hiperparámetros, con los regresores optimizados. . . . .	43
5.5.	Dimensión del estado oculto de la celda LSTM seleccionadas en la optimización. . . . .	44
5.6.	Número de parámetros de los modelos de predicción. . . . .	45
5.7.	Error de predicción de los distintos modelos de predicción de radiación solar a 10 minutos, evaluados en el conjunto de test. . . . .	46
5.8.	Error de predicción de los distintos modelos de predicción de temperatura a 10 minutos, evaluados en el conjunto de test. . . . .	48
5.9.	Error de predicción de los distintos modelos de predicción de radiación solar diaria, evaluados en el conjunto de test. . . . .	50
5.10.	Error de predicción de los distintos modelos de predicción de temperatura diaria, evaluados en el conjunto de test. . . . .	51
5.11.	Error de predicción de los distintos modelos de predicción de precipitación diaria, evaluados en el conjunto de test. . . . .	52
6.1.	Número de parámetros de los modelos de intervalos de predicción. . . . .	58

6.2.	Comparación entre los distintos modelos de intervalos en los datos de radiación solar a 10 minutos, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno). . . . .	60
6.3.	Comparación entre los distintos modelos de intervalos en los datos de temperatura a 10 minutos, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno). . . . .	62
6.4.	Comparación entre los distintos modelos de intervalos en los datos de radiación solar diaria, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno). . . . .	64
6.5.	Comparación entre los distintos modelos de intervalos en los datos de temperatura diaria, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno). . . . .	65
6.6.	Comparación entre los distintos modelos de intervalos en los datos de precipitación diaria, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno). . . . .	67
C.1.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de radiación solar a 10 minutos evaluado en los conjuntos de test . . . . .	86
C.2.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de radiación solar a 10 minutos evaluado en los conjuntos de test . . . . .	87
C.3.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de temperatura a 10 minutos evaluado en los conjuntos de entrenamiento, validación y test. . . . .	88
C.4.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de temperatura a 10 minutos evaluado en los conjuntos de entrenamiento, validación y test. . . . .	89
C.5.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de radiación solar diaria evaluado en los conjuntos de entrenamiento, validación y test. . . . .	90
C.6.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de radiación solar diaria evaluado en los conjuntos de entrenamiento, validación y test. . . . .	91
C.7.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de temperatura diaria evaluado en los conjuntos de entrenamiento, validación y test. . . . .	92
C.8.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de temperatura diaria evaluado en los conjuntos de entrenamiento, validación y test. . . . .	93
C.9.	Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de precipitación diaria evaluado en los conjuntos de entrenamiento, validación y test. . . . .	94

# Índice de Figuras

1.1.	Esquema de control. . . . .	2
2.1.	Modelo del perceptrón [6]. . . . .	7
2.2.	Arquitectura del <i>multiperceptrón</i> [7]. . . . .	8
2.3.	Arquitectura de una celda LSTM [4]. . . . .	10
2.4.	Ejemplo de intervalo de predicción [10]. . . . .	11
2.5.	Ejemplo de dropout en una estructura de una red neuronal. . . . .	21
4.1.	Datos meteorológicos registrados en la comunidad José Painecura, a una tasa de muestreo de 10 minutos. . . . .	27
4.2.	Ubicación de las fuentes de datos encontradas. Se limitan las distintas cuencas hidrográficas de la zona, marcando con rojo la cuenca <i>Costera límite regional y Río Imperial</i> . . . . .	29
4.3.	Precipitación anual acumulada para las fuentes de <i>CR2 Met, Satelital 3 (ERA5) y Puerto Saavedra</i> . . . . .	33
4.4.	Curva doble másica entre Cr2 Met y la estación de Puerto Saavedra. . . . .	34
4.5.	Mapa de elevaciones de la novena región. En un círculo rojo se demarca la ubicación de la comunidad mapuche (Fuente: <a href="https://es-cl.topographic-map.com/maps/6iyt/Chile/">https://es-cl.topographic-map.com/maps/6iyt/Chile/</a> ). . . . .	35
5.1.	Comparación de la predicción de los distintos modelos de predicción de radiación solar de 10 minutos a 144 pasos, evaluados en el conjunto de test. . . . .	46
5.2.	Comparación de la predicción de los distintos modelos de predicción de temperatura de 10 minutos a 144 pasos, evaluados en el conjunto de test. . . . .	48
5.3.	Comparación de la predicción de los distintos modelos de predicción de radiación solar diaria a 14 pasos, evaluados en el conjunto de test. . . . .	49
5.4.	Comparación de la predicción de los distintos modelos de predicción de temperatura diaria a 14 pasos, evaluados en el conjunto de test. . . . .	50
5.5.	Comparación de la predicción de los distintos modelos de predicción de precipitación diaria a 14 pasos, evaluados en el conjunto de test. . . . .	52
6.1.	Intervalos de predicción para los datos de radiación solar de 10m a 144 horas (PM: Perceptrón multicapa). . . . .	59
6.2.	Intervalos de predicción para los datos de temperatura de 10m a 144 horas (PM: Perceptrón multicapa). . . . .	61
6.3.	Intervalos de predicción para los datos de radiación solar diaria a 14 días (PM: Perceptrón multicapa). . . . .	63
6.4.	Intervalos de predicción para los datos de temperatura diaria a 14 días (PM: Perceptrón multicapa). . . . .	66
6.5.	Intervalos de predicción para los datos de precipitación diaria a 14 días (PM: Perceptrón multicapa). . . . .	68



A.1.	Promedio de RMSE de la validación cruzada en los datos de radiación solar a 10 minutos, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores. . . . .	80
A.2.	Promedio de RMSE de la validación cruzada en los datos de temperatura a 10 minutos, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores. . . . .	81
A.3.	Promedio de RMSE de la validación cruzada en los datos de radiación solar diaria, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores. . . . .	81
A.4.	Promedio de RMSE de la validación cruzada en los datos de temperatura diaria, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores. . . . .	82
A.5.	Promedio de RMSE de la validación cruzada en los datos de precipitación diaria, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta. . . . .	82
B.1.	Gradientes absolutos promedio de temperatura cada 10 minutos para cada uno de los regresores evaluados en el conjunto de entrenamiento. . . . .	83
B.2.	Gradientes absolutos promedio de radiación solar cada 10 minutos para cada uno de los regresores evaluados en el conjunto de entrenamiento. . . . .	83
B.3.	Gradientes absolutos promedio de temperatura diaria para cada uno de los regresores evaluados en el conjunto de entrenamiento. . . . .	84
B.4.	Gradientes absolutos promedio de radiación solar diaria para cada uno de los regresores evaluados en el conjunto de entrenamiento. . . . .	84
B.5.	Gradientes absolutos promedio de precipitación diaria para cada uno de los regresores evaluados en el conjunto de entrenamiento. . . . .	84

# Capítulo 1

## Introducción

### 1.1. Formulación del problema

Actualmente Chile está atravesando por una de las peores sequías desde que se tiene registro, con una variación en los patrones de precipitaciones que están tributando a problemas de cantidad y calidad que se hacen especialmente severos en las zonas rurales. Esto ha generado numerosas consecuencias negativas y es imperativo tomar medidas para contrarrestar sus efectos [1].

Sumado a esto se encuentran los efectos de una mala gestión del agua por parte de las entidades correspondientes y una desventaja económica en el acceso al agua. Esto da a lugar una extracción superior a la sostenible, provocando un agotamiento del recurso [2].

La combinación de ambos fenómenos inside de manera catastrófica en la vida de muchos pobladores que necesitan el agua para distintas necesidades. En Chile, una de estas consecuencias ha sido la disminución de la disponibilidad de agua que ha afectado principalmente a las zonas rurales en los últimos años, tanto para el consumo humano como para el riego de cultivos. Históricamente, el riego artificial no había sido necesario en el sur del país, debido a que las lluvias estaban bien distribuidas a lo largo del año. Sin embargo, esta tendencia se rompió en el período 2010-2018, que fue el más seco registrado en la zona, con un déficit de precipitaciones del 20-40 % [3].

El presente informe se enfoca en la comunidad mapuche lafkenche José Painecura Hueñalihuen, ubicada en la comuna de Carahue, en la IX región de Chile. Esta comunidad se encuentra en un terreno costero montañoso, donde los cultivos se desarrollan en las colinas y el agua debe ser bombeada desde los acuíferos para regarlos. Actualmente, la comunidad cuenta con un invernadero y se busca implementar un sistema de control de agua y energía (EWMS) para optimizar el uso de estos recursos, especialmente en el riego de los cultivos. El esquema del sistema se presenta en la Figura 1.1.

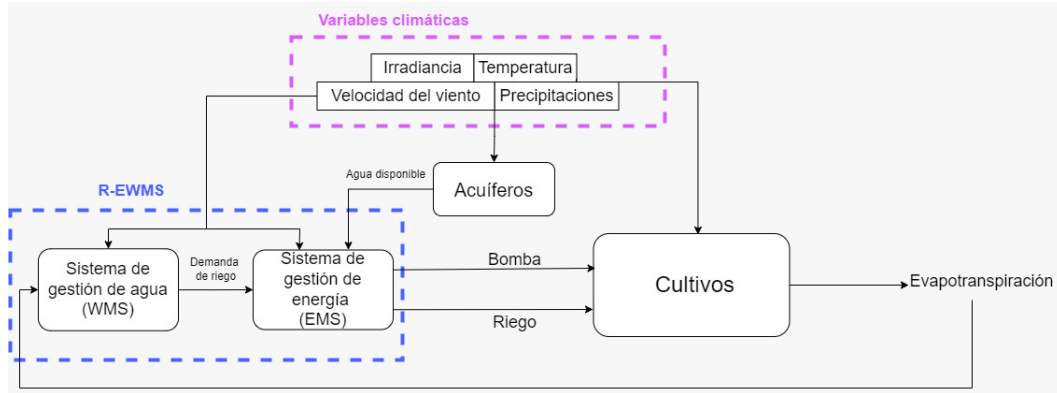


Figura 1.1: Esquema de control.

Como se aprecia en la figura 1.1 este sistema de control consta de dos etapas: *Water Management System (WMS)* y *Energy Management System (EMS)*, que se encargan del manejo del agua y la energía, respectivamente, para optimizar el cultivo de diferentes plantaciones considerando los recursos hídricos disponibles, la eficiencia energética y económica.

El objetivo del WMS es controlar el uso del agua de los estanques y proporcionar instrucciones diarias al EMS sobre las necesidades de agua para cada sistema hídrico (pozo, bomba, estanque). En este sistema de control, la optimización se realiza mediante la predicción de un horizonte de 28 días con un tiempo de muestreo diario. Esta elección se debe a que los efectos del riego de los cultivos pueden observarse semanas después, ya que las plantaciones siguen un ciclo biológico lento. Para este sistema, se requieren predicciones de variables meteorológicas como radiación solar, temperatura ambiente del aire, humedad relativa, precipitaciones y velocidad del viento.

Por otro lado, el EMS proporciona instrucciones para la compra y venta de energía, la operación del banco de baterías y las cargas controlables, así como la consigna de volumen de riego. En este caso, se utiliza un tiempo de muestreo de 10 minutos y se optimiza mediante la predicción de un horizonte de 48 horas. La microred utiliza varios recursos renovables, como la energía solar y eólica, que dependen de la radiación solar y la velocidad del viento, siendo esta última estocástica. El horizonte de predicción de 48 horas se debe a la dinámica de las baterías. Para el EMS, se requieren predicciones de variables meteorológicas como radiación solar y temperatura del aire.

En la comunidad, hay una estación meteorológica que registra datos meteorológicos de radiación solar, temperatura del aire y velocidad del viento cada 10 minutos. Sin embargo, debido a problemas técnicos, hay periodos de muestreo sin datos, además de otros corruptos. Por esta razón, se evalúa la necesidad de buscar otras fuentes para el entrenamiento de los modelos.

Esta memoria se centrará en las variables de entrada necesarias para un controlador de gestión de agua y energía, cuya manipulación depende de datos meteorológicos. Estos datos incluyen predicciones de radiación solar y temperatura del aire a 48 horas, y predicciones de radiación solar, temperatura del aire, velocidad del viento y precipitaciones a 28 días. Además, se busca caracterizar la incertidumbre en la predicción de estas variables para ante-

ponerse a situaciones adversas, utilizando modelos de predicción para redes neuronales que proporcionan un rango donde se encuentra el valor esperado con cierta probabilidad.

Por lo tanto, en esta memoria se ha desarrollado un sistema de intervalos de predicción para las variables climáticas, que se utilizará como entrada en un controlador de gestión agua-energía existente. Los resultados obtenidos de esta memoria contribuirán al diseño mejorado del controlador y permitirán ahorros significativos de energía, agua y costos para la comunidad. Esto se debe a que una mejor predicción de las variables climáticas permitirá que el controlador tenga un menor margen de error, lo que resultará en un uso más eficiente y sostenible de los recursos disponibles.

La metodología empleada en esta memoria comienza con un análisis de los datos disponibles de la comunidad y una búsqueda exhaustiva de datos de variables climáticas relevantes como radiación solar, temperatura del aire y velocidad del viento. Las fuentes de datos incluyen sensores, datos satelitales y modelos estadísticos que hacen uso de datos satelitales y/o de sensores. Luego se lleva a cabo una comparación entre estos datos y los datos de la comunidad para establecer cuáles son los más adecuados para su uso en los modelos implementados.

Posteriormente, se implementan modelos de predicción en conjuntos de datos de radiación solar diaria y registrada cada 10 minutos, temperatura del aire diaria y registrada cada 10 minutos y precipitación diaria. Se utilizan dos tipos de modelos de predicción: el perceptrón multicapa y la red LSTM, aunque en el caso de la precipitación diaria solo se utiliza el perceptrón multicapa. Se llevan a cabo optimizaciones de hiperparámetros, entrenamiento y se comparan los resultados obtenidos, además de compararlos con un modelo difuso entrenado con los mismos datos.

Finalmente, se implementan modelos de intervalos de predicción utilizando los modelos de *Joint Supervision* y *Quality Driven* basados en los modelos de perceptrón multicapa y LSTM, lo que da lugar a cuatro tipos de modelos en total. Los modelos base de perceptrón multicapa y LSTM utilizan los mismos hiperparámetros encontrados en los modelos de predicción. Se comparan los resultados entre los modelos y también se considera el desempeño de un modelo de intervalos de predicción difuso entrenado con los mismos datos.

## 1.2. Hipótesis

En esta sección se describen las hipótesis que se hacen a lo largo de la memoria. La propuesta de esta memoria se centra en la implementación de modelos de predicción de intervalos que puedan predecir valores futuros y la incertidumbre asociada a variables climatológicas. Teniendo esto en consideración, las hipótesis planteadas en este trabajo son las siguientes:

1. Hay variables climatológicas que presentan una mayor estacionalidad, lo que se espera se obtendrán mejores resultados de predicción que las variables más estocásticas.
2. La incertidumbre aumenta a medida que se predice a un mayor número de pasos.
3. Implementar un modelo de predicción de Deep Learning, como lo es el modelo LSTM, obtendrá mejor desempeño que otros modelos.

## 1.3. Objetivos

A continuación se enuncian el objetivos generales y los objetivos específicos de la memoria.

### 1.3.1. Objetivos generales

El objetivo general de la memoria es evaluar un sistema de predicción de variables climatológicas (radiación solar, temperatura del aire, precipitaciones y humedad relativa) tomando en cuenta la incertidumbre de estas, para ser utilizadas como entrada a un controlador EWMS de una microrred basada en la comunidad José Painecura. Para esto, se diseña e implementan modelos de predicción de intervalos neuronales para el corto y mediano plazo del controlador.

### 1.3.2. Objetivos específicos

Los objetivos específicos de la memoria son:

1. Analizar los datos de la estación meteorológica de la comunidad mapuche José Painecura Hueñalihuen y evaluar otras fuentes de datos, a partir de estaciones meteorológicas y modelos mesoescalares, para validar y comparar con los obtenidos en la estación de la comunidad.
2. Diseñar e implementar algoritmos de predicción clásicos y de *deep learning* basados en modelos neuronales.
3. Diseñar e implementar intervalos de predicción basados en algoritmos de redes neuronales clásicos y del estado del arte.
4. Validar el desempeño de los algoritmos en un conjunto de test y evaluar el desempeño de los modelos implementados.

### 1.3.3. Estructura memoria

La estructura de este trabajo de tesis se divide en los siguientes 6 capítulos:

**Capítulo 1:** Se presenta la introducción de la memoria, así como la motivación, la formulación del problema, hipótesis y objetivo general y específicos.

**Capítulo 2:** Se desarrolla el Marco Teórico de la investigación. Se describen los fundamentos de: Modelos de Regresión, Intervalos de Predicción, Clasificación de métodos de intervalos de predicción, concluyendo con la discusión de los contenidos.

**Capítulo 3:** Describe el Estado del Arte de los distintos modelos de predicción y modelos de intervalos de predicción, enfocándose en aquellos basados en modelos neuronales.

**Capítulo 4:** Se detallan y establece la fuente de datos. Esto es: los datos de la comunidad, la búsqueda de fuente de datos y la selección de datos.

**Capítulo 5:** Se describe la etapa de diseño y implementación de modelos de predicción. Específicamente se trata el Modelo de Predicción de Perceptrón Multicapa, entrenamiento LSTM, Modelo de referencia y validación de resultados. Finalizando con los resultados de la predicción, y su posterior discusión.

**Capítulo 6:** Describe el diseño e implementación de intervalos de predicción. Ahonda en los Modelos de intervalos de predicción neuronales y en la Implementación de modelos de intervalos neuronales. Termina con la sección de discusión de todo lo expuesto en el capítulo.

**Capítulo 7:** Este capítulo entrega las conclusiones del trabajo y propone trabajos futuros basados en la presente memoria.

#### 1.3.4. Publicaciones generadas durante esta memoria

1. Ocaranza, J., Cartagena O., Sáez D., Daniele L., Ahumada C. y Rojas L., “Robust Energy-Water Management System with Prediction Interval based on Deep Learning,” Enviado a una conferencia internacional *International Joint Conference in Neural Networks 2023 (IJCNN 2023)*.

# Capítulo 2

## Marco Teórico

A continuación se describe el marco teórico que permite dar los fundamentos teóricos de esta memoria. Primero se introducen los modelos de predicción, en particular los modelos de redes neuronales. Se inicia explicando el modelo más básico para luego dar cabida a modelos más complejos con arquitecturas más idóneas a problemas de series temporales. Luego se explica acerca de los intervalos de predicción y su importancia, detallando varias implementaciones.

### 2.1. Modelos de regresión

En esta sección se explica la teoría de algunos de los modelos de regresión más usados.

Los modelos de regresión son aquellos capaces de encontrar una función de correspondencia entre variables de entrada, o regresores, y una salida de carácter continuo, o predictores. Uno de los modelos de regresión más usados son las redes neuronales con sus distintas variantes [4].

A continuación se explica la teoría de las redes neuronales:

#### 2.1.1. Redes Neuronales

Una red neuronal artificial (RNA) es un modelo matemático inspirado en el comportamiento biológico de las neuronas y en cómo se organizan formando la estructura del cerebro [5]. Así como las neuronas están interconectadas en el cerebro, esta arquitectura se basa en las interconexiones de un componente mínimo: el perceptrón multicapa.

El perceptrón es el análogo a una neurona del cerebro en un modelo computacional. Recibe varias señales de entrada  $x_i$  procedentes de otras neuronas, y cada señal es ponderada por un peso  $w_i$  que depende de la importancia de la señal para el perceptrón. Además de las entradas ponderadas, cada neurona tiene un término propio llamado sesgo o *bias*, que indica su nivel de susceptibilidad a activarse. En la práctica, el sesgo se trata como otra componente del peso  $w$ , y la entrada que se pondera por este peso es un 1 por medio de cálculo vectorial.

La suma de estas señales son procesadas a través de una función de activación, indicando la salida. Tal como se aprecia en la figura 2.1:

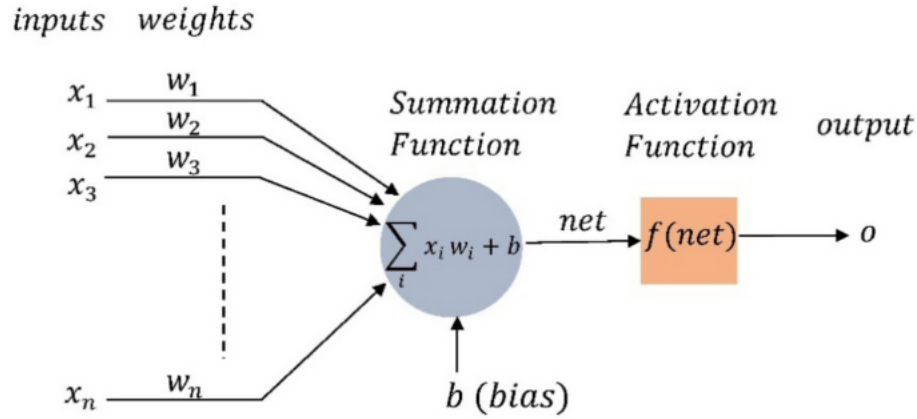


Figura 2.1: Modelo del perceptron [6].

En donde la salida es de la forma:

$$o = f \left( \sum_{i=0}^N x_i \cdot w_i + b \right) \quad (2.1)$$

El objetivo del uso de un sólo perceptrón para clasificación es lograr una separación lineal del espacio de características. Sin embargo, el cerebro no funciona de esta manera, sino que tiene millones de neuronas interconectadas entre sí, lo que inspiró al modelo de perceptrón multicapa.

#### 2.1.1.1. Modelo Perceptron Multicapa

En particular esta arquitectura se llama *Multi-layer Perceptron* (MLP) o *Feedforward Neural Net* (FFNN). Es una generalización del modelo de perceptrón.

En la figura 2.2 se puede observar que este modelo consta de 3 capas: la capa de entrada, las capas ocultas y la capa de salida. La capa de entrada consiste simplemente en el vector de entradas. En las capas ocultas, los perceptrones de una misma capa no están conectados entre sí, pero sí están conectados con cada perceptrón de la capa anterior y con cada perceptrón de la capa siguiente, mediante los pesos respectivos explicados anteriormente. Finalmente, la capa de salida para el caso de clasificación usa la función de activación *softmax* y tiene un nodo (perceptrón) por cada clase. Para el caso de regresión, se suele utilizar la función sigmoidea en cada neurona de salida.



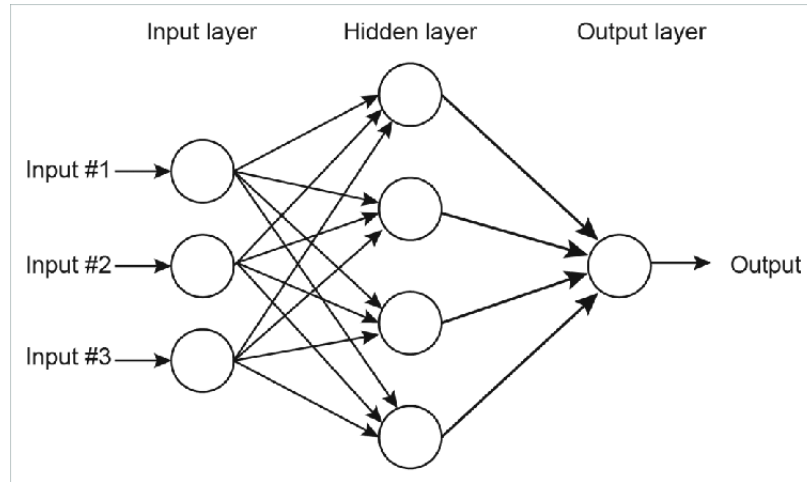


Figura 2.2: Arquitectura del *multiperceptrón* [7].

Hay distintas funciones de activaciones que se utilizan, las más comúnmente usadas son:

- Sigmoidea:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

- Función tangente hiperbólica (tanh):

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (2.3)$$

La diferencia radica en el rango en que se mueven, la sigmoidea entre 0 y 1, mientras que la tanh entre -1 y 1.

Para el entrenamiento de este tipo de red existen varios métodos. Uno de los más usados es el *Backpropagation* con el *Descenso del Gradiente*.

*Backpropagation* hace uso de la regla de la cadena para poder calcular la derivadas de la función de costo con respecto a los parámetros [8] y *Descenso del gradiente* utiliza estas derivadas para actualización de los pesos de la red, logrando el aprendizaje. Para un estudio en más detalle de su funcionamiento, se puede consultar [9]. El proceso de entrenamiento corresponde al ajuste de los componentes de  $w$  para mejorar la precisión del modelo. En este caso, los valores de los pesos se van ajustando de manera iterativa. Asumiendo una sola capa oculta, las ecuaciones de entrenamiento son de la siguiente forma:

$$\Delta w_{ij} = -1 \cdot \eta \cdot \frac{\partial E}{\partial w_{ij}} \quad (2.4)$$

$$w_{ij} := w_{ij} + \Delta w_{ij} \quad (2.5)$$

Cada peso  $w_{ij}$  representa la conexión entre la neurona  $i$  de la capa de entrada y la neurona  $j$  de la capa oculta. El término  $E$  representa el error de salida y  $\eta$  es la tasa de aprendizaje, que determina la magnitud del ajuste de los pesos en cada iteración del entrenamiento. El proceso de aprendizaje se realiza a través de varias iteraciones, en las cuales el error tiende a disminuir.

Es importante tener en cuenta que, si se realizan demasiadas iteraciones, la red neuronal

puede aprender tanto la dinámica de la señal como el ruido presente en los datos de entrenamiento, lo que se conoce como sobreajuste. Para evitar este problema, es necesario contar con un conjunto de validación, el cual permite evaluar el desempeño de la red en datos no utilizados en el entrenamiento. En un principio, el error en el conjunto de validación disminuye junto con el error en el conjunto de entrenamiento, pero en un momento dado el error en el conjunto de validación comienza a aumentar. Para evitar este aumento, se pueden utilizar técnicas como el *early stopping*.

## 2.1.2. Redes Recurrentes

Las redes neuronales recurrentes (RNN) son un tipo de red neuronal en donde la salida en un momento  $t$  pasa a ser la entrada en el momento  $t + 1$  logrando capturar dependencias temporales [4].

Su expresión matemática es la siguiente:

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b) \quad (2.6)$$

$$y_t = h_t \quad (2.7)$$

donde  $x_t$  es la entrada,  $h_t$  el estado recurrente de la red y  $y_t$  la salida en el tiempo  $t$ ;  $W_h$  y  $W_x$  son los pesos; y  $b$  es el sesgo.

Un problema que tienen las formas más básicas de este tipo de redes es que no pueden manejar dependencias temporales largas. Esto se debe a los problemas de desvanecimiento del gradiente y la explosión del gradiente. Estos se presentan a la hora entrenar la red mediante el método de *backpropagation* en donde en este caso se hace a lo largo del tiempo. El problema recae en que, al calcular el gradiente del error, este tiende a desvanecerse o explotar, afectando el aprendizaje de las dependencias de largo plazo.

### 2.1.2.1. Modelo *Long Short Term Memory* (LSTM)

El modelo *Long Short Term Memory* (LSTM) es un tipo de RNN que fue propuesto para resolver los problemas de desvanecimiento del gradiente y la explosión del gradiente [4]. Se logró mejorar la capacidad de aprendizaje de largo plazo mediante la inclusión de compuertas a la celda. Existen distintas variaciones en la arquitectura de la celda, en donde la más usada consta de 3 compuertas distintas: la compuerta de entrada, la compuerta de salida y la compuerta de olvido. Esta arquitectura se observa en la figura 2.3.

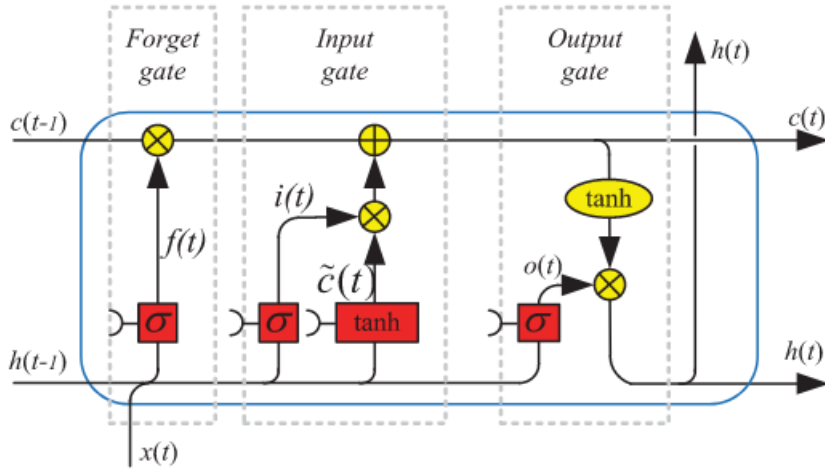


Figura 2.3: Arquitectura de una celda LSTM [4].

Las ecuaciones matemáticas de este modelo se enuncian a continuación:

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \quad (2.8)$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \quad (2.9)$$

$$\tilde{c}_t = \tanh(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}) \quad (2.10)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (2.11)$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o) \quad (2.12)$$

$$h_t = o_t \cdot \tanh c_t \quad (2.13)$$

en donde:  $\sigma$  es la función sigmoidea,  $\tanh$  la función tangente hiperbólica,  $c_t$  l estado de la celda LSTM;  $W_{\tilde{c}x}$ ,  $W_{\tilde{c}h}$ ,  $W_{oh}$ ,  $W_{ox}$  son pesos de la red y  $\cdot$  denota el producto punto entre vectores.

Cuando se actualiza el estado de la celda, la compuerta de entrada  $i_t$  puede decidir qué información nueva será guardada en este estado. La compuerta del olvido  $f_t$  puede decidir qué información olvidar del estado de la celda. Cuando el valor de la compuerta de olvido es 1, mantiene toda la información, mientras que cuando el valor es cero, olvida todo. La compuerta de salida  $o_t$  puede decidir que información puede entregar como salida basado en la información del estado de la celda [4].

Con estas compuertas la red es capaz de aprender dinámicas más complejas, aunque esto viene acompañado de un aumento sustancial del número de parámetros, lo que hace necesario un mayor número de datos de entrenamiento.

## 2.2. Intervalos de predicción

A diferencia de la predicción puntual común, en donde solo se calcula el valor predicho para un tiempo futuro puntual, los intervalos de predicción (IP) entregan un rango en donde se encuentra la predicción con cierta probabilidad [10]. Esto se ilustra en la 2.4 figura:

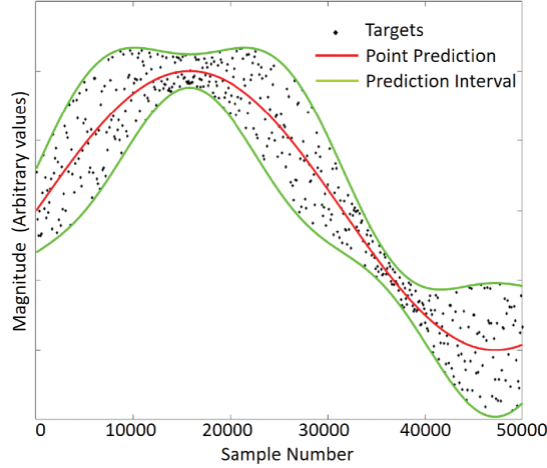


Figura 2.4: Ejemplo de intervalo de predicción [10].

Las bandas representan el intervalo de predicción, que varía para cada predicción puntual. Un intervalo más angosto significa menos incertidumbre en la predicción, por otra parte es importante considerar la cantidad de datos que se encuentran en dicha banda. Se busca maximizar el número de predicciones dentro del intervalo de predicción, además de que sea lo más acotado posible. La forma de calcular la predicción  $\hat{y}$  y los intervalos superior  $\bar{y}$  e inferior  $\underline{y}$ , dependerá del método que se utilice.

### 2.2.1. Métricas de intervalos de predicción

En la literatura han sido utilizadas distintas métricas para comparar el desempeño de los intervalos de predicción [11]. El error de predicción del modelo es uno de los factores más importante que afecta el comportamiento de la banda de incertidumbre. Para poder representar la precisión del modelo predictivo, se usa como métrica el error cuadrático medio (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(z_k))^2} \quad (2.14)$$

en donde  $N$  es el número de datos,  $y(k)$  es a medida real de la salida del sistema y  $\hat{y}(z_k)$  es la predicción cuando es recibido como vector de entrada  $z_k$ .

Otra métrica usualmente utilizada es el Error absoluto medio (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{k=1}^N |y(k) - \hat{y}(k)| \quad (2.15)$$

Con respecto a los intervalos de predicción, las métricas más importantes son el ancho del

intervalo y los niveles de cobertura.

En [12],[13], el ancho promedio normalizado del intervalo de predicción (PINAW) y la probabilidad de cobertura del intervalo de predicción (PICP) son definidos de la siguiente manera:

$$\text{PINAW} = \frac{1}{NR} \sum_{k=1}^N (\bar{y}(z_k) - \underline{y}(z_k)) \quad (2.16)$$

en donde,

$$R = \max\{y(k)\} - \min\{y(k)\} \quad (2.17)$$

es el rango de valores del conjunto de entrenamiento,  $\bar{y}(z_k)$  intervalo superior,  $\underline{y}(z_k)$  es el intervalo inferior,  $N$  es el número de datos, y

$$\text{PICP} = \frac{1}{N} \sum_{k=1}^N c_k \quad (2.18)$$

en donde,

$$c_k = \begin{cases} 1, & \text{si } \underline{y}(z_k) \leq y_k \leq \bar{y}(z_k) \\ 0, & \text{si no} \end{cases} \quad (2.19)$$

indica si el intervalo  $[\underline{y}(z_k), \bar{y}(z_k)]$  contiene al dato medido  $y(k)$ .

Lo ideal es tener un ancho de intervalo pequeño, indicando poca incertidumbre, junto a una cobertura grande, indicando que las predicciones se encuentran en el intervalo con una alta probabilidad.

## 2.3. Clasificación de métodos de intervalos de predicción

Los métodos de intervalos de predicción pueden clasificarse de distintas maneras: en [11] lo clasifican según dos criterios principales: el tipo de modelo usado para describir el intervalo de predicción y el tipo de entrenamiento para su construcción.

Dentro de los tipos de modelos usados, se puede encontrar los modelos basados en lógica difusa y los modelos basados en redes neuronales.

Considerando el tipo de entrenamiento, la primera categoría corresponde a los métodos secuenciales, los cuales reciben un modelo base identificado previamente que estima  $f(x)$  y lo utiliza para construir un IP. La segunda categoría, denominado métodos directos, realizan la construcción de los modelos de intervalo en paralelo con la identificación del modelo de predicción. Por lo tanto, para este caso no es necesario un modelo previamente identificado.

En este tipo de métodos, el entrenamiento del modelo incorpora los cálculos para estimar la banda de incertidumbre de las predicciones, de forma tal que, los modelos obtenidos son capaces de entregar como salida, directamente, los intervalos de predicción sin la necesidad de pasos adicionales.

### 2.3.1. Métodos secuenciales

Los métodos secuenciales son los que necesitan de un modelo de predicción entrenado previamente.

#### 2.3.1.1. Método Delta

Tal como se señala en [11], este método usa teoría de regresión no lineal clásica para obtener una aproximación de la incertidumbre de predicción [14].

Esto se puede hacer mediante el cálculo de las expansiones de las series de Taylor de la función de costo de error cuadrático medio de la red. Luego, bajo la hipótesis de que la varianza del ruido es homogénea y Gaussiana, la varianza total de predicción para cada predicción  $\sigma_{tot}^2(z_k)$  puede ser representada como:

$$\sigma_{tot}^2(z_k) = \sigma_\epsilon(1 + g_o^T(J^T(z_k)J(z_k))^{-1}g_o) \quad (2.20)$$

en donde  $g_o$  representa los parámetros de la red neuronal (pesos y sesgos),  $J(z_k)$  es la matriz Jacobiana de la red evaluada en el dato de entrada  $z_k$ , y  $\sigma_\epsilon$  es la varianza del ruido de los datos, el cual puede ser obtenido a partir de la siguiente expresión:

$$\sigma_\epsilon = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}(z_k))^2 \quad (2.21)$$

en donde  $y_k$  representa el dato de entrenamiento para la muestra  $k$ ,  $\hat{z}_k$  es la predicción del modelo para el dato de entrada  $z_k$ .

Luego de calcular este valor, se puede construir el intervalo de confianza de intervalo de  $(1 - \alpha)\%$ , asumiendo que es simétrica y Gaussiana:

$$\bar{y}(z_k) = \hat{y}(z_k) + t_{n-p}^{1-\frac{\alpha}{2}} \sigma_{tot}(z_k) \quad (2.22)$$

$$\underline{y}(z_k) = \hat{y}(z_k) - t_{n-p}^{1-\frac{\alpha}{2}} \sigma_{tot}(z_k) \quad (2.23)$$

en donde  $n$  es el número de datos de entrada que fueron usadas para entrenar el modelo,  $p$  es el número de parámetros de la red, y  $t_{n-p}^{1-\frac{\alpha}{2}}$  representa el cuantil  $\frac{\alpha}{2}$  de la función acumulativa de la distribución t-Student con  $n - p$  grados de libertad.

Debido a la hipótesis de que la incertidumbre es homogénea y Gaussiana, este método puede fallar en producir intervalos de predicciones de calidad en donde el ruido de los datos varía con el tiempo.

Por otra parte, dado que este método requiere el cálculo de la matriz Jacobiana del modelo, el cálculo del intervalo de predicción puede ser computacionalmente costoso, por lo que en ocasiones se considera un método no factible para aplicaciones en línea.

#### 2.3.1.2. Método Bayesiano

Tal como se señala en [11], el método Bayesiano trata el modelo de red neuronal como un estimador Bayesiano [15],[16], lo cual significa que los pesos de la red son modelados como una distribución de probabilidad en vez de valores fijos, con lo que la incertidumbre de predicción y, consecuentemente los IP, pueden ser derivados intuitivamente a partir de esta distribución.

Para optimizar los parámetros del modelo, el método Bayesiano intenta estimar una distri-

bución a posteriori para los pesos de la red  $w$ , dado los datos del conjunto de entrenamiento  $z_k$ , aplicando el teorema de Bayes.

Usando esta estrategia, es posible obtener una expresión que es proporcional a la distribución a posteriori deseada, en la cual se puede aplicar el algoritmo de Markov Chain Monte Carlo para estimar la distribución a posterior real, mediante un procedimiento de muestreo [15].

Una vez que la distribución posterior aproximada es determinada, la varianza total del error puede ser calculada como:

$$\sigma_{tot}^{z_k} = \sigma_D^2 + \sigma_{w^{MP}}^2(z_k) = \frac{1}{\beta} + \nabla \hat{y}_{w^{MP}}^T(z_k) (H^{MP})^{-1} \nabla \hat{y}_{w^{MP}} \quad (2.24)$$

en donde  $\sigma_D^2$  es la varianza del ruido de los datos,  $w^{MP}$ , representa el máximo de la distribución a posteriori del peso de la red,  $\nabla \hat{y}_{w^{MP}}$  es el gradiente de la red con respecto a los parámetros  $w^{MP}$  evaluados en el dato de entrada  $z_k$ , y  $H^{MP}$  es la matriz Hessiana del modelo.

Finalmente, especificando la probabilidad de cobertura deseada, se puede construir un IP calculando los cuantiles correspondientes a partir de la distribución a posteriori.

$$\bar{y}(z_k) = \hat{y}(z_k) + q^{1-\frac{\alpha}{2}} \sigma_{tot}(z_k) \quad (2.25)$$

$$\underline{y}(z_k) = \hat{y}(z_k) - q^{1-\frac{\alpha}{2}} \sigma_{tot}(z_k) \quad (2.26)$$

en donde  $q^{1-\frac{\alpha}{2}}$  representa el cuantil  $1 - \frac{\alpha}{2}$  de la distribución a posterior normalizada.

El cálculo de la matriz Hessiana hace de este método un método computacionalmente demandante, lo cual dificulta su aplicación a problemas reales.

### 2.3.1.3. Mean-Variance Estimation Method (MVEM)

El Método de la Estimación de la Media-Varianza (MVEM) [17] propone un IP que se puede construir a partir de dos redes neuronales entrenadas de manera secuencial [18]: una para estimar las predicciones de los datos ( $NN_y$ ) y otra para estimar la varianza de predicción del error ( $NN_\sigma$ ).

La metodología usada para entrenar este modelo es: primero, entrenar la red neuronal  $NN_y$  con una función de pérdida de error cuadrático medio estándar. Luego, la red  $NN_\sigma$  es entrenada para estimar la varianza del error  $\sigma_k^2$  entre cada par de predicción  $\hat{y}(z_k)$  y salida objetivo  $y_k$ .

Esto se consigue considerando la siguiente función de costo:

$$C(y, \hat{y}) = \frac{1}{2} \sum_{k=1}^N \left( \log(\hat{\sigma}^2(z_k)) + \frac{(y_k - \hat{y}(z_k))^2}{\hat{\sigma}^2(z_k)} \right) \quad (2.27)$$

Finalmente, luego de obtener los predictores  $\hat{\sigma}^2(z_k)$  y  $\hat{y}(z_k)$ , un IP puede ser construido de la misma manera que en el método Bayesiano:

$$\bar{y}(z_k) = \hat{y}(z_k) + q^{1-\frac{\alpha}{2}} \hat{\sigma}^2(z_k) \quad (2.28)$$

$$\underline{y}(z_k) = \hat{y}(z_k) - q^{1-\frac{\alpha}{2}} \hat{\sigma}^2(z_k) \quad (2.29)$$

Dado que MVEM estima la incertidumbre de predicción a partir de la salida de una red neuronal, es menos costoso, computacionalmente, que los métodos Delta y Bayesiano, además de ser capaz de adaptarse a ruidos no homogéneos, ya que las redes neuronales son aproximadores universales. La principal desventaja es que este método trata al modelo  $NN_y$  como un estimador insesgado.

#### 2.3.1.4. Método Bootstrap

En consonancia con lo manifestado por [11], se evidencia que el Método Bootstrap fusiona un conjunto de modelos de redes neuronales con procedimientos de muestreo, a fin de establecer estimaciones más precisas de las predicciones del modelo y de la incertidumbre asociada, en comparación con modelos convencionales fundados en la predicción de un solo modelo [19].

De acuerdo con [20], los modelos de enjambre se componen de una serie de metodologías que permiten el uso de múltiples redes neuronales de manera conjunta para resolver un problema. Entre los diversos tipos de modelos de enjambre existentes, se destaca la técnica bootstrap debido a su capacidad para cuantificar la incertidumbre en las predicciones del enjambre.

A pesar de que existen varias versiones existentes de bootstrap para cumplir este objetivo, el algoritmo básico se describe a continuación:

1. Muestrear el conjunto entero de datos en  $S$  subconjuntos de entrenamientos mediante un muestreo uniforme con reemplazo.
2. Entrenar un modelo de red neuronal para cada subconjunto remuestreado (para un total de  $S$  modelos).
3. Una vez que cada modelo ha sido entrenado, la salida total del enjambre se puede calcular de la siguiente manera:

$$\hat{y}(z_k) = \frac{1}{S-1} \sum_{s=1}^S \hat{y}_s(z_k) \quad (2.30)$$

en donde  $\hat{y}(z_k)$  es la predicción total del enjambre para el dato de entrada  $z_k$  y  $\hat{y}_s(z_k)$  es la predicción del  $s^{\text{ésimo}}$  modelo de red neuronal para una muestra de entrada  $z_k$ .

Una vez que el enjambre ha sido entrenado, el método Bootstrap propone que la incertidumbre epistémica  $\sigma_{\hat{y}}^2$  puede ser estimada a partir de la varianza de las salidas de las redes individuales, de acuerdo con la siguiente fórmula:

$$\sigma_{\hat{y}}^2 = \frac{1}{S} \sum_{s=1}^S (\hat{y}_s(z_k) - \hat{y}(z_k))^2 \quad (2.31)$$

Después de esto, la incertidumbre aleatoria es estimada por una nueva red  $\hat{\sigma}_D^2(z_k)$  para ajustar un estimador con los residuales remanentes:

$$\hat{r}^2(z_k) = \max\left([y_k - \hat{y}_{\text{validation}}(z_k)]^2 - \hat{\sigma}_{\hat{y}}^2(z_k), 0\right) \quad (2.32)$$



en donde  $y_k$  es el dato de entrenamiento para una muestra  $k$  y

$$\hat{y}_{\text{validation}}(z_k) = \frac{\sum_{s=1}^S q_s(z_k) \hat{y}_s(z_k)}{\sum_{s=1}^S q_s(z_k)} \quad (2.33)$$

$$q_s(z_k) = \begin{cases} 1, & \text{si } z_k \text{ está en el conjunto de validación del modelos} \\ 0, & \text{si no} \end{cases} \quad (2.34)$$

Este paso final es alcanzado usando una función de costo:

$$C(y, \hat{y}) = \frac{1}{2} \sum_{k=1}^N \left( \log(\hat{\sigma}_D^2(z_k)) + \frac{\hat{r}^2(z_k)}{\hat{\sigma}_D^2(z_k)} \right) \quad (2.35)$$

Finalmente, una vez que ambas incertidumbres han sido estimadas, un IP puede ser construido de una manera similar a los otros métodos secuenciales:

$$\hat{\sigma}_{tot}^2(z_k) = \sigma_D^2 + \sigma_{\hat{y}}^2 \quad (2.36)$$

$$\bar{y}(z_k) = \hat{y}(z_k) + t_S^{1-\frac{\alpha}{2}} \sigma_{tot}(z_k) \quad (2.37)$$

$$\underline{y}(z_k) = \hat{y}(z_k) - t_S^{1-\frac{\alpha}{2}} \sigma_{tot}(z_k) \quad (2.38)$$

en donde  $t_S$  representa el cuantil  $\frac{\alpha}{2}$  de una función acumulativa de la distribución t-Student con  $S$  grados de libertad.

El método Bootstrap, que implica el entrenamiento de múltiples modelos, conlleva un mayor costo computacional en comparación con el MVEM. Sin embargo, el impacto negativo de esta exigencia se minimiza en la fase de evaluación. Por lo tanto, el enfoque Bootstrap sigue siendo un método viable para la resolución de problemas en línea.

Además, este método puede ofrecer resultados más precisos que el MVEM tradicional, ya que tiene la capacidad de estimar con mayor precisión tanto las predicciones como la incertidumbre asociada. Sin embargo, esta hipótesis depende en gran medida de la premisa de que cada uno de los  $S$  modelos de redes neuronales del enjambre ha sido entrenado con suficientes datos para lograr un buen rendimiento y evitar el sobreajuste.

Finalmente, es importante destacar que debido a la forma en que es entrenado, el método Bootstrap es el único método de intervalos neuronales capaz de separar las componentes epistémicas y aleatorias de la incertidumbre, lo que lo convierte en una herramienta valiosa para el modelado de sistemas complejos donde la incertidumbre es una preocupación importante.

### 2.3.1.5. Método de la Covarianza

Tal como se señala en [11], el método de la Covarianza [21] es muy similar al método Delta, pero toma un enfoque más estadístico basado en el trabajo hecho de Škrjanc y Sáez *et al.* [22] en modelos de intervalos difusos.

Usando esta notación y considerando una función de activación tangente hiperbólica, el vector de salida de las neuronas de salida  $Z_k$  y el vector de parámetros del modelo, pueden ser definidos como:

$$Z_k = [Z_{1k}, \dots, Z_{N_h k}, 1] \quad (2.39)$$

$$\theta = [w_1, \dots, w_{N_h}, b] \quad (2.40)$$

en donde:

$$Z_{jk} = \tanh \left( \sum_{i=1}^{n_z} w_{ji} z_k^{(i)} + b_j \right) \quad (2.41)$$

$$\hat{y}(z_k) = Z_k^T \theta \quad (2.42)$$

Luego, asumiendo que el error de predicción es homogéneo y distribuido de manera Gaussiana, se puede calcular la varianza del error de predicción:

$$\sigma_k = \sigma_e (1 + Z_k^T (Z^* Z^{*T})^{-1} Z_k)^{\frac{1}{2}} \quad (2.43)$$

en donde  $\sigma_e$  representa la varianza del ruido de los datos,  $Z_k^T$  es calculado usando los datos de validación y  $Z^*$  es la matriz de todas las salidas  $Z_k$  de las neuronas ocultas para cada muestra del conjunto de entrenamiento.

Finalmente, un IP puede ser construido como:

$$\bar{y}(z_k) = \hat{y}(z_k) + \alpha \sigma_k \quad (2.44)$$

$$\underline{y}(z_k) = \hat{y}(z_k) - \alpha \sigma_k \quad (2.45)$$

En donde  $\alpha$  es un parámetro de sintonización que necesita ser ajustado a través de un proceso iterativo para ajustar la probabilidad de cobertura deseada.

El método de Covarianza se asemeja al método Delta en cuanto a la cuantificación de la incertidumbre, dado que ambos métodos intentan estimar la varianza de la predicción del error. La principal diferencia radica en que el método Delta emplea una metodología basada en parámetros, mientras que el método de Covarianza estima la varianza a partir de los datos.

No obstante, estas similitudes también implican que ambos métodos comparten las mismas limitaciones, asumiendo que el error de predicción es homogéneo y sigue una distribución gaussiana.

Es importante destacar que la inclusión de un proceso iterativo durante el entrenamiento del método de Covarianza puede aumentar significativamente su costo computacional en comparación con el método Delta.

## 2.3.2. Métodos directos

Los métodos directos son métodos de intervalos que no requieren de un modelo base de predicción entrenado previamente, sino que los intervalos se entrenan junto a él.

### 2.3.2.1. Lower Upper Bound Estimation (LUBE)

En [13], se introduce el método de la Estimación de los Límites Superiores e Inferiores (LUBE), que propone una solución para construir un intervalo de predicción mediante el entrenamiento directo de una red neuronal con una arquitectura modificada capaz de predecir los límites superiores e inferiores del intervalo.

La red propuesta en este método consta de una arquitectura similar a la tradicional, pero en este caso se utilizan dos unidades en la capa de salida para generar las dos extremidades del intervalo.

Dado que algunos parámetros del modelo son compartidos entre ambas salidas, la rutina tradicional de entrenamiento de *backpropagation* no es compatible con esta arquitectura.

Para solucionar este problema, las redes de múltiples salidas calculan el gradiente de los pesos compartidos como el promedio entre los gradientes derivados de cada salida del modelo.

Luego, este método propone entrenar la arquitectura con múltiples salidas utilizando como función de pérdida el Criterio de Cobertura Basado en el Ancho (CWC), representado a continuación:

$$\text{CWC} = \text{PINAW} \{1 + \gamma(\text{PICP}) \exp(\nu(\mu - \text{PICP}))\} \quad (2.46)$$

en donde  $\gamma(\text{PICP})$  es representado como:

$$\gamma = \begin{cases} 0, & \text{PICP} \geq \mu \\ 1, & \text{PICP} < \mu \end{cases} \quad (2.47)$$

en donde  $\nu$  y  $\mu$  son dos hiperparámetros que controlan el comportamiento de la métrica CWC. Es importante destacar que el hiperparámetro  $\mu$  representa el nivel de cobertura nominal del intervalo, mientras que el hiperparámetro  $\nu$  establece el peso de penalización para las situaciones en donde la PICP no alcanza los valores deseados.

Además, cabe mencionar que el método LUBE ha demostrado tener un mejor desempeño y mayor eficiencia en comparación con otros métodos de intervalos secuenciales. Debido a esto, en la mayoría de las investigaciones recientes dedicadas a los modelos de IP neuronales, se ha enfocado en mejorar el procedimiento de optimización para el método LUBE.

Dado que la función de costo CWC es altamente no lineal, se han empleado numerosos algoritmos de optimización no convencionales, tales como Simulated Annealing [13], Particle Swarm Optimization [23] y Algoritmos Genéticos [24], entre otros.

### 2.3.2.2. Quality Driven (QD)

El método *Quality Driven (QD)* [25] es una mejora del método LUBE mediante una función de costo modificada:

$$L_{QD} = MPIW_{capt} + \lambda \frac{n}{\alpha(1-\alpha)} \max(0, (1-\alpha) - PICP)^2 \quad (2.48)$$

en donde  $\lambda$  es un hiperparámetro que no depende del conjunto de datos ni de la probabilidad de cobertura deseada,  $n$  es la cantidad total de datos,  $(1-\alpha)$  es la probabilidad de cobertura deseada y  $MPIW_{capt}$  es el ancho del intervalo de predicción capturado promedio:

$$MPIW_{capt} = \frac{\sum_{k=1}^N (\bar{y}(z_k) - \underline{y}(z_k)) i_k}{\sum_{k=1}^N i_k} \quad (2.49)$$

en donde,

$$i_k = \begin{cases} 1, & \text{si } \underline{y}(z_k) \leq y_k \leq \bar{y}(z_k) \\ 0, & \text{si no} \end{cases} \quad (2.50)$$

Con esta función de costo, el método es capaz de utilizar el método de descenso del gradiente durante el entrenamiento, lo que acelera el proceso en comparación con el método LUBE, manteniendo su calidad de intervalo. Además, se incluye el término  $\alpha$  de manera explícita en la función de costo, lo que evita la necesidad de optimizar hiperparámetros, a diferencia del método de *Joint Supervision*.

Una desventaja de este método es que la modificación a la función de costo de LUBE se realiza bajo la hipótesis de que los datos de entrenamiento son independientes e idénticamente distribuidos.

### 2.3.2.3. Joint Supervision

Como se indica en [11], el método de *Joint Supervision* propone la identificación de un único modelo compuesto por tres salidas. La versión de este método presentada en [26] utiliza una red neuronal como modelo base, en el cual dos de sus salidas se utilizan para los límites del intervalo superior e inferior, mientras que la tercera salida se asocia con el valor predicho.

En cuanto al procedimiento de entrenamiento, se utiliza una función de costo diferente para cada salida  $(\bar{y}, \hat{y}, \underline{y})$ . En primer lugar, todas las salidas del modelo comparten una componente de costo de error cuadrático medio para minimizar el error de predicción. Además, las salidas asociadas a los límites del intervalo incluyen un costo de error cuadrático medio que solo se aplica cuando los datos se encuentran fuera del intervalo predicho.

$$\bar{L} = \frac{1}{N} \sum_{k=1}^N (y_k - \bar{y}(z_k))^2 + \lambda \frac{1}{N} \sum_{k=1}^N ReLU^2(y_k - \bar{y}(z_k)) \quad (2.51)$$

$$\hat{L} = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}(z_k))^2 \quad (2.52)$$

$$\underline{L} = \frac{1}{N} \sum_{k=1}^N (y_k - \underline{y}(z_k))^2 + \lambda \frac{1}{N} \sum_{k=1}^N ReLU^2(\underline{y}_k - y(z_k)) \quad (2.53)$$

en donde:

$$ReLU(x) = \begin{cases} 0, & \text{si } x < 0 \\ x, & \text{si } x \geq 0 \end{cases} \quad (2.54)$$

La función de activación utilizada en la red neuronal base es la función ReLU, definida como:

$$ReLU(x) = \begin{cases} 0, & \text{si } x < 0 \\ x, & \text{si } x \geq 0 \end{cases} \quad (2.55)$$

y  $\lambda$  es un hiperparámetro que debe ajustarse para cada experimento individual durante un proceso iterativo. En cada iteración de este proceso de ajuste, se entrena un nuevo modelo de *Joint Supervision* con un valor dado de  $\lambda$ , el cual continúa evolucionando hasta que se alcanzan los valores deseados de cobertura de intervalo.

La principal ventaja del método *Joint Supervision* sobre otros métodos directos es que la simplicidad de su función de costo permite la optimización utilizando algoritmos de descenso del gradiente estocástico altamente optimizados para el entrenamiento de redes neuronales, lo que resulta en una construcción más rápida del IP.

A pesar de esta disminución en el costo computacional, la inclusión del hiperparámetro  $\lambda$ , que debe optimizarse a través de un algoritmo de búsqueda logarítmica, introduce la necesidad de reentrenar el modelo varias veces para obtener una solución óptima. Esto se debe a que en cada iteración del algoritmo de búsqueda logarítmica para el hiperparámetro  $\lambda$ , se entrena un nuevo modelo de *Joint Supervision* con un valor diferente de  $\lambda$  en la función de costo.

Esto tiene un impacto directo en los tiempos de entrenamiento del modelo, ya que a pesar de tener una función de costo simple que permite el entrenamiento eficiente mediante el método del gradiente, se ve afectado por los repetidos entrenamientos.

### 2.3.3. Intervalos de predicción basados en modelos de *Deep Learning*

#### 2.3.3.1. Cuantización de la incertidumbre mediante *Monte Carlo Dropout (MCD)*

En las redes neuronales, la técnica de *dropout* se ha utilizado principalmente como técnica de regularización para evitar el sobreajuste del modelo, como se propuso en [27]. En la versión clásica de la técnica, se realiza un muestreo por cada neurona a partir de la distribución de *Dirichlet* con probabilidad  $p$ , y el valor resultante se multiplica por la salida de la neurona. De esta forma, actúa como si se apagarán algunas neuronas, lo que permite que el aprendizaje sea más homogéneo a lo largo de la red. Se puede ejemplificar la representación de esta técnica en la figura:

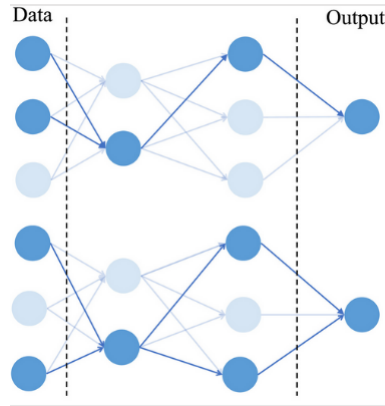


Figura 2.5: Ejemplo de dropout en una estructura de una red neuronal.

La técnica de *dropout* se puede utilizar como un medio para cuantificar la incertidumbre, reinterpretando la técnica como una aproximación bayesiana de un proceso gaussiano, como se describe en [28].

Este proceso implica un número  $K$  de pasos hacia adelante a través de la red, en los cuales se aplica una máscara de *dropout* distinta en cada iteración. Al final de este proceso, en lugar de tener un vector de probabilidades de clase, se obtienen  $K$  salidas del modelo para cada muestra. A partir de estas salidas, se puede calcular el promedio y la varianza como estimaciones de la incertidumbre de predicción del modelo.

## 2.4. Discusión

En este capítulo se discuten distintos métodos relacionados con la predicción mediante redes neuronales, así como intervalos de predicción basados en modelos neuronales. Existen dos tipos de métodos, dependiendo de si se requiere un modelo base ya entrenado previamente o no. Se observa que hay muchos métodos de complejidad variable que asumen diferentes hipótesis relacionadas con el modelo.

En el siguiente capítulo se llevará a cabo un estudio bibliográfico del estado del arte tanto de los métodos presentados en este capítulo como de otros métodos. Se describirán principalmente sus aplicaciones en la literatura, con el fin de aclarar qué métodos pueden ser apropiados para su uso en esta memoria, dadas las condiciones específicas de la misma.

En el siguiente capítulo se desarrolla un estudio bibliográfico del estado del arte tanto de métodos presentados en este capítulo como de otros métodos. Se enuncia principalmente sus aplicaciones que han tenido en la literatura, con el fin de esclarecer cuales métodos pueden ser apropiados para ser utilizados en esta memoria dado las condiciones de esta.

# Capítulo 3

## Estado del Arte

En esta sección se estudian distintas soluciones de la literatura en lo que respecta a la predicción de intervalos, enfocándose en los métodos basados en modelos neuronales.

En [11] se realiza una clasificación dependiendo de la necesidad de un modelo base para el cálculo del IP. Los métodos secuenciales son los que necesitan de un modelo base ya entrenado para establecer el IP, mientras que los métodos directos calculan el IP de manera directa sin la necesidad de este.

Dentro de los métodos secuenciales se encuentra el método Delta, que usa teoría de regresión no lineal clásica para obtener una aproximación de la incertidumbre de predicción. Este método se ha usado en el pronóstico de carga eléctrica [14], monitoreo del proceso de deposición de pasta de soldar [29], predicción de temperatura interior y humedad relativa [30], monitoreo de la calidad de aguas subterráneas [31] y predicción de tiempo de viaje [32]. En [59] Khosravi *et. al* se propone una nueva función de costo para la optimización, llamada función de costo basada en el error de IP (*PICF*). A diferencia de la función de costo original que optimiza mediante el error, esta nueva función contiene como componentes las métricas de *PICP* y *PINAW*. Se define como:

$$PICF = CLC + e^{(PICF_{opt} - WDCF_{trad})}$$

Donde  $WDCF_{trad}$  se obtiene mediante la minimización de la función de costo original. El  $CLC$  es el criterio basado en el largo y la cobertura, definido como:  $CLC = \frac{PINAW}{\sigma(PICP, \eta, \mu)}$ , siendo  $\sigma(\cdot)$  la función sigmoidea definida como:  $\sigma(PICP, \eta, \mu) = \frac{1}{1 + e^{-\eta(PICP - \mu)}}$ , en el cual  $\eta$  y  $\mu$  son dos hiperparámetros que regulan donde y que tan empinado asciende la función sigmoidea.

Este método de optimización consta de dos etapas. En la primera etapa, las redes neuronales son entrenadas utilizando el algoritmo de Levenberg-Marquardt para minimizar la función de costo original. En la segunda etapa, se procede a reentrenar las redes neuronales utilizando un conjunto de datos diferente, minimizando la función *PICF*.

Otro método Secuencial es el Método Bayesiano. Este método se ha usado, por ejemplo, en el pronóstico de la carga eléctrica [33], donde se obtiene un peor resultado de cobertura con respecto al método Delta, pero una mayor robustez ante cambios en la estructura de la red neuronal. También es usado en la para la predicción del tiempo de viaje en buses [32], en donde se emplean diferentes variantes de las métricas de desempeño mencionadas anterior-

mente, llegando a conclusiones similares.

Otro método Secuencial es el de la Estimación de la Media-Varianza (MVEM). Se ha utilizado en: la predicción de la tasa de cambio [34] en donde se concluye que las redes recurrentes tienen un mejor desempeño que las redes feedforward comunes, además de no necesitar un proceso de selección del número de entradas; en la predicción de la vida útil restante de artefactos [35] y predicción del deslizamiento de tierra [36], utilizando como modelo base una red *Echo State Network* (ESN), tipo de red neuronal recurrente, obteniendo mejores resultados que las redes estáticas.

En [37] Khosravi y Nahavandi extienden el *MVEM* para la construcción de intervalos de predicción de la potencia eólica. Para la optimización de la red neuronal  $NN_\sigma$  en vez de usar la estimación de máxima verosimilitud, usan una nueva función de costo que desarrollan en [13]. Esta función incluye tanto *PINAW* como *PICP*, mejorando sustantivamente el desempeño de los intervalos de predicción construidos.

Otro método Secuencial es el método *Bootstrap*. Este método se ha usado en el pronóstico de potencia eólica [38], pronóstico del nivel de un lago [39] en donde usan como modelos además de las redes feedforward, redes neuronales difusas; predicción del tiempo de viaje en bus [40], predicción del caudal de agua en una alimentación nuclear [41], estimación de la degradación de componentes sujetos a una carga [42], diseños de súper aleaciones basadas en níquel [43], obtención del espesor óptico de aerosoles [44] y pronóstico del precio de la electricidad [45] en donde se muestra un mejor desempeño en el ancho del intervalo con respecto al método Delta. En [46] introducen un nuevo método para mejorar la calidad de los IP construidos a partir del método *Bootstrap*. Se propone un nuevo criterio de optimización en el entrenamiento de las redes neuronales, conocido como criterio basado en el ancho y cobertura (CWC), definido ya anteriormente en la ecuación 2.46.

Otro método Secuencial es el método de la Covarianza. Este método se ha usado en aplicaciones como el pronóstico de la carga eléctrica [26].

En [47] L G. Marín *et al.* proponen un sistema *fuzzy* de tipo 2 para un sistema con predicción de generación eólica y carga del sistema.

La función de costo tiene la siguiente expresión:

$$\min_{NN} \quad \beta_1 PINAW + \beta_2 ||e||^2 + \exp[-\eta(PICP - (1 - \alpha))] \quad (3.1)$$

en donde  $\beta_1$  y  $\beta_2$  son hiperparámetros. También se usa la normalización de *PINAW*, y del error definido como  $||e|| = \sqrt{\sum_{j=1}^n \left| t_j - \frac{\bar{y}_j + y_j}{2} \right|^2}$  en el cual  $t_j$  es el valor de la  $j$ -ésima muestra. La función de costo penaliza tanto valores elevados como bajos de *PICP* y la sintonización de  $\beta$  dependen en gran medida de la naturaleza de los datos. Además, como  $||e||$  no está normalizado con el número de muestras, se deben ajustar distintos parámetros  $\beta$  para un distinto número de muestras.

Dentro de los métodos Directos se encuentra el método de la Estimación de los Límites Superiores e Inferiores (*LUBE*). Este método se ha usado en el pronóstico de la potencia eólica [23], [48], pronóstico de la velocidad de viento [49], [50], pronóstico de la potencia solar



[51], [52], pronóstico de la carga eléctrica [23],[53],[54],[55], predicción de temperatura de un reactor [56], [57] y la estimación de la compensación de potencia reactiva de los hornos de arco eléctrico [58].

En [56] Hosen *et al.* proponen un sistema de agregación óptima para mejorar el desempeño de los intervalos de predicción resultantes de *LUBE* mediante la combinación de estos. Para llevarlo a cabo, se utilizan dos técnicas de optimización, Algoritmos Genéticos (AG) y *Simulated annealing* (SA), con el fin de optimizar los pesos de ponderación minimizando la función de optimización *CWC*. En los resultados se obtiene un reducción del *CWC* para los casos de estudio, además de mejorar el desempeño del IP en comparación al caso cuando se usa como técnica de agregación un promedio simple.

En [59] Quan et al. se propone una mejora del método *LUBE* mediante el uso de optimización por enjambre de partículas (*PSO*) y una modificación en la función de costo. En vez de utilizar *PINAW*, utilizan *prediction interval normalized root-mean-square width (PINRW)*, definida como:  $PINRW = \frac{1}{R} \sqrt{\frac{1}{n} \sum_{j=1}^n (\bar{y}_j - \underline{y}_j)^2}$ . modificacndo así su respectiva función de costo  $\min_{NN} CWC = PINRW \{1 + \gamma(PICP)e^{\eta(\mu - PICP)}\}$

Se valida en seis casos de estudio, en donde se alcanza una reducción en el costo de cómputo y una mejora en la calidad del IP.

Otro método directo es el de la Supervisión Conjunta (*Joint Supervision*). Este método ha visto aplicaciones principalmente en problemas de pronóstico de cargas [26], [60].

En [61] usan *Joint Supervision* para predecir la demanda eléctrica de un barrio residencial del Reino Unido y en [60] utilizan el método para construir predicción de intervalos de demanda eléctrica usando redes *Long Short Term Memory (LSTM)*, otro tipo de red recurrente.

En [62] se utiliza *Monte Carlo Dropout* como método de aproximación Bayesiana mediante el uso de. Para esto dejan *dropout* activo en al momento de realizar la inferencia, realizándolo múltiples veces, de modo que cada vez que exista un resultado distinto. A partir de la distribución de estas salidas es posible establecer un intervalo. Un problema de este tipo de métodos es que tienden a subestimar la incertidumbre en muchos casos, en particular de muestras fuera del conjunto de entrenamiento.

También existen métodos, que hace uso del *Ensemble Learning* o aprendizaje de conjunto. El aprendizaje de conjunto consiste en entrenar múltiples modelos de aprendizaje automático y combinar sus resultados. Los diferentes modelos se utilizan como base para crear un modelo predictivo óptimo.

En [63] usan una técnica de *Deep Ensemble* en donde se propone entrenar múltiples redes neuronales con distinta inicialización de parámetros, para generar un conjunto diverso de modelos. Combinándolo con el método de MVEM, se es capaz de aprender la varianza de predicción.

Para evitar los métodos de *Ensemble*, en [64] usan unos métodos de *Split-train* en el cual se usan redes MVEM con un entrenamiento modificado con mini-batches. Se modela la varianza de la predicción usando una distribución *gamma* en donde se usa la distancia euclidiana entre una muestra de prueba y el conjunto de entrenamiento, para determinar la similitud de la red con la entrada. Se produce una varianza más pequeña si la entrada es cercana a los datos de entrenamiento, y una varianza más grande si no. Esta medida de similitud está basado en la distancia euclidiana lo que puede ser difícil de estimar para dimensiones grandes. En [65] se utiliza una arquitectura desacoplada *Decoupled Prediction Interval Network (DPIN)*, se entrena primero la predicción utilizando como función de costo el RMSE. Luego en una fase posterior. Una de las premisas del modelo es que la mayoría de modelos pierden desempeño en

la predicción puntual. En este trabajo comparan los métodos de *MC-Dropout*, *Deep Ensemble*, *Split-train*, *Quality Driven*, *DPIN* en 9 datasets públicos:

viviendas en Boston, fuerza de compresión de concreto, eficiencia energética, cinemática de un brazo robótico, propulsión naval, planta de energía de ciclo combinado, propulsión naval, dataset de vinos, hidrodinámica de un yate y estructura de proteínas.

Se obtiene que el método DPIN tiene el mejor resultado de error de predicción. Lo que demuestra la utilidad de entrenar explícitamente utilizando la función de costo de RMSE, para minimizar el error de predicción. Pero en la métrica de PICP, se obtiene que *Quality Driven* obtiene levemente mejores resultados, pero a costa de un ancho de cobertura mayor. Con respecto a los otros métodos, solo *Split-Train* tuvo resultados comparables, pero con un peor desempeño en general.

Otro tipo de modelos para predicción de series de tiempo es *Temporal Fusion Transformer*. Este está basado en una arquitectura distinta a los modelos anteriores, basado en la arquitectura de *Transformers* [66]. Una novedad de este modelo es la interpretabilidad que tiene debido a la arquitectura que utiliza, además es capaz de operar con distintos horizontes de predicción. Para la predicción de intervalos, existe una función de costos que toma en cuenta el cuantil de los datos que se predicen, logrando obtener intervalos de confianza de las predicciones.

En [67] comparan distintos modelos de predicción de *deep learning* para la construcción de intervalos mediante *MonteCarlo Dropout* para predicción multivariada.

Se usa *Echo State Network (ESN)*, *Long Short Term Memory (LSTM)*, *Cascaded Neural Network (CCN)* y lo comparan con modelos lineales. Se utilizan tres distintos conjunto de datos: datos de calidad de aire, consumo energético de un barrio residencial y el de un programa de bicicletas compartidas. Algunas variables de estos conjuntos tienen una marcada estacionalidad y tendencia, lo cual es perjudicial para el aprendizaje en particular para las redes neuronales. En el preprocesamiento entre otras técnicas, se opta por utilizar un método llamado *Seasonal and Trend decomposition using Loess (STL)*, el cual separa la señal en sus componentes de estacionalidad, de tendencia y el resto.

Una de las conclusiones del trabajo fue que los modelos de *deep learning* tienen un mejor desempeño que los modelos lineales. En particular LSTM y ESN.

Otra conclusión que obtienen, es que se tienen mejores resultados en la predicción de intervalos cuando se tiene mejores resultados de predicción de los datos, por lo que este paso es crucial.

Teniendo estas dos conclusiones en cuenta, cabe pensar en usar modelos de *deep learning* y no solo la estructura clásica de redes neuronales. En particular, debido a que existe un código base con el método de LSTM y de perceptrón multicapa, se decide utilizar LSTM y el modelo de perceptrón multicapa como modelo de predicción base.

En [68] se realiza una comparación entre distintos modelos para estudiar capacidad de aprendizaje en series de tiempo con estacionalidad y tendencias.

Los modelos estudiados son *Feed Forward Neural Network (FFNN)*, *NonPooling Convolutional Neural Network (NPCNN)* y *Convolutional Neural Network (CNN)* y se comparan con modelos lineales. Se concluye que las redes *FFNN* son suficientes para modelar la no linealidad presente en datos con estacionalidad y tendencias. En particular, las redes *NPCNN* tienen un desempeño levemente mejor.

Este resultado reafirma la elección de redes neuronales como método para predecir las variables.

Con respecto a los intervalos de predicción, no se encontraron publicaciones que comparen

el desempeño de todos los modelos aquí mostrados en conjunto.

La mayoría de los modelos aquí planteados hacen algún supuesto en lo que respecta a la distribución de los datos. Mientras que las funciones de costo de los métodos de *LUBE*, *Quality Driven*, *Joint Supervision* optimizan directamente con las métricas de intervalos lo que significa que no se asume distribución alguna. Se descarta el uso de *LUBE* debido a que utiliza una optimización no lineal, además de que *Quality Driven* es una variante de este método que sí puede ser entrenada por el algoritmo de optimización *Gradient Descend*, convirtiéndolo en un modelo más fácil de usar.

Con respecto al *Quality Driven* no existe una variable que defina la salida de predicción propiamente tal, y se toma como valor predicho simplemente el promedio de los intervalos. Esto puede afectar el desempeño en lo que respecta a al error de predicción, pero aún así tiene desempeños comparables al *Joint Supervision*.

En el caso de *Decoupled Prediction Interval Network*, se tiene una arquitectura similar a *Joint Supervision*, en el sentido de tener una salida encargada solamente de la predicción utilizando como función de costo el RMSE. Pero al ser una publicación reciente no se tiene mucha información acerca de comparación además de la que usan en su misma publicación, en comparación al *Joint Supervision* que sí tiene comparación en otras publicaciones.

En el caso de *Temporal Fusion Transformer* si bien es un modelo del estado del arte con buen desempeño y muchas características positivas, como su capacidad de interpretabilidad o poder operar con distintos horizontes de tiempo, se descarta su uso debido al costo computacional por la alta complejidad del modelo.

Un punto a considerar de gran importancia, es que ya hay código programado de los métodos *Joint Supervision* y *Quality Driven*, lo que permite un avance más rápido.

Considerando esta gran ventaja, además de lo dicho anteriormente, se decide utilizar los métodos de intervalos neuronales *Quality Driven* y *Joint Supervision*, usando como modelos base de predicción el modelo perceptrón multicapa y la red LSTM.

# Capítulo 4

## Establecimiento de fuente de datos

A continuación se presenta el desarrollo de la memoria con respecto a la obtención de datos para el posterior entrenamiento del modelo.

### 4.1. Datos de la comunidad

Los datos disponibles de la comunidad proceden de la estación meteorológica localizada en la comunidad José Painecura Hueñalihuen, registrando variables como temperatura del aire, humedad relativa del aire, velocidad del viento y radiación solar con una frecuencia de muestreo de 10 minutos. Sin embargo, se debe tener en cuenta que la comunidad Lafkenche no cuenta con medidores pluviométricos, por lo que no existen registros de precipitación.

En la figura 4.1 se muestra los datos disponibles de cada una de estas variables:

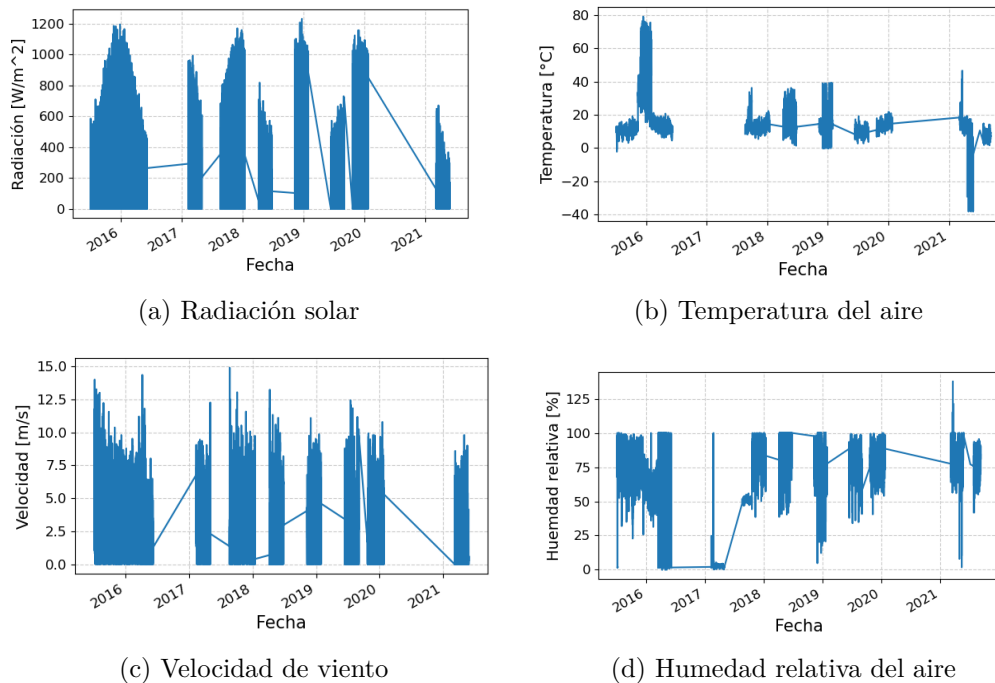


Figura 4.1: Datos meteorológicos registrados en la comunidad José Painecura, a una tasa de muestreo de 10 minutos.

La figura 4.1 muestra la disponibilidad de datos de cada variable mencionada anteriormente, abarcando aproximadamente un periodo de 5 años, desde 2015 hasta 2021. Los datos no presentan una continuidad en el tiempo, y es necesario realizar un análisis detallado para asegurar su calidad. De hecho, en el caso de la temperatura se observan valores anómalos que corresponden a errores instrumentales.

Para entrenar modelos de predicción y intervalos de predicción, es necesario contar con una cantidad de datos que permita cubrir la dinámica que se desea aprender. Debido a la periodicidad de las variables climatológicas estudiadas, tales como la temperatura del aire, la radiación solar y las precipitaciones, los datos utilizados para entrenar el modelo deben extenderse a lo largo de periodos lo suficientemente extensos para abarcar los ciclos estudiados. Por ejemplo, el modelo a corto plazo correspondiente a EMS trabaja con ventanas de dos días, ya que a lo largo de un día se producen ciclos naturales de temperatura más baja durante las noches y las mañanas, y de temperatura más alta durante la tarde. Asimismo, el modelo a largo plazo necesita datos extendidos a lo largo de un año para captar los cambios en estas variables.

Aunque es posible entrenar modelos de corto plazo con una frecuencia de muestreo de 10 minutos, para modelos de mediano plazo que se basan en datos diarios, la falta de datos y los vacíos hacen que sea difícil entrenarlos. Por lo tanto, es necesario buscar nuevas fuentes de datos, tales como estaciones meteorológicas adicionales y datos satélites para poder entrenar los modelos de predicción; además de complementar la información de terreno del área de estudio con datos adicionales que permitan disponer de información más continua en el tiempo.

## 4.2. Búsqueda de fuentes de datos

Para los datos climatológicos se tienen las bases de datos de las siguientes fuentes públicas disponibles: I) *Dirección Meteorológica de Chile* [69], II) *Dirección General de Aguas* [70], y III) los exploradores Solar [71] y Eólico [72]. Además, se consideran datos satelitales provenientes de la fuente ERA5-Land [73, 74].

En la figura 4.2 se muestra la ubicación de las distintas fuentes de datos así como la comunidad José Painecura Hueñalihuen.

En la primera base de datos, las estaciones encontradas más cercanas a la comunidad Lafkenche son las estaciones de Quiripio y Tranapunte, ubicadas a una distancia de la comunidad de 42 km y 26 km respectivamente, en la comuna de Carahue. Ambas pertenecen al *Instituto de Investigación de Agropecuaria (INIA)* (ver figura 4.2)

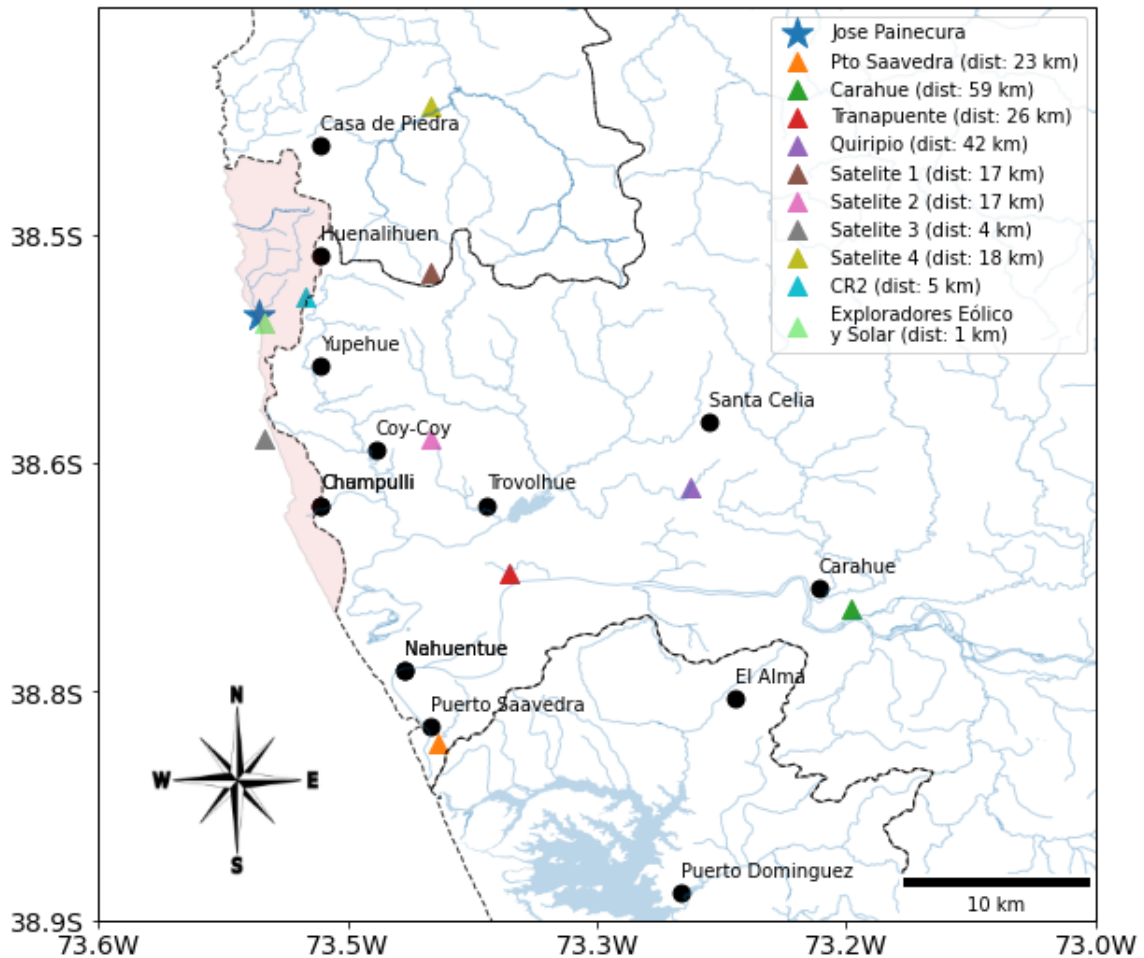


Figura 4.2: Ubicación de las fuentes de datos encontradas. Se limitan las distintas cuencas hidrográficas de la zona, marcando con rojo la cuenca *Costera límite regional y Río Imperial*.

Las variables de interés obtenidas de estas estaciones son:

- Agua caída, acumulada cada 6 horas: Cantidad de Precipitación Líquida Caída o Acumulada durante un Período de 6 Horas;
- Evapotranspiración total diaria: elementos para el cálculo de evapotranspiración total diaria. Calculada con valores registrados durante un día cronológico, es decir, entre las 00:00 y 23:59 hora local;
- Humedad Relativa del Aire: humedad que contiene una masa de aire en relación (cociente) a la máxima humedad absoluta que puede contener;
- Radiación solar global instantánea;
- Temperatura del Aire Seco;
- Viento a 2 metros de altura: dirección e intensidad del viento, medido a 2 metros de altura.

Estas variables son importantes para la predicción de las variables climatológicas de interés en la comunidad lafkenche y se utilizarán para entrenar y validar los modelos de predicción e intervalos de predicción.

La resolución temporal de los datos es horaria excepto para la evapotranspiración total y precipitación, que son diarias y cada 6 horas, respectivamente. El inicio de la medición fue en el año 2019 y continua hasta la actualidad.

La segunda base de datos que se usa es la perteneciente al DGA, en donde se hace uso del explorador climático CR2 [75] y de los reportes climatológicos del Ministerio de Obras Públicas [70]. En estas hay datos de estaciones pluviométricas con datos de precipitaciones además de datos mesoescales.

Las estaciones pluviométricas más cercanas a la comunidad Lafkenche son las de Puerto Saavedra y Carahue, ubicadas a 23 km y 59 km de la comunidad respectivamente (ver figura 4.2). En estas se encuentran datos de precipitaciones diarias. En el caso de la estación de Puerto Saavedra, se registran datos que remontan a 1950 y llegan hasta fines del año 2020. Mientras que en el caso de la estación de Carahue, los datos abarcan el período 2018-2020 [70].

También se hace uso de la base de datos de CR2 MET, el cual fue creado a través del método de regionalización estadística. De este repositorio se obtienen datos diarios de temperatura del aire y precipitaciones [75]. Estos valores se encuentran disponibles en un grillado cuadrulado con una resolución de un km cuadrado. Los valores de la grilla se aproximan en base a los datos combinados de estaciones meteorológicas, datos satelitales y topografía del lugar.

El Explorador Solar [71] es una plataforma del Ministerio de Energía que entrega información acerca de la radiación y temperatura. Se basa en modelos que combinan datos satelitales, modelos físicos y datos topográficos. Los datos tienen una resolución horaria, en donde se ocupa la radiación global horizontal y temperatura del aire.

El Explorador Eólico [72], es el análogo del explorador solar pero para información de viento. Entrega velocidad de viento para distintas alturas. En particular se selecciona la velocidad de viento en las cercanías de la comunidad, a una altura de 5 metros sobre el nivel de tierra.

Para los datos satelitales, se utilizan datos del modelo *ERA5-Land* del servicio europeo satelital *Copernicus* [73, 74]. Este modelo contiene distintas variables de superficies continentales en donde se combinan observaciones satelitales con un modelo geofísico para obtener mediciones más precisas. El modelo es global, y su resolución espacial es de  $0.1^\circ \times 0.1^\circ$ . Los datos están disponibles a partir del año 1950 con una resolución temporal horaria, con aproximadamente 3 meses de desfase a partir del tiempo actual debido a los ajustes de calibración del modelo. Algunas de las variables meteorológicas de interés que incluye son:

- Componente  $u$  de velocidad y dirección de viento a 10 metros de altura.
- Componente  $v$  de velocidad y dirección de viento a 10 metros de altura.
- Temperatura del aire a 2 metros de altura.

- Radiación superficial solar neta.
- Humedad relativa del aire.

De toda la grilla disponible se seleccionan los 4 puntos más cercanos a la comunidad, denominándolos *Sateiltal 1*, *Sateiltal 2*, *Sateiltal 3* y *Sateiltal 4*, los que se pueden observar en la figura 4.5. Finalmente se elige como dato satelital el punto más cercano a la comunidad, siendo este *Satelital 3*.

Luego, se obtiene el equivalente diario mediante una agregación de los datos, tomando el promedio de los valores diarios. A modo de resumen, se muestran en las tablas 4.1, 4.2, 4.3 y 4.4, la cantidad de datos horaria disponibles para las diversas fuentes junto con el rango de tiempo y el tiempo de muestreo original de las diversas fuentes.

Tabla 4.1: Cantidad de datos disponibles de las distintas fuentes de datos para la radiación solar agregadas de manera diaria.

Fuente de datos	Rad. Solar Diaria # Datos	Rango temporal	Tiempo de muestreo
<i>Datos Comunidad</i>	999	04/07/2015 - 27/05/2021	10 minutos
<i>Explorador solar [71]</i>	4744	01/01/2004 - 26/12/2016	1 hora
<i>Tranapuente [76]</i>	1203	01/01/2019 - 19/04/2022	1 hora
<i>Quiripio [77]</i>	1204	01/01/2019 - 19/04/2022	1 hora
<i>Satelital 3 [74]</i>	4139	01/01/2011 - 01/05/2022	1 hora

De la tabla 4.1 se observa que el Explorador solar es el que tiene la mayor cantidad de datos, junto a los datos satelitales.

Tabla 4.2: Cantidad de datos disponibles de las distintas fuentes de datos para la temperatura del aire agregadas de manera diaria.

Fuente de datos	T° Aire Diaria # Datos Totales	Rango temporal	Tiempo de muestreo
<i>Datos Comunidad</i>	960	04/07/2015 - 09/09/2021	10 minutos
<i>Datos Comunidad Filtr</i>	436	04/07/2015 - 24/01/2020	10 minutos
<i>Explorador solar [71]</i>	4748	01/01/2004 - 30/12/2016	1 hora
<i>Tranapuente [76]</i>	1199	01/01/2019 - 19/04/2022	1 hora
<i>Quiripio [77]</i>	1204	01/01/2019 - 19/04/2022	1 hora
<i>Satelital 3 [74]</i>	4139	01/01/2011 - 01/05/2022	1 hora
<i>CR2 Met [75]</i>	15096	01/01/1979 - 30/04/2020	1 día

De la tabla 4.2 se observa que los datos de *CR2 Met* parte del año 1979, lo que se traduce en muchos datos un orden de magnitud más grande que los demás. Luego está el Explorador solar que tiene aproximadamente 12 años de datos, abarcando los años 2004-2016.



Tabla 4.3: Cantidad de datos disponibles de las distintas fuentes de datos para la velocidad de viento agregadas de manera diaria.

Fuente de datos	Vel. Viento Diaria	Rango temporal	Tiempo de muestreo
	# Datos		
<i>Datos Comunidad</i>	999	04/07/2015 - 27/05/2021	10 minutos
<i>Explorador eólico [72]</i>	13880	01/01/1980 - 31/12/2017	1 hora
<i>Tranapunte [76]</i>	1203	01/01/2019 - 19/04/2022	1 hora
<i>Quiripio [77]</i>	1204	01/01/2019 - 19/04/2022	1 hora
<i>Satelital 3 [74]</i>	4139	01/01/2011 - 01/05/2022	1 hora

De la tabla 4.3 se observa que los datos del Explorador eólico se registran desde 1980, lo que conlleva a tener muchos datos en comparación con el resto.

Tabla 4.4: Cantidad de datos disponibles de las distintas fuentes de datos para la precipitación diaria.

Fuente de datos	Precipitación Diaria	Rango temporal	Tiempo de muestreo
	# Datos Totales		
<i>Pto Saavedra [70]</i>	14808	01/01/1979 - 31/10/2020	1 día
<i>Carahue [70]</i>	639	01/01/2019 - 31/10/2020	1 día
<i>Tranapunte [76]</i>	1205	01/01/2019 - 19/04/2022	6 horas
<i>Quiripio</i>	1205	01/01/2019 - 19/04/2022	6 horas
<i>Satelital 3 [74]</i>	12174	01/01/1989 - 01/05/2022	1 día
<i>CR2 Met [75]</i>	15096	01/01/1979 - 30/04/2020	1 día

De la tabla 4.4 se observa que los datos de *Puerto Saavedra*, *Satelital 3* y *CR2 Met* tienen un largo registro de datos, aunque *Puerto Saavedra* es el único con mediciones físicas de datos.

## 4.3. Selección de datos

### 4.3.1. Precipitaciones

Dado que no hay datos de precipitación disponibles para la comunidad, se debe realizar un análisis para determinar qué fuente de datos utilizar. Para ello, se consulta con expertos en hidrología con el objetivo de seleccionar la fuente de datos más adecuada para las necesidades del estudio.

En primera instancia, debido a la cantidad más alta que se dispone sobre las demás, se evalúa utilizar los datos de *Satelital 3 (ERA5)*, *CR2 Met* y el de *Puerto Saavedra*.

En la figura 4.3 se muestra la precipitación anual acumulada entre los años 1989 y 2020 (2020 incompleto).

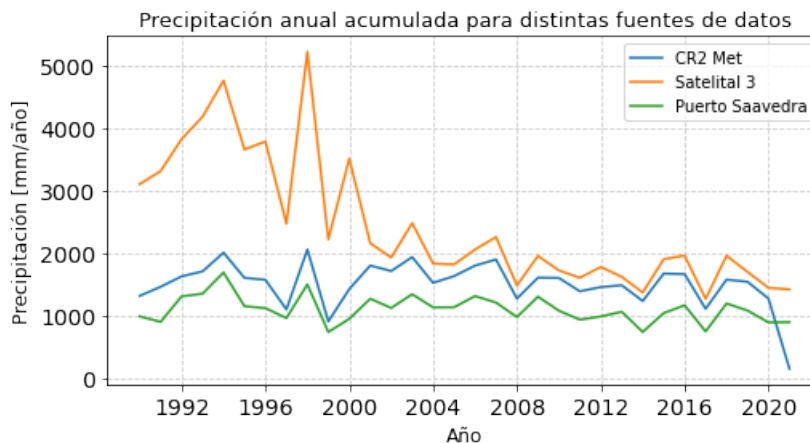


Figura 4.3: Precipitación anual acumulada para las fuentes de *CR2 Met*, *Satelital 3 (ERA5)* y *Puerto Saavedra*

De la figura 4.3 se observa que los datos de *ERA5* y *CR2Met* son considerablemente más altos que los de Puerto Saavedra, sobre todo los de *ERA5*. Cabe decir, que en los datos de *Puerto Saavedra* existe una cantidad importante de datos faltantes en los años 1991 y 2017. Pero aún así se mantiene la tendencia que la magnitud de precipitación acumulada es menor que las demás.

Asimismo, se observa que los datos de *CR2 Met* son muy similares a los datos de Puerto Saavedra en su forma. Esto se debe a que el cálculo de los datos de *CR2 Met* utiliza datos medidos de estaciones pluviométricas, en particular de *Puerto Saavedra*, así como datos satelitales de *ERA5*.

Dado que existen diferencias en los datos disponibles y su distribución espacial, los datos medidos in situ son considerados los más adecuados para la escala de trabajo de esta memoria. Esto se debe a que estos datos se posicionan en un escenario menos favorable (menos lluvias, menos agua disponible) que es más parecido a la realidad. Además, los datos medidos siempre son más confiables que los datos satelitales o los datos modelados a partir de modelos estadísticos. Por lo tanto, se descartan los datos de *ERA 5* para la precipitación.

Por otra parte, se realizan curvas dobles mágicas entre los datos de Puerto Saavedra y *CR2 Met*, tomando como estación patrón la estación de Puerto Saavedra. En este tipo de curva, se grafica la precipitación acumulada de dos fuentes distintas de datos para poder comparar su comportamiento. En la elección de la estación patrón se asume que sus datos son correctos y se compara con una estación de prueba. Se dibuja una línea 1:1 que representa una correlación con un coeficiente de determinación, indicando qué tan bien se ajusta la estación de prueba a los datos de la estación patrón. Si los datos se encuentran en la diagonal, significa que ambas estaciones han medido la precipitación de manera correcta. A mayor distancia de la diagonal con respecto a los datos, existe más error en las mediciones en la estación de prueba.

En la figura 4.4 se muestra la curva con zoom. Teniendo en cuenta que ha ocurrido una sequía desde aproximadamente el año 2009, se hace la distinción entre los datos previos y posteriores a este año

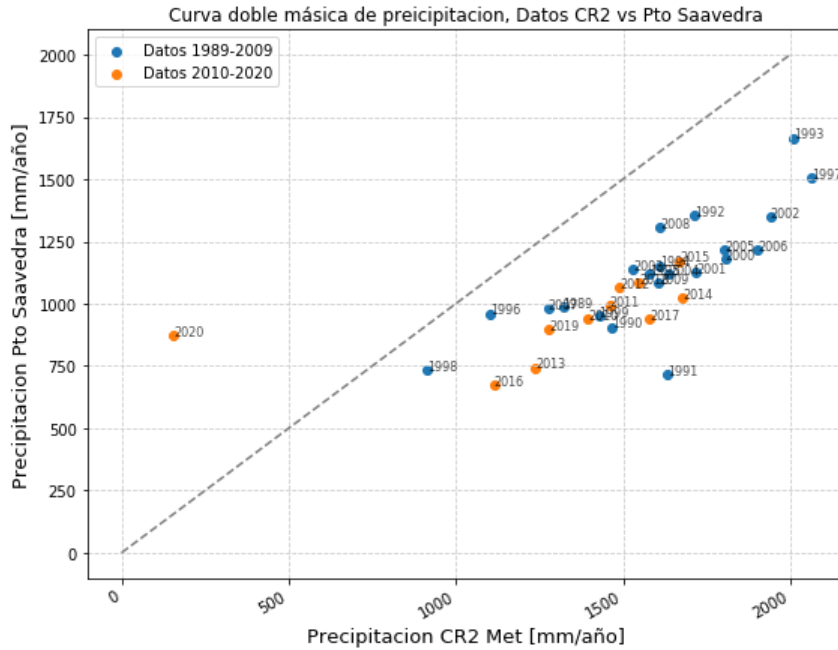


Figura 4.4: Curva doble másica entre Cr2 Met y la estación de Puerto Saavedra.

De la figura 4.4, se puede observar que los puntos se ubican a la derecha de la diagonal. Esto indica que los datos de precipitación anual acumulada de *CR2* son mayores en magnitud que los de *Puerto Saavedra* y por tanto que las mediciones de precipitaciones de ambas no fueron medidas de la misma manera. Debido a que representan puntos geográficos distintos, asumir que los datos de la estación Puerto Saavedra son correctos no asegura que los de *CR2* son incorrectos.

La fuente *CR2 Met* contiene datos climatológicos más cercanos a la comunidad que la estación Puerto Saavedra, sin embargo, se debe considerar que esta información es el resultado de una interpolación en la grilla geográfica. Por otra parte, esta fuente también presenta su repositorio de datos de forma continua entre 1979 hasta el año 2020 sin vacíos.

Finalmente para lo que respecta a esta memoria se decide trabajar con los datos medidos en la estación Puerto Saavedra considerando que reflejan mejor la realidad climática de la comunidad.

Con este fin, se obtiene información acerca de las cuencas hidrológicas de la zona. Estas constituyen zonas de aguas subterráneas compartidas. Puerto Saavedra y la comunidad se encuentran en cuencas distintas, pero ambas son costeras y vecinas mutuamente, por lo que sus niveles de precipitación deberían ser similares.

Las cuencas de las zonas se pueden observar en la figura 4.2. Se verifica que ambas se encuentran en cuencas separadas, y más aún, la comunidad se encuentra en una cuenca pequeña y alejada (cuenca *Costera límite regional y Río Imperial*), existiendo casi nula información al respecto, imposibilitando un análisis más a fondo con respecto a las cuencas.

Se designó a Puerto Saavedra como la fuente más confiable para aproximar las precipitaciones de la cuenca en la que se localiza la comunidad por las siguientes razones: 1) Distancia, entre esta estación y el lugar de estudio, hay una distancia de 23 kilómetros, (ver figura 4.2); 2) Zona costera: ambas localizaciones se encuentran en zonas costeras y 3) se encuentran en cotas similares, por lo que estas zonas deberían presentar variables climatológicas similares.

En la figura 4.5 se muestra el mapa de elevaciones de la zona.

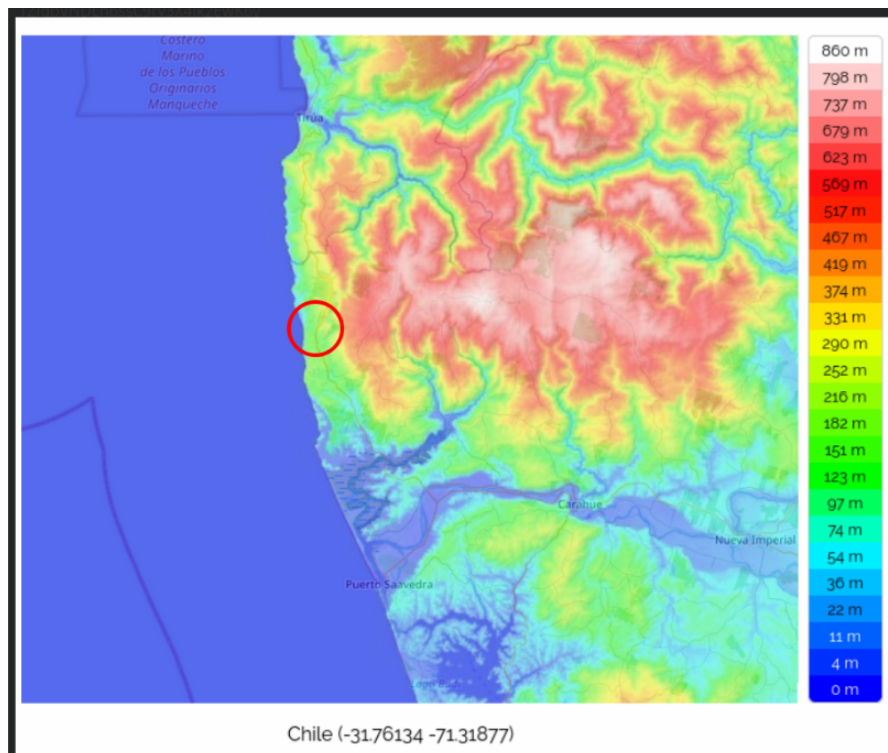


Figura 4.5: Mapa de elevaciones de la novena región. En un círculo rojo se demarca la ubicación de la comunidad mapuche (Fuente: <https://es-cl.topographic-map.com/maps/6iyt/Chile/>).

Se observa en la figura 4.5 que la cota de la comunidad está entre el nivel del mar y unos 100-150m. Este es el mismo rango de la estación de Puerto Saavedra, por lo que finalmente se valida su uso.

### 4.3.2. Temperatura, radiación solar y velocidad de viento

Para seleccionar los datos de las variables de temperatura, velocidad del viento y radiación solar, se comparan los datos registrados en la comunidad con las fuentes externas de datos, considerando únicamente los tramos de datos que se encuentren en el mismo rango temporal. Para ello, se utilizan datos con diferentes resoluciones temporales, los cuales se homogeneizaron para luego calcular el promedio diario de los casos en que los datos tengan un tiempo de muestreo menor a un día, para garantizar una comparación consistente y diaria. Como métricas de comparación se emplean el error cuadrático medio (RMSE) y el error absoluto medio (MAE) (ecuaciones 2.14 y 2.15, respectivamente).

En la tabla 4.5 se presenta el error cuadrático medio y el error absoluto medio para los datos de radiación solar agregados diariamente mediante el promedio, calculados entre los datos de la comunidad y las demás fuentes externas. En la columna *# Datos Comparación* se indica el número de datos que comparten un intervalo temporal entre la comunidad y la fuente de datos externa, y se realiza la comparación.

Tabla 4.5: Error entre datos de radiación solar agregadas de manera diaria de la comunidad y las demás fuentes de datos (en color azul los errores de menor magnitud).

Fuente de datos	Radiación Diaria		
	# Datos Comparación	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]
<i>Explorador solar</i>	340	24.7	21.6
<i>Tranapunte</i>	291	77.3	58.3
<i>Quiripio</i>	291	211.1	175.9
<i>Satelital 3</i>	999	64.4	52.1

En la tabla 4.5 se puede observar que para el caso de la radiación solar, los datos obtenidos a través del *Explorador Solar* presentan un error menor para ambas métricas, en comparación con los datos obtenidos por *Satelital 3*. Es importante destacar que, aunque *Satelital 3* proporciona tres veces más datos, lo que permite una estimación más precisa del error, la diferencia en el error obtenido es significativa, casi tres veces mayor que el obtenido por el *Explorador Solar*.

Es preciso destacar que el *Explorador Solar* cuenta con una gran cantidad de datos disponibles, lo cual le confiere una mayor robustez a la hora de estimar la radiación solar diaria. Es por esta razón que se elige al *Explorador Solar* como la fuente de datos para la estimación de la radiación solar diaria en este estudio.

En la tabla 4.6 se presentan los valores del error cuadrático medio y del error absoluto medio para los datos de radiación solar agregados diariamente mediante el promedio, calculados entre los datos de la comunidad y el explorador solar.

Tabla 4.6: Error entre datos de temperatura agregadas de manera diaria de la comunidad y las demás fuentes de datos (en color azul los errores de menor magnitud).

Fuente de datos	Temperatura Diaria		
	# Datos Comparación	RMSE [°C]	MAE [°C]
<i>Explorador solar</i>	258	1.34	0.86
<i>Tranapunte</i>	178	1.44	1.11
<i>Quiripio</i>	178	2.77	2.55
<i>Satelital 3</i>	436	1.29	0.96
<i>CR2 Met</i>	436	1.53	1.23

En la tabla 4.6, se observa que la fuente *Satelital 3* presenta un menor RMSE para los datos diarios de radiación solar, mientras que la fuente del Explorador solar arroja un menor MAE. Es importante destacar que la diferencia en RMSE entre ambas fuentes es muy pequeña, menor a 0.05 grados. Dado que se decidió utilizar los datos del Explorador solar para la radiación solar diaria, se considera conveniente emplear esta fuente también para los datos diarios de temperatura.

En la tabla 4.7 se presentan los resultados del cálculo del error cuadrático medio y el error absoluto medio para los datos de velocidad de viento agregados diariamente mediante el promedio, y comparados entre los datos de la comunidad y las demás fuentes de datos.

Tabla 4.7: Error entre datos de velocidad de viento agregadas de manera diaria de la comunidad y las demás fuentes de datos (en color azul los errores de menor magnitud).

Fuente de datos	Velocidad Viento Diaria		
	# Datos Comparación	RMSE [m/s]	MAE [m/s]
<i>Explorador Eólico</i>	557	1.65	1.34
<i>Tranapunte</i>	291	1.42	1.15
<i>Quiripio</i>	291	2.75	2.13
<i>Satelital 3</i>	999	1.86	1.52

De la tabla 4.7 se observa que el error calculado es menor para la estación Tranapunte. Pero considerando que el *Explorador Eólico* tiene una cantidad de datos diez veces más cuantiosa que Tranapunte, como se indicó en la tabla 4.3, y además fue el que obtuvo el segundo menor error, se decide utilizar estos datos.

### 4.3.3. Discusión

En este capítulo se describe el proceso de búsqueda y selección de nuevas fuentes de datos para utilizarlos en la implementación de los modelos. Para las variables de radiación solar, temperatura y velocidad del viento se realizó una comparación del error entre los datos registrados en la comunidad y los datos de las nuevas fuentes disponibles. Se decidió utilizar los datos del Explorador Solar para la radiación solar diaria y la temperatura diaria debido a que presentaron un menor error comparado con los datos de otras fuentes, mientras que para la velocidad del viento se optó por utilizar los datos del Explorador Eólico.

En cuanto a las precipitaciones, no se encontraron datos registrados en la comunidad, por lo que después de un análisis pertinente se decidió utilizar los datos de la estación física de Puerto Saavedra para las estimaciones deseadas. Estos datos estarán disponibles para la comunidad de investigadores.

# Capítulo 5

## Diseño e implementación de modelos de predicción

En este capítulo se describe el diseño e implementación de modelos de predicción, basado en el estado del arte se consideran las redes perceptrón multicapa y las redes LSTM.

### 5.1. Modelo de predicción de perceptrón multicapa

#### 5.1.1. Selección de datos

En cuanto a la temperatura registrada cada 10 minutos en la comunidad, se observa una cantidad considerable de datos faltantes, así como la presencia de algunos datos corruptos con magnitudes extremas, como se observa en la figura 4.1. Por lo tanto, se procede a seleccionar el rango con la mayor continuidad, es decir, el tramo de octubre del año 2019 a enero del año 2020, abarcando 3 meses. Para interpolar los datos faltantes dentro de este rango, se utiliza el promedio de los datos vecinos

En el caso de la radiación registrada cada 10 minutos en la comunidad, se elige el rango de datos entre 2015 y 2016 debido a que es el mayor período continuo de datos, abarcando aproximadamente un año. En este caso, no hay datos faltantes en este rango.

Para los datos de precipitación, se opta por rellenar los datos faltantes con el promedio histórico diario calculado en el rango de interés. Sin embargo, debido a que desde aproximadamente el año 2009 se ha vivido un periodo de sequía, se decide tomar los datos solo a partir de ese año, ya que la dinámica de las precipitaciones ha cambiado en comparación con el pasado.

En cuanto a la velocidad del viento, se dispone de datos del Explorador Eólico desde 1980 hasta 2016. No obstante, después de analizar la alta variabilidad de los datos y la dificultad de obtener resultados aceptables, se decide no implementar modelos para estos datos.

Tanto para los datos de temperatura diaria como para los de radiación solar, se utilizan los datos del Explorador Solar, que contienen información desde el año 2004 al 2016. No hay datos faltantes en este rango.

En la tabla 5.1 se muestra el número de datos disponibles y estadísticas de cada uno de

los conjuntos que se utilizarán en los modelos.

Tabla 5.1: Estadísticas de las fuentes de datos seleccionadas para usarlas en la implementación de los modelos de predicción.

Estadísticas	Fuente de datos				
	Radiación solar a 10m	Temperatura a 10m	Radiación solar diaria	Temperatura diaria	Precipitación diaria
Nro de datos	48836	13485	4744	4748	4349
Promedio	193.8 [W/m <sup>2</sup> ]	14.10 [°C]	186.4 [W/m <sup>2</sup> ]	12.2 [°C]	2.7 [mm/día]
Máximo	1196.3 [W/m <sup>2</sup> ]	21.85 [°C]	415.0 [W/m <sup>2</sup> ]	23.9 [°C]	96.8 [mm/día]
Mínimo	0.0025 [W/m <sup>2</sup> ]	6.78 [°C]	11.3 [W/m <sup>2</sup> ]	2.3 [°C]	0 [mm/día]

En la tabla 5.1 se observa que la cantidad disponible para los datos a 10 minutos es mayor a la de los datos diarios. Esto es esperable, ya que un solo día de datos contiene 144 datos de 10 minutos y solo 1 dato diario. También se aprecia que los datos diarios tienen una cantidad de datos muy similar entre ellos.

### 5.1.2. Selección de modelos

A partir de lo analizado en el estado del arte, se decide utilizar el modelo de perceptrón multicapa (PM) y el modelo de *deep learning Long Short Term Memory* (LSTM) para los casos de radiación solar y temperatura, tanto diaria como a 10 minutos. El modelo LSTM ha demostrado tener buenos resultados para series de tiempo, pero requiere de más datos que el modelo de perceptrón multicapa. Debido a la menor cantidad de datos de precipitación, solo se utiliza el modelo de perceptrón multicapa por las razones descritas anteriormente. Para implementar el modelo de perceptrón multicapa, se hace uso del lenguaje *Python* junto con la librería *Pytorch* de *Deep Learning*. Para el caso del modelo LSTM, usaremos *Matlab* junto con sus respectivas librerías de *Deep Learning*.

### 5.1.3. Metodología utilizada para el entrenamiento del perceptrón multicapa

La metodología utilizada para el entrenamiento en el caso de los modelos de perceptrón multicapa se resume a continuación:

1. Separación de conjuntos de datos;
2. Preprocesamiento de datos;
3. Entrenar con un número de regresores fijo, iterando para distintos valores de neuronas en la capa oculta para obtener el número óptimo;
4. Análisis de sensibilidad de los regresores;
5. Entrenar con los regresores óptimos, iterando para distintos neuronas en la capa oculta;
6. Entrenar con un número de regresores óptimas, iterando para distintos valores de neuronas en la capa oculta para obtener el nuevo óptimo.



Este tipo de entrenamiento no es tan exhaustivo, pero se decide utilizarlo debido a limitaciones de tiempo. La búsqueda exhaustiva de hiperparámetros puede llevar mucho tiempo y no es factible en este caso.

A continuación se explica cada uno de los pasos:

#### 5.1.3.1. Separación de conjuntos

Los conjuntos de datos se separan en 80 % datos de entrenamiento y 20 % conjunto de test. Para el entrenamiento, se utiliza la validación cruzada con 5 subconjuntos.

#### 5.1.3.2. Preprocesamiento

En el caso de los conjuntos de datos de temperatura, se normalizan aplicando la fórmula:

$$y_{std} = \frac{y - \bar{y}}{\sigma^2} \quad (5.1)$$

en donde  $\bar{y}$  es el promedio, y  $\sigma^2$  es la desviación estándar.

Para los demás casos, radiación solar y precipitación, debido a la naturaleza física de las variables no pueden tener valores negativos. Por lo que, se procede a normalizar dividiendo por su máximo, es decir:

$$y_{std} = \frac{y}{y_{max}} \quad (5.2)$$

#### 5.1.3.3. Búsqueda del número de neuronas óptimo

Se entrena la red con un número fijo de regresores y se varía el número de neuronas en la capa oculta. Para los datos registrados cada 10 minutos, el número de regresores es de 200, lo que equivale a aproximadamente a 1 día y medio hacia el pasado. Mientras que para los datos diarios, se fija el número de regresores en 60, lo que equivale a 2 meses hacia el pasado. Se prueban múltiplos de 5 en el rango de 10 a 300 para el número de neuronas en la capa oculta.

Se utiliza el método de optimización Adam, se usan lotes de 10 datos y se entrena durante 100 épocas. La función de pérdida utilizada es el error cuadrático medio. Los valores de error de entrenamiento y validación obtenidos para cada subconjunto de la validación cruzada se promedian para obtener el error de entrenamiento y validación promedio para cada época en cada iteración de  $n_h$ .

#### 5.1.3.4. Optimización del número de regresores

La optimización del número de regresores se efectúa en dos etapas. En la primera etapa, se realiza un análisis de sensibilidad para calcular la importancia de cada regresor. En la segunda etapa, se entrenan los modelos utilizando únicamente los  $k$  regresores más importantes determinados en la etapa anterior.

En el análisis de sensibilidad se calcula el gradiente de la salida de la red con respecto a

cada una de las entradas (regresores), esto es:

$$\frac{\partial y}{\partial x_i} i \in \{0, \dots, L\}$$

En donde  $L$  es el número de regresores. Esto con el fin de establecer cuánta información aporta en promedio a la salida cada una de las entradas. Para llevar a cabo este análisis, se utiliza el conjunto de datos de entrenamiento y se obtienen vectores por cada regresor. Posteriormente, se aplica valor absoluto y se calcula tanto el promedio como la desviación estándar.

En la siguiente etapa, se procede a reentrenar los modelos, modificando los regresores mientras se mantiene constante el número de neuronas optimizadas obtenido en la fase anterior. Los regresores se modifican seleccionando los  $k$  mejores regresores. En el caso de los datos con intervalo de 10 minutos, se varía  $k$  en múltiplos de 5 en un rango de 5 a 200. Por otro lado, en el caso de los datos diarios como la radiación solar, temperatura y precipitación, el valor de  $k$  varía entre 1 y 60, sin saltos.

El modelo que obtenga el menor error indica cuál es el número de regresores óptimos.

Una vez que se obtienen los regresores óptimos, se procede a optimizar nuevamente el número de neuronas manteniendo fijo estos nuevos regresores.

#### 5.1.4. Resultados búsqueda inicial de número de neuronas en la red perceptrón multicapa

En esta sección se muestra los resultados de la primera optimización de hiperparámetros del perceptrón multicapa.

En la tabla 5.2 se muestra el número de neuronas seleccionadas en la optimización inicial de hiperparámetros. Dejando fijo el número de regresores: 200 para los datos a 10 minutos y 60 regresores para el caso de los datos diarios.

Tabla 5.2: Número de neuronas en la capa oculta seleccionadas en la optimización inicial de hiperparámetros, con los regresores originales.

Conjunto de datos	Número de neuronas con regresores originales	Número de regresores
Radiación solar a 10m	100	200
Temperatura a 10m	70	200
Radiación solar diaria	70	60
Temperatura diaria	25	60
Precipitación diaria	50	60

De acuerdo con la tabla 5.2, se puede observar que, a pesar de haber probado con valores tan grandes como 300 neuronas, existe una tendencia a que el óptimo se encuentre en valores más pequeños. Esto podría deberse a que, para números más grandes, existe una tendencia al sobreajuste de los datos, lo cual afecta el desempeño general del modelo.

### 5.1.5. Optimización del número de regresores

En esta sección se presentan los resultados de la optimización del número de regresores de la red perceptrón multicapa. Para ello, se llevó a cabo un análisis de sensibilidad a fin de determinar el orden de importancia de los regresores. Las figuras B.2, B.1, B.4 y B.3 del anexo muestran la magnitud de los gradientes de cada regresor para los diferentes conjuntos de datos.

Luego, se procede a entrenar la red utilizando los  $k$  mejores regresores, variando el valor de  $k$  y manteniendo fijo el número óptimo de neuronas obtenido en la fase anterior. Finalmente, se compara el desempeño de los modelos en función del error obtenido.

En la tabla 5.3 se muestran los regresores seleccionados para cada una de las fuentes.

De aquí se observa que en todos los casos se registran como regresores importantes los de menos pasos y más cercanos al tiempo actual. Esto tiene sentido debido a la correlación que existe entre datos cercanos.

Tabla 5.3: Regresores seleccionados en el análisis de sensibilidad para cada uno de los conjuntos (m: denota la cantidad de minutos, h: la cantidad de horas, d: la cantidad de días).

Conjunto de datos	Nro regresores	Regresores
Rad 10m	35	10m, 20m, 30m, 50m, 1:00h, 1:10h, 2:10h, 2:20h, 4:30h, 5:00h, 6:00h, 6:20h, 6:50h, 18:10h, 20:50h, 21:10h, 21:20h, 21:50h, 22:00h, 22:40h, 22:50h, 23:40h, 24:00h, 24:20h, 24:40h, 24:50h, 25:20h, 25:40h, 26:00h, 26:10h, 27:50h, 28:00h, 28:10h, 29:20h, 30:00h
Temp 10m	10	10m, 20m, 30m, 40m, 50m, 7:40h, 9:00h, 15:20h, 20:50h, 24:20h
Rad 1d	25	1d, 3d, 4d, 5d, 6d, 7d, 8d, 10d, 12d, 14d, 15d, 18d, 19d, 22d, 24d, 25d, 27d, 34d, 38d, 39d, 40d, 45d, 55d, 57d, 59d
Temp 1d	29	1d, 3d, 5d, 6d, 7d, 10d, 11d, 14d, 15d, 18d, 25d, 27d, 29d, 31d, 33d, 34d, 36d, 38d, 40d, 43d, 48d, 50d, 51d, 52d, 53d, 56d, 58d, 59d, 60d
Precip 1d	10	1d, 3d, 5d, 11d, 13d, 19d, 27d, 36d, 49d, 60d

Además, en el caso de los datos de corto plazo se constata que los regresores cercanos a las 24 horas también fueron seleccionados como regresores importantes. Esto era esperable ya que los datos a corto plazo presentan una estacionalidad de carácter diario. En el caso de la radiación solar, su dinámica es muy similar todos los días, en donde la magnitud aumenta desde el amanecer partiendo de valores casi nulos, alcanzando un máximo para luego descender a valores cercanos a cero nuevamente al anochecer. Cabe destacar que la magnitud máxima varía a lo largo del año, siendo mayor en verano y menor en invierno. Además, hay días en los que la magnitud de la radiación es extremadamente baja, ya sea de forma parcial o durante todo el día, y esto se debe principalmente a la nubosidad que bloquea la radiación solar.

La temperatura, por otro lado, exhibe una dinámica que está influenciada por la hora del día, con un aumento durante las horas de sol y un descenso durante la noche.

### 5.1.6. Optimización final del número de neuronas

En esta sección se muestran los resultados de la optimización final de hiperparámetros de la red perceptrón multicapa.

La tabla a continuación, se muestra el número de neuronas seleccionadas en la optimización final de hiperparámetros. Dejando fijo los valores de los regresores según la selección realizada en la fase anterior.

Tabla 5.4: Número de neuronas en la capa oculta seleccionadas en la optimización final de hiperparámetros, con los regresores optimizados.

Conjunto de datos	Nro neuronas para regresores optimizados	Número de regresores
Radiación solar a 10m	45	35
Temperatura a 10m	95	10
Radiación solar diaria	55	25
Temperatura diaria	45	60
Precipitación diaria	15	10

En la tabla 5.4 se aprecia que en algunos casos el número de neuronas ocultas baja considerablemente como es el caso de radiación solar a 10 minutos y precipitación diaria, en donde se reducen a la mitad. Esto coincide con una reducción significativa con el número de regresores en la optimización previa. Esto se podría deber a que un número menor de regresores puede conllevar a recibir menos ruido de la señal, y por lo tanto un número menor de neuronas puede ser suficiente para aprender de igual o mejor manera la señal.

## 5.2. Entrenamiento LSTM

En esta sección se muestran los resultados de la optimización de la red LSTM para los distintos conjuntos.

Para la red LSTM, a diferencia de la red perceptrón multicapa, debido a los costos computacionales y el tiempo derivado de esto, solo se optimiza una única vez la dimensión del estado oculto y la dimensión del estado de la celda, en donde ambas comparten el mismo valor. Para llevar a cabo esta búsqueda de hiperparámetros, se prueban distintos valores de dimensión de estado oculto y estado de la celda, Finalmente, se elige la dimensión según el error obtenido en el conjunto de validación.

En la tabla 5.5 se muestra la dimensión del estado oculto y de la celda oculta seleccionadas en la optimización de hiperparámetros.

Tabla 5.5: Dimensión del estado oculto de la celda LSTM seleccionadas en la optimización.

Conjunto de datos	Dimensión del estado de la celda óptima
Radiación solar a 10m	56
Temperatura a 10m	126
Radiación solar diaria	66
Temperatura diaria	41
Precipitación diaria	41

En la tabla 5.5 se observa que la dimensión del estado oculto seleccionado es relativamente baja, a pesar de haber probado con un número mayor. Esto podría deberse, al igual que en el caso de la red de perceptrón multicapa, a la presencia de sobreajuste, lo que resultaría en una disminución del rendimiento de la red para un número excesivamente elevado de unidades ocultas.

### 5.3. Modelo de referencia y validación de resultados

Para poder validar los resultados obtenidos, se emplean los resultados de un modelo de predicción implementado en la tesis [78], pero entrenado con los conjuntos de datos de esta memoria. La elección de este modelo se debe a que ya se cuenta con código en el lenguaje *Matlab*, además de haber obtenido resultados satisfactorios en la predicción de datos de radiación solar [11].

Este modelo de predicción adoptado en la presente investigación es el modelo difuso *Takagi Sugeno* tipo 2 [79], cuya naturaleza difusa se diferencia de los modelos implementados en esta memoria. Específicamente, el modelo se implementó mediante el algoritmo *Fuzzy C-Means* como algoritmo de *clustering* difuso, una función de membresía tipo *Gaussiana* y una consecuencia del modelo lineal.

### 5.4. Resultados de predicción

En esta sección se presentan los resultados de los modelos de predicción obtenidos en las distintas fuentes de datos.

Para obtener los resultados, se realiza la evaluación de las predicciones a diferentes horizontes de tiempo, específicamente a 72, 144 y 288 pasos hacia adelante en el caso de los datos de radiación solar y temperatura del aire, los cuales son registrados cada 10 minutos. En cuanto a los datos diarios de radiación solar, temperatura del aire y precipitación, se realizan predicciones a 1, 14 y 28 pasos hacia adelante, lo que corresponde a 1, 14 y 28 días, respectivamente. Esta elección se justifica por la necesidad de mostrar la evolución dinámica de las predicciones a medida que se acerca al horizonte de predicción deseado. Para los gráficos, se opta por mostrar los resultados a 144 pasos en el caso de los datos a 10 minutos, y 14 días para el caso de los datos diarios, ya que representan el punto intermedio del horizonte de predicción intermedia.

Con respecto a los parámetros, en la tabla 5.6 se expone la cantidad de parámetros correspondientes a cada modelo de predicción. Para el caso de perceptrón multicapa, estos se refieren a los pesos de las capas ocultas y los sesgos, mientras que para los modelos LSTM, estos incluyen los pesos de la entrada, los pesos de la capa de recurrencia y los pesos de la capa de salida.

En el caso específico del modelo difuso *Takagi Sugeno*, los parámetros involucran la desviación estándar y el centro de cada conjunto difuso, así como las constantes de las consecuencias del modelo.

Tabla 5.6: Número de parámetros de los modelos de predicción.

<b>Modelos de predicción</b> <b>Fuente de datos</b>	<b>Perceptrón multicapa</b> # parámetros	<b>LSTM</b> # parámetros	<b>Difuso Takagi Sugeno</b> # parámetros
<b>Radiación solar a 10m</b>	1666	13049	441
<b>Temperatura a 10m</b>	1141	64639	846
<b>Radiación solar diaria</b>	1486	18019	366
<b>Temperatura diaria</b>	1396	7094	318
<b>Precipitación diaria</b>	181	-	108

De la tabla 5.6 se desprende que la cantidad de parámetros del modelo difuso es inferior a la del modelo de perceptrón multicapa, y a su vez, este último tiene menos parámetros que el modelo LSTM. En particular, el modelo de perceptrón multicapa posee entre dos y cuatro veces más parámetros que el modelo difuso. En contraste, el modelo LSTM cuenta con una cantidad considerablemente mayor, con entre 13 y 63 veces más parámetros que el modelo de perceptrón multicapa, y entre 20 y 80 veces más parámetros que el modelo difuso.

A continuación se muestra la comparación de la predicción de los 3 modelos de predicción para el caso de la radiación solar de 10 minutos a 24 horas (144 pasos), evaluados en el conjunto test:

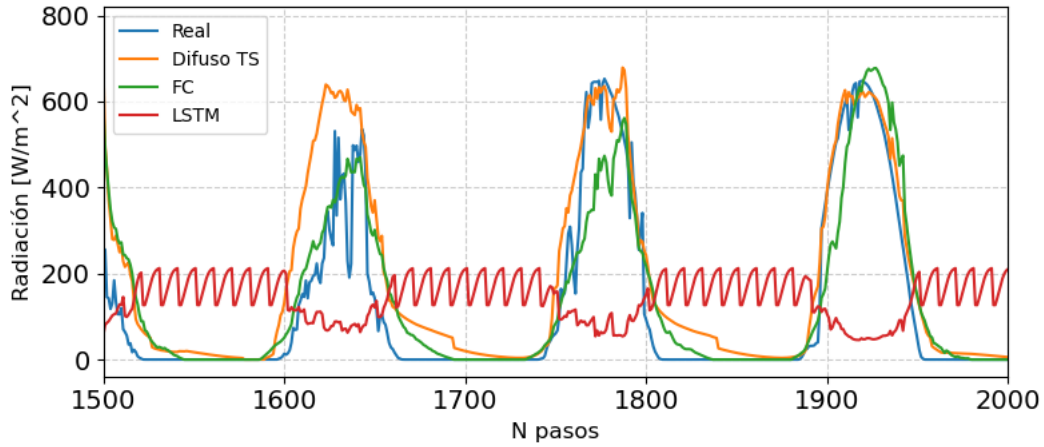


Figura 5.1: Comparación de la predicción de los distintos modelos de predicción de radiación solar de 10 minutos a 144 pasos, evaluados en el conjunto de test.

De la figura 5.1 se puede observar que tanto el modelo de perceptrón multicapa como el modelo difuso, son capaces de capturar la dinámica de la radiación solar, siguiendo la curva característica de radiación durante el día y alcanzando valores cercanos a cero durante la noche. En cambio, la red LSTM presenta un desempeño deficiente en la predicción de la radiación solar, obteniendo resultados que parecen ser valores promedio de los datos.

A partir de la figura 5.1 se puede apreciar que tanto el modelo de perceptrón multicapa como el modelo difuso exhiben habilidades para rastrear con precisión la dinámica de la radiación solar, alcanzando valores próximos a cero durante la noche y siguiendo una curva característica de radiación durante el día. En contraste, se puede observar que la red LSTM no se desempeña satisfactoriamente, presentando predicciones que, posiblemente, corresponden a valores cercanos al promedio de los datos.

En la tabla 5.7 se muestran los errores de predicción de los 3 modelos de predicción para el caso de la radiación solar de 10 minutos evaluados a 12 horas (72 pasos), a 24 horas (144 pasos) y 48 horas (288 pasos), evaluados en el conjunto test:

Tabla 5.7: Error de predicción de los distintos modelos de predicción de radiación solar a 10 minutos, evaluados en el conjunto de test.

Radiación solar a 10m	Difuso TS		Perceptrón multicapa		LSTM	
	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]
Número de pasos						
a 12 horas (72 Pasos)	<b>105.68</b>	68.52	119.48	<b>68.28</b>	175.95	146.92
a 24 horas (144 Pasos)	147.73	91.95	<b>125.82</b>	<b>70.09</b>	215.52	184.69
a 48 horas (288 Pasos)	247.86	187.04	180.20	<b>110.01</b>	<b>162.66</b>	134.29

En la tabla 5.7 se puede apreciar que para las predicciones a 12 horas, el modelo difuso

presenta el menor RMSE, mientras que para el MAE, el modelo perceptrón multicapa es el que muestra un desempeño ligeramente superior, aunque la diferencia es mínima en comparación con el modelo difuso. Por otro lado, para las predicciones a 24 y 48 horas, el modelo perceptrón multicapa presenta el menor MAE, mientras que el modelo LSTM muestra el menor RMSE.

De acuerdo con la tabla 5.7, se observa que el modelo perceptrón multicapa logra un desempeño significativamente mejor que el modelo difuso en las predicciones a 24 y 48 horas, aunque en las predicciones a 12 horas obtiene un desempeño ligeramente peor que el modelo difuso. La capacidad de la red perceptrón multicapa para aprender mejor la dinámica de la radiación solar a corto plazo puede explicarse por su mayor número de parámetros, el cual es aproximadamente cuatro veces mayor que el del modelo difuso. En este caso, a diferencia del modelo LSTM, la cantidad de datos disponibles fue suficiente para un entrenamiento adecuado. Como resultado, se observa una diferencia de aproximadamente el 40 % en RMSE y del 60 % en MAE entre ambos modelos en las predicciones a 48 horas.

En la figura 5.1 se puede apreciar que la red LSTM no logra capturar la dinámica de la señal de radiación, limitándose a predecir valores cercanos al promedio de los datos. No obstante, a pesar de no aprender la dinámica, obtiene el menor RMSE para el caso de la predicción a 48 horas. Esta aparente contradicción se explica por la forma en que se calculan los errores de predicción. El MAE castiga de igual manera todos los errores, mientras que el RMSE penaliza más fuertemente los errores que se encuentran más alejados del dato real. En el caso de la red LSTM, al predecir valores cercanos al promedio, el error se mantiene acotado en todo momento, sin alcanzar grandes magnitudes. En cambio, para los otros modelos, las predicciones a 48 horas se alejan más de los datos reales, y pueden surgir situaciones en las que el día estuvo nublado pero el modelo predice una radiación solar alta, contribuyendo a un RMSE elevado.

En las predicciones a corto y mediano plazo de la radiación solar, se observa una relación directa entre la amplitud de las observaciones pasadas y la amplitud de las predicciones. Por otro lado, es posible notar en los datos reales que existen días con una radiación significativamente baja en comparación con los días contiguos, lo que puede deberse a la presencia de nubes. En estos casos, el modelo de predicción puede fallar al predecir la radiación, especialmente si los cambios en la señal son irregulares o atípicos. No obstante, si estos cambios son temporales, la señal se vuelve más fácil de predecir. Cabe destacar que la información sobre la cobertura de nubes es difícil de capturar únicamente con los datos de radiación solar, lo que presenta un desafío para la predicción tanto a corto como a mediano plazo.

A continuación se muestra la comparación de la predicción de los 3 modelos de predicción para el caso de la temperatura de 10 minutos a 24 horas (144 pasos), evaluados en el conjunto test:



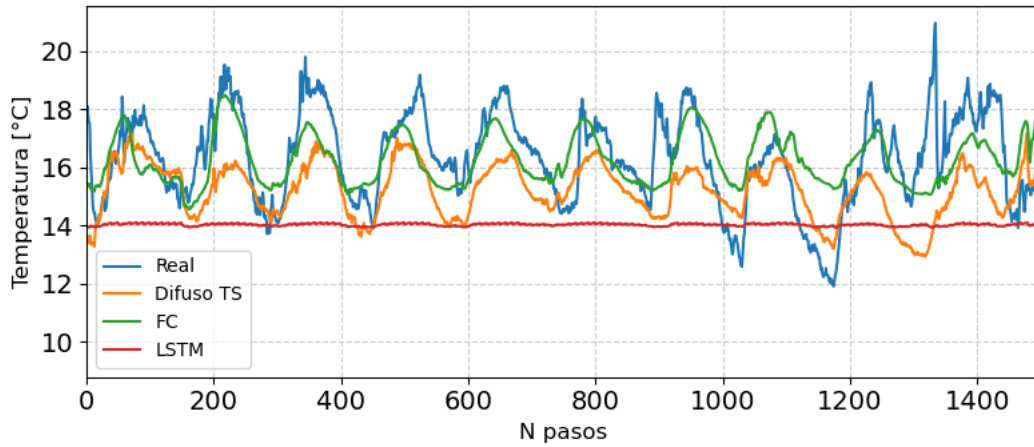


Figura 5.2: Comparación de la predicción de los distintos modelos de predicción de temperatura de 10 minutos a 144 pasos, evaluados en el conjunto de test.

De la figura 5.2 se observa que tanto el modelo difuso como el modelo de perceptrón multicapa siguen la dinámica de la temperatura real, aunque con una magnitud menor que la real. En el caso del modelo LSTM este no fue capaz de seguir la dinámica de la señal real.

En la siguiente tabla se muestran los errores de predicción de los 3 modelos de predicción para el caso de temperatura de 10 minutos evaluados a 12 horas (72 pasos), a 24 horas (144 pasos) y 48 horas (288 pasos), evaluados en el conjunto test:

Tabla 5.8: Error de predicción de los distintos modelos de predicción de temperatura a 10 minutos, evaluados en el conjunto de test.

Temperatura a 10m	Difuso TS		Perceptrón multicapa		LSTM	
	RMSE [°C]	MAE [°C]	RMSE [°C]	MAE [°C]	RMSE [°C]	MAE [°C]
Número de pasos						
a 12 horas (72 Pasos)	1.94	1.57	<b>1.77</b>	<b>1.39</b>	3.03	2.61
a 24 horas (144 Pasos)	<b>1.76</b>	<b>1.40</b>	1.81	1.41	2.49	2.09
a 48 horas (288 Pasos)	2.06	1.70	<b>1.94</b>	<b>1.51</b>	2.51	2.09

Según la tabla 5.8, se puede observar que el modelo de perceptrón multicapa presenta el menor error en las predicciones a 12 y 48 horas, medido por las métricas MAE y RMSE. Por otro lado, el modelo difuso obtiene el menor error para la predicción a 24 horas.

En términos generales, el modelo de perceptrón multicapa demuestra un mejor desempeño en la tarea de predicción de temperatura a corto plazo, específicamente para las predicciones a 12 y 48 horas. Sin embargo, se observa que su desempeño es levemente peor que el modelo difuso en la predicción a 24 horas.

La diferencia en el desempeño entre ambos modelos también puede ser atribuida a la cantidad de parámetros que poseen. En este caso, se encuentra que el modelo de perceptrón

multicapa cuenta con 1.3 veces más parámetros que el modelo difuso. Sin embargo, la diferencia es menor en comparación con el caso de predicción de radiación solar a 10 minutos, lo que explica en parte la similitud en los resultados obtenidos.

En la figura 5.2 se puede visualizar que las predicciones generadas por los modelos de perceptrón multicapa y difuso siguen la dinámica de la señal real, aunque presentan una magnitud ligeramente inferior en comparación con los valores reales. Esto podría ser explicado por el hecho de que el conjunto de entrenamiento se encuentra compuesto principalmente por datos recopilados en primavera, mientras que el conjunto de prueba corresponde a datos recolectados en invierno. Es decir, la dinámica de la temperatura en ambos periodos puede ser diferente, por ejemplo, la magnitud de la temperatura puede ser más elevada en verano en comparación con la primavera.

A continuación se muestra la comparación de la predicción de los 3 modelos de predicción para el caso de la radiación solar diaria a 14 días (14 pasos), evaluados en el conjunto test:

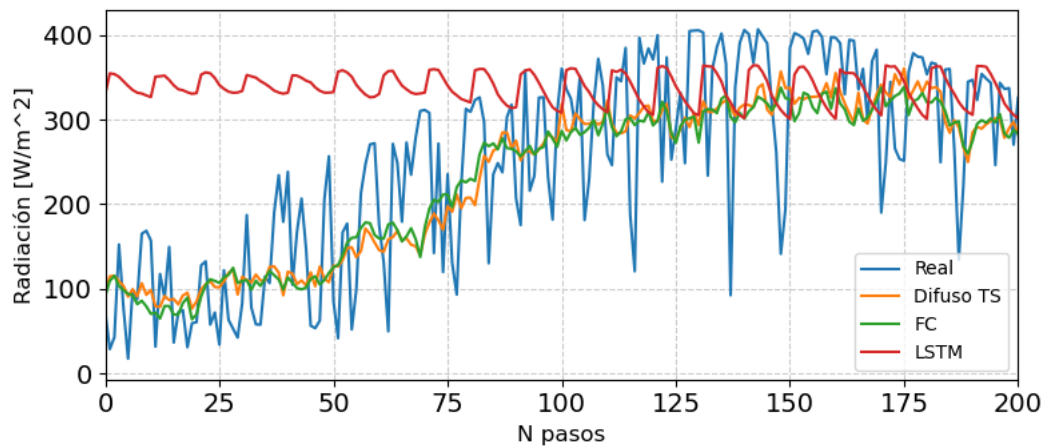


Figura 5.3: Comparación de la predicción de los distintos modelos de predicción de radiación solar diaria a 14 pasos, evaluados en el conjunto de test.

De la figura 5.3 se observa, que los datos reales presentan una gran variación día a día, aunque siguen una tendencia. Esta tendencia es capturada por los modelos de perceptrón multicapa y difuso. Mientras que el la red LSTM predice valores erróneos en torno a un valor.

En la tabla 5.9 se muestra los errores de predicción de los 3 modelos de predicción para el caso de la radiación solar diaria evaluados a 1 día (1 paso), a 14 días (14 días) y 28 días (28 pasos), evaluados en el conjunto test:

Tabla 5.9: Error de predicción de los distintos modelos de predicción de radiación solar diaria, evaluados en el conjunto de test.

Radiación solar diaria Número de pasos	Difuso TS		Perceptrón multicapa		LSTM	
	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]	RMSE [W/m <sup>2</sup> ]	MAE [W/m <sup>2</sup> ]
a 1 día (1 Paso)	69.00	55.40	68.13	55.03	117.94	98.53
a 14 días (14 Pasos)	72.40	59.41	73.63	60.14	199.85	169.75
a 28 horas (28 Pasos)	76.48	63.67	78.81	65.40	179.18	151.29

De la tabla 5.9 se puede apreciar que el modelo difuso tiene un mejor desempeño, aunque con una diferencia leve con respecto al perceptrón multicapa. El mejor modelo para las predicciones a 1 día es el perceptrón multicapa, pero para 14 días y 28 días, el modelo difuso tiene un mejor desempeño. Esto podría deberse a que, en este caso, la cantidad de datos disponibles es considerablemente menor que en los casos anteriores de datos a 10 minutos. Por tanto, el perceptrón multicapa en este caso se ve perjudicado al tener cuatro veces la cantidad de parámetros que el modelo difuso.

De la figura 5.3 se observa que la radiación solar diaria tiene una gran variación entre días contiguos, sin embargo, sigue una marcada tendencia dada por la estacionalidad. Cuando la amplitud es mayor, esta tendencia se debe a datos de primavera-verano y cuando son bajos, a los datos de invierno-otoño. Los modelos difuso y perceptrón multicapa no son capaces de aprender la dinámica completa de la señal, pero sí su tendencia. Esta variación posiblemente se deba, como se mencionó anteriormente, a la nubosidad presente durante el día.

A continuación se muestra la comparación de la predicción de los 3 modelos de predicción para el caso de la radiación solar diaria a 14 días (14 pasos), evaluados en el conjunto test:

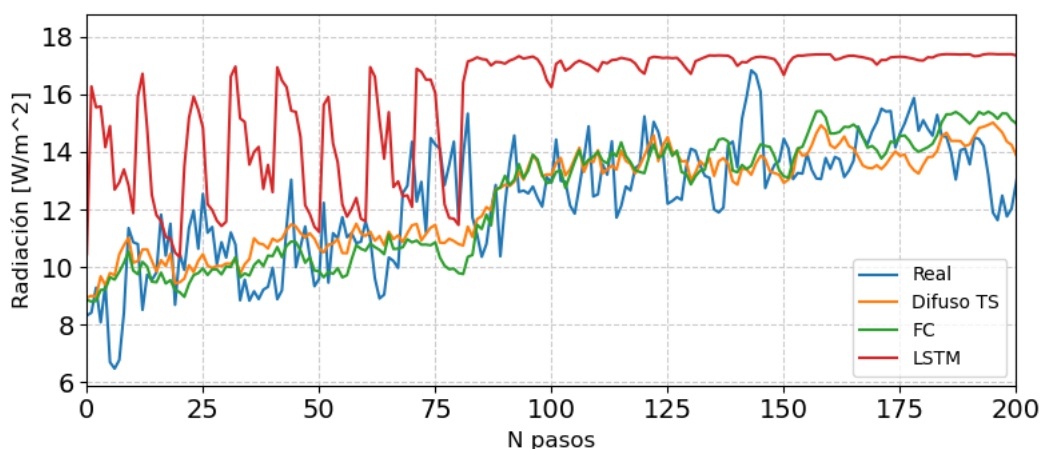


Figura 5.4: Comparación de la predicción de los distintos modelos de predicción de radiación solar diaria a 14 pasos, evaluados en el conjunto de test.

De la figura 5.4 se observa que los datos de radiación solar diaria siguen cierta tendencia

ascendente que es capturada por los modelos de perceptrón multicapa y difuso. Mientras que la red LSTM no es capaz de capturarlo.

En la tabla 5.10 se muestra los errores de predicción de los 3 modelos de predicción para el caso de la temperatura diaria evaluados a 1 día (1 paso), a 14 días (14 días) y 28 días (28 pasos), evaluados en el conjunto test:

Tabla 5.10: Error de predicción de los distintos modelos de predicción de temperatura diaria, evaluados en el conjunto de test.

Temperatura diaria	Difuso TS		Perceptrón multicapa		LSTM		
	Número de pasos	RMSE [°C]	MAE [°C]	RMSE [°C]	MAE [°C]	RMSE [°C]	MAE [°C]
a 1 día (1 Paso)		1.03	0.81	<b>1.01</b>	<b>0.79</b>	1.52	1.23
a 14 días (14 Pasos)		<b>1.41</b>	<b>1.13</b>	1.47	1.19	4.11	3.66
a 28 horas (28 Pasos)		<b>1.49</b>	<b>1.18</b>	1.66	1.32	4.43	3.92

De la tabla 5.10 se aprecia que el mejor modelo para las predicciones a 1 día es el de perceptrón multicapa pero para las predicciones a 14 días y 28 días, el modelo difuso, aunque la diferencia es leve.

De la figura 5.4 se puede observar que la temperatura diaria presenta variaciones entre días consecutivos, aunque estas no son tan pronunciadas como en el caso de la radiación solar. Es posible que esto se deba a las diferencias climáticas que se presentan día a día y que tienen un carácter aleatorio. Los modelos no son capaces de seguir esta dinámica rápida, sino que aprenden la tendencia general. La subida y bajada de magnitud de esta tendencia depende de la estación del año.

En la tabla 5.10 se puede apreciar que el modelo difuso presenta el mejor desempeño para la predicción de la temperatura diaria. Sin embargo, la diferencia con el modelo perceptrón multicapa es pequeña. La razón de esto, al igual que para el caso de la radiación solar, podría ser la cantidad limitada de datos disponibles. En comparación con los datos de radiación solar a 10 minutos, los datos de temperatura diaria son escasos, lo que limita el aprendizaje de patrones más complejos en los modelos. Además, el modelo perceptrón multicapa tiene cuatro veces más parámetros que el modelo difuso, lo que le permite aprender mejor la tendencia de los datos a corto plazo.

En la figura 5.5 se muestra la comparación de la predicción de los 2 modelos de predicción para el caso de la precipitación diaria a 14 días (14 pasos), evaluados en el conjunto test. Como se había explicado anteriormente, para esta variable no se utiliza el modelo LSTM debido a que en pruebas preliminares los resultados con errores muy grandes, lo que se debe probablemente a la poca cantidad de datos, usando un modelo con muchos parámetros intentando resolver un problema tan complejo como es la precipitación, que tiene un carácter más estocástico que las variables de radiación solar y temperatura del aire.

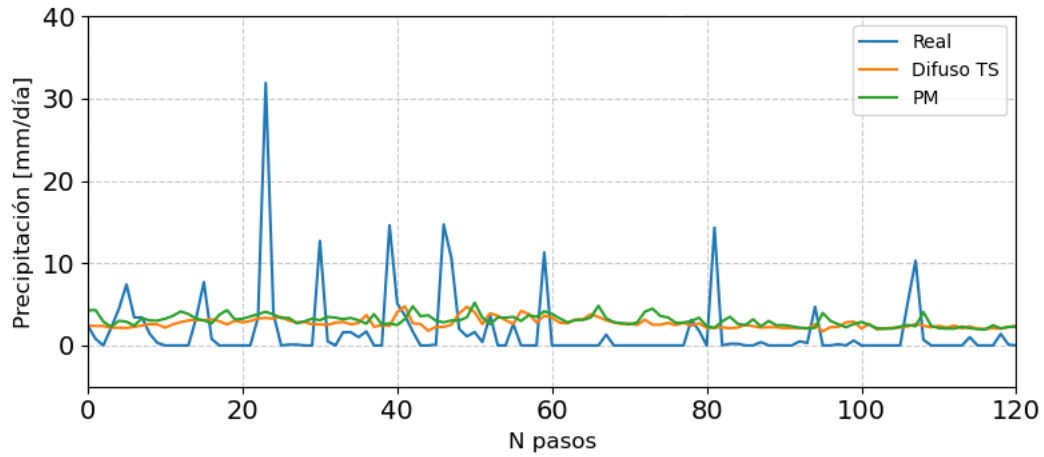


Figura 5.5: Comparación de la predicción de los distintos modelos de predicción de precipitación diaria a 14 pasos, evaluados en el conjunto de test.

En la tabla 5.5 se muestra los errores de predicción de los 2 modelos de predicción para el caso de la precipitación diaria evaluados a 1 día (1 paso), a 14 días (14 días) y 28 días (28 pasos), evaluados en el conjunto test:

Tabla 5.11: Error de predicción de los distintos modelos de predicción de precipitación diaria, evaluados en el conjunto de test.

Precipitación diaria	Difuso TS		Perceptrón multicapa	
	RMSE [mm/día]	MAE [mm/día]	RMSE [mm/día]	MAE [mm/día]
Número de pasos				
a 1 día (1 Paso)	<b>5.41</b>	<b>3.19</b>	5.46	3.28
a 14 días (14 Pasos)	<b>5.62</b>	<b>3.48</b>	5.67	3.60
a 28 horas (28 Pasos)	<b>5.65</b>	<b>3.47</b>	5.77	3.77

De la figura 5.5, se aprecia que ambos modelos presentan un comportamiento similar, en donde se predice un valor cercano al promedio de los datos, que según la tabla 5.1, es de 2.7 [mm/día]. Este comportamiento era esperable debido a que la naturaleza de las precipitaciones es de carácter caótico, lo que las hace muy variables. Por lo tanto, intentar obtener una predicción a varios pasos, en particular 14 pasos, es una tarea sumamente complicada, sobre todo si solo se utiliza información de las precipitaciones y no de otras variables climatológicas que podrían tener alguna correlación. Además, los datos utilizados son pocos, lo que implica que no se puede extraer tanta información.

De la tabla 5.6, se puede notar que la cantidad de parámetros en este caso es menor en comparación con otras fuentes de datos. Esto podría deberse a que, en la optimización de la arquitectura del modelo, se concluye que, a partir de la poca cantidad de datos disponibles y la naturaleza aleatoria de las precipitaciones, no se puede extraer tanta información de la dinámica real, sino solo de su promedio. Por lo tanto, no es necesario contar con tantos parámetros.

De la tabla 5.11, se aprecia que los valores de errores RMSE y MAE entre ambos modelos son muy similares para todos los pasos, mientras que existe una marcada diferencia en los valores de RMSE y MAE. Al predecir valores casi constantes, el error se mantendrá relativamente constante ya que, en general, no hay mucha lluvia, pero en los casos en los que sí llueve realmente, el error aumenta drásticamente, incrementando el valor de RMSE.

En esta sección se analizan los resultados de la sección de modelos de predicción.

De las figuras 5.1, 5.2, 5.3 y 5.4, se aprecia que la red LSTM obtiene un mal desempeño, no siendo capaz de aprender la dinámica de la señal real. En teoría, esta red puede aprender dependencias temporales más complejas que los otros modelos, pero a costa de tener un mayor número de parámetros y, por ende, de más datos de entrenamiento.

De la tabla 5.6, se puede observar que la cantidad de parámetros de la red LSTM excede la cantidad de datos mostrados en la tabla 5.1. Sin importar la cantidad de épocas usadas en el entrenamiento, debido al hecho descrito anteriormente, la red se sobreajustará a los datos de entrenamiento de 1 paso, lo que explica por qué se obtienen esas predicciones para varios pasos.

Por otra parte, se observa una tendencia en la que las predicciones a menos pasos obtienen mejores resultados que las predicciones a más pasos. Este resultado era esperado, ya que para realizar las predicciones se está utilizando un modelo que solo predice a 1 paso. Para la predicción de varios pasos, se usa la salida del modelo como entrada para calcular el siguiente paso. De esta manera, en cada iteración se agrega ruido a la señal y, por lo tanto, error, que aumenta a medida que el horizonte de predicción aumenta.

## 5.5. Discusión

En este capítulo se lleva a cabo el entrenamiento de los modelos de predicción Perceptrón Multicapa y LSTM, junto con sus respectivas optimizaciones de hiperparámetros. Adicionalmente, se realiza una comparación de desempeño con un modelo difuso de tipo *Takagi Sugeno* tipo 2. Para llevar a cabo la comparación, se realizan predicciones recursivas a 12 horas, 24 horas y 48 horas para los datos registrados cada 10 minutos, y a 1 día, 14 días y 28 días para los datos diarios. La comparación se lleva a cabo mediante el cálculo del error de predicción utilizando las métricas de RMSE y MAE.

Como resultado, se ha determinado que para los datos de radiación solar registrados cada 10 minutos, el Perceptrón Multicapa tiene el mejor desempeño. En cuanto a los datos de temperatura registrados cada 10 minutos, se ha obtenido que el modelo Perceptrón Multicapa tiene el mejor desempeño para predicciones a 12 horas y 48 horas, aunque a 24 horas su desempeño es levemente peor que el de la red difusa. En el caso de los datos diarios de radiación solar diaria, temperatura diaria y precipitación diaria, el modelo difuso es el que presenta el mejor resultado, lo cual se debe a su menor número de parámetros y al hecho de que hay menos datos diarios disponibles.

En todos los casos, la red LSTM se ha mostrado incapaz de predecir de manera correcta a muchos pasos, lo que podría deberse a que no puede aprender de manera efectiva la dinámica de la señal debido a su gran cantidad de parámetros en comparación con los datos disponibles.

Los resultados de las optimizaciones de hiperparámetros de los modelos Perceptrón Multicapa y LSTM serán utilizados en la arquitectura de los modelos de intervalo de predicción en el siguiente capítulo. En particular, se utilizarán los resultados del Perceptrón Multicapa para los modelos de intervalo de predicción *Joint Supervision* basado en Perceptrón Multicapa (PM-JS) y *Quality Driven* basado en el modelo Perceptrón Multicapa (PM-QD), mientras que los resultados de la red LSTM serán utilizados por los modelos de intervalo de predicción *Joint Supervision* basado en el modelo LSTM (LSTM-JS) y *Quality Driven* basado en el modelo LSTM (LSTM-QD).

# Capítulo 6

## Diseño e implementación de intervalos de predicción

En este capítulo se desarrolla el diseño e implementación de los intervalos de predicción.

### 6.1. Modelos de intervalos de predicción neuronales

En esta sección se presenta el diseño de los modelos de intervalos neuronales que se utilizaron en este trabajo.

Considerando el estado del arte, en un análisis bibliográfico se evaluaron distintos modelos de intervalos de predicción comparando su desempeño [11]. Dentro de los métodos de intervalos neuronales, se encontró que los modelos de *Joint Supervision* (JS), *Quality Driven* (QD) y *Lower Upper Bound Estimation* (LUBE) presentaban el mejor desempeño para modelos de perceptrón multicapa y LSTM.

Sin embargo, LUBE tiene la particularidad de que su función de costo es no lineal, lo que requiere de una forma de optimización no lineal más compleja que el descenso del gradiente, utilizado en QD y JS. Debido a que QD es una variante de LUBE, se optó por utilizar los modelos QD y JS para este trabajo.

#### 6.1.1. Preprocesamiento

Los datos son separados en conjuntos de entrenamiento, validación y test en la misma proporción que se utilizó para los modelos de predicción, es decir, 60 %, 20 % y 20 %, respectivamente.

Posteriormente, se normalizan utilizando las mismas técnicas que se emplearon en los modelos de predicción. Los datos de temperatura se normalizan mediante la fórmula 5.1, mientras que los datos de radiación solar y precipitación diaria se normalizan utilizando la fórmula 5.2.

#### 6.1.2. *JointSupervision* (JS)



### 6.1.2.1. Entrenamiento

Se requiere fijar el PICP deseado antes de iniciar el entrenamiento, ya que éste no finaliza hasta que se cumpla una condición de término. Existen dos condiciones posibles: la primera es alcanzar un valor de PICP evaluado en el conjunto de validación cercano al PICP deseado y la segunda es alcanzar un valor muy alto de  $\lambda$ .

En todos los casos, se fija el PICP en un 90 %. Esto significa que se entrena hasta lograr que la predicción de intervalos contenga el 90 % de los datos en el conjunto de validación. Se escoge este valor porque si se selecciona uno demasiado alto, como el 99 %, el PINAW resultante podría ser demasiado grande y no aportar información valiosa en cuanto a la incertidumbre. Si, por otro lado, se selecciona un valor muy bajo, se excluyen muchos datos del intervalo.

Para el entrenamiento, se realiza una iteración sobre distintos valores de  $\lambda$ , con una diferencia logarítmica entre ellos. Específicamente, se prueban los valores  $\lambda = 1, 10, 10^2, 10^3, \dots, 10^{15}$ .

Si los valores de PICP son mayores al valor deseado, se modifica el valor actual de  $\lambda$  disminuyéndolo hasta que se cumpla la condición de PICP en el conjunto de validación.

### 6.1.2.2. Arquitectura

Para los modelos de intervalo se usarán las arquitecturas óptimas encontradas en los modelos de predicción. En el caso de los intervalos basados en el perceptrón multicapa, esto sería utilizar el número de neuronas en la capa oculta  $n_h$  óptimos de la tabla 5.4, y los regresores óptimos que se puede observar en la tabla 5.3.

Para el caso de los intervalos basados en la red LSTM, esto sería utilizar el número de neuronas en el estado de la celda mostradas en la tabla 5.5.

## 6.1.3. *Quality Driven* (QD)

### 6.1.3.1. Entrenamiento

A diferencia del método *Joint Supervision*, el método *Quality Driven* requiere entrenar el modelo solo una vez.

Existen dos hiperparámetros:  $\lambda$  y el factor de suavizamiento  $s$ . En [25], se utilizaron los mismos valores ( $\lambda = 15$ ,  $s = 160$ ) para varios conjuntos de datos de diferentes tamaños. Aunque en algunos de los conjuntos de datos se variaron los valores, se mantuvieron cercanos al valor de  $\lambda = 15$ . Por tanto, se opta por utilizar estos valores y dejarlos fijos para todos los conjuntos de datos. En caso de que no se cumpla el requisito de alcanzar el porcentaje de cobertura requerido en el conjunto de entrenamiento, se probarán distintos valores de  $\lambda$  y  $s$  hasta alcanzar dicho requisito.

Otro hiperparámetro que debe seleccionarse es el porcentaje de cobertura deseado. En este caso, se elige un 90 %, al igual que en el método *Joint Supervision*, por las mismas razones.

### 6.1.3.2. Arquitectura del modelo

Para los modelos de intervalo se usarán las arquitecturas óptimas encontradas en los modelos de predicción. En el caso de los intervalos basados en el perceptrón multicapa, esto sería utilizar el número de neuronas en la capa oculta  $n_h$  óptimos de la tabla 5.4, y los regresores óptimos que se pueden observar en la tabla 5.3.

Para el caso de los intervalos basados en la red LSTM, esto sería utilizar el número de neuronas en el estado de la celda mostradas en la tabla 5.5

## 6.2. Implementación de modelos de intervalos neuronales

### 6.2.1. Comparación y validación de resultados

Para poder validar los resultados obtenidos, se utilizan los resultados de un modelo de predicción de intervalo implementado en la tesis [78], pero entrenado en los conjuntos de datos de esta memoria. La elección de este se debe a que ya se dispone de código en el lenguaje *Matlab*, además de haber obtenido buenos resultados en la predicción de intervalos en datos de radiación solar [11]. Este modelo de predicción de intervalos se llama *Fuzzy Numbers* [80], que, a diferencia de los modelos de intervalos usados en la memoria, requiere de un modelo base ya entrenado. Para este fin, se utiliza el modelo difuso *Takagi Sugeno* ya enunciado en el capítulo anterior.

### 6.2.2. Resultados de predicción de intervalos

En esta sección se presentan los resultados de los modelos de predicción de intervalos obtenidos en las distintas fuentes de datos. Los resultados se evalúan a 12 horas, 24 horas y 48 horas para el caso de los datos a 10 minutos, y a 1, 14 y 28 días para el caso de los datos diarios. Esta elección se hizo para mostrar la evolución dinámica de los resultados a medida que el número de pasos llega al horizonte de predicción requerido.

En cuanto a los gráficos, se decidió mostrar los resultados a 24 horas para el caso de los datos a 10 minutos y a 14 días para el caso de los datos diarios. Esta elección se hizo ya que estos representan el punto intermedio del horizonte de predicción.

En la tabla 6.1, se muestra la cantidad de parámetros para cada modelo de predicción de intervalo. En el caso de JS basado en perceptrón multicapa, estos corresponden a los pesos de las capas ocultas, los sesgos y la capa de salida, que en este caso tiene 3 salidas en lugar de 1. Para el caso de perceptrón multicapa basado en JS, estos son los pesos de las capas ocultas, los sesgos y la capa de salida, que en este caso tiene 2 salidas en lugar de 1.

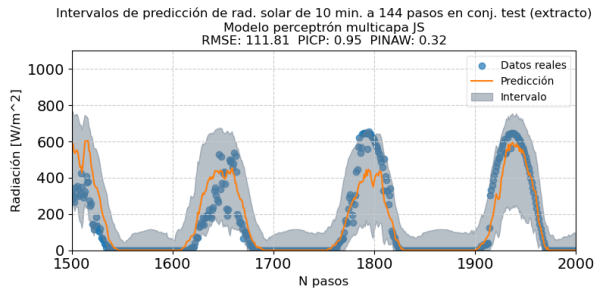
En cuanto a los modelos de JS basados en LSTM, estos se refieren a los pesos de la entrada, los pesos de la capa de recurrencia y los pesos de la capa de salida, que en este caso tiene 3 salidas. Finalmente, en el caso del método *Fuzzy Numbers* basado en el modelo *Takagi Sugeno*, tal como en el modelo solo de predicción *Takagi Sugeno*, se tienen centros y desviaciones estándar de los centros difusos, pero la diferencia radica en que las constantes

de las consecuencias se definen con dos parámetros: el centro y el *spread*. Esto se hace con el fin de establecer el intervalo. Se aprecia que la diferencia con respecto a los modelos de solo predicción es pequeña debido al tipo de intervalo que se utiliza, lo que conlleva un número muy similar de parámetros, solo ligeramente superior.

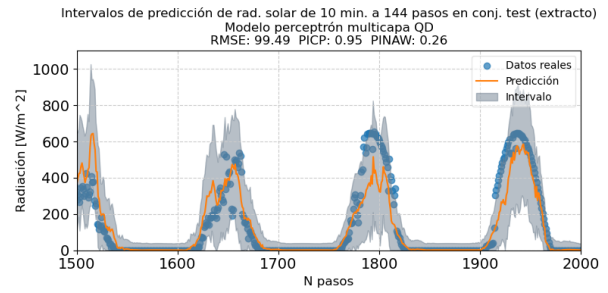
Tabla 6.1: Número de parámetros de los modelos de intervalos de predicción.

Modelos de intervalos	JS -Perceptrón multicapa	QD-Perceptron multicapa	JS-LSTM	QD-LSTM	Fuzzy Numbers - Difuso TS
Fuente de datos	# parámetros	# parámetros	# parámetros	# parámetros	# parámetros
Radiación solar a 10m	1758	1712	13163	13106	588
Temperatura a 10m	1333	1237	64893	64766	1128
Radiación solar diaria	1598	1542	18153	18086	488
Temperatura diaria	1488	1442	7178	7136	424
Precipitaciones diarias	213	197	-	-	180

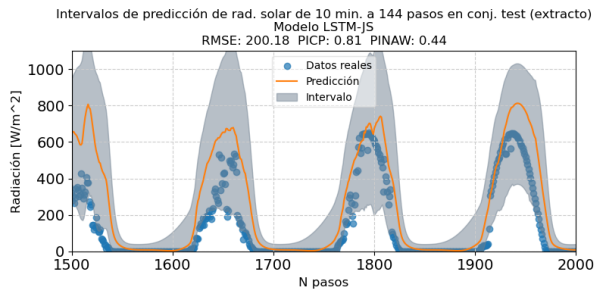
A continuación se muestra la comparación de la predicción de intervalos de los distintos modelos para el caso de la radiación solar de 10 minutos a 24 horas (144 pasos), evaluados en el conjunto test:



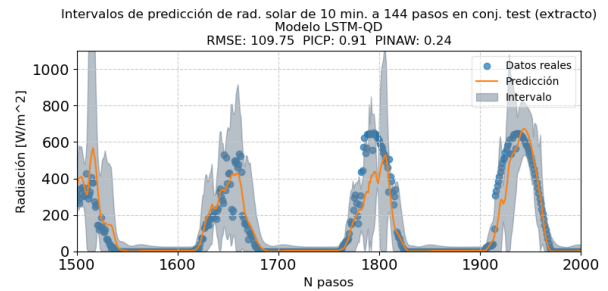
(a) PM - JointSupervision



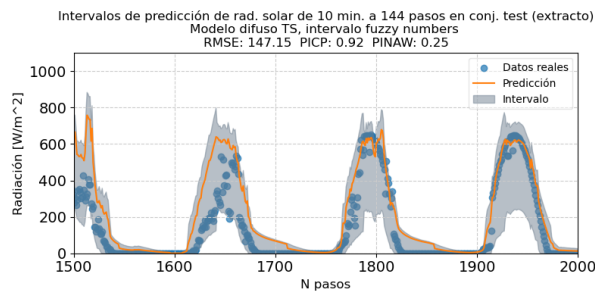
(b) PM - Quality Driven



(c) LSTM - JointSupervision



(d) LSTM - Quality Driven



(e) Fuzzy Numbers (Referencia)

Figura 6.1: Intervalos de predicción para los datos de radiación solar de 10m a 144 horas (PM: Perceptrón multicapa).

De la figura 6.1 se observa que todos siguen la dinámica de manera correcta, aunque el LSTM-QD tiene un comportamiento levemente ruidoso. A pesar de esto es el tiene menor ancho.

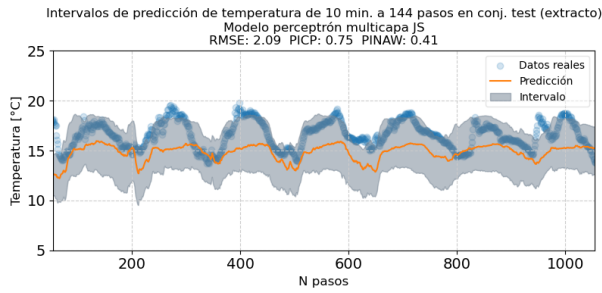
En la siguiente tabla se muestra los errores de predicción, porcentaje de cobertura del intervalo (PICP) y el ancho promedio normalizado del intervalo (PINAW) de los distintos modelos de predicción de intervalos. Para el caso de la radiación solar de 10 minutos evaluados a 12 horas (72 pasos), a 24 horas (144 pasos) y 48 horas (288 pasos), evaluados en el conjunto test:

Tabla 6.2: Comparación entre los distintos modelos de intervalos en los datos de radiación solar a 10 minutos, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno).

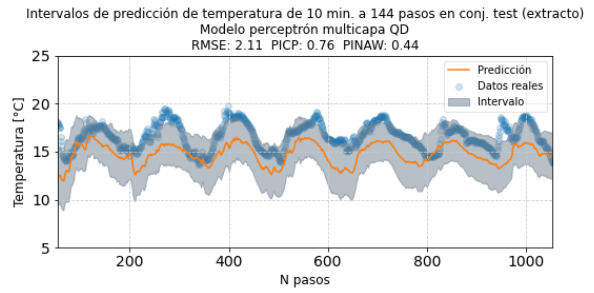
Horizonte de Predicción	Métricas	Radiación solar a 10m				
		PM-JS	PM-QD	LSTM-JS	LSTM-QD	FN-TS
a 12 horas (72 Pasos)	RMSE [W/m <sup>2</sup> ]	332.79	239.68	260.603	197.039	105.682
	PICP (%)	53.55	78.47	52.32	88.62	90.55
	PINAW (%)	52.60	64.59	93.82	47.26	21.59
a 24 horas (144 Pasos)	RMSE [W/m <sup>2</sup> ]	111.81	99.49	198.8459	108.822	146.0425
	PICP (%)	94.64	94.64	80.96	91.34	91.96
	PINAW (%)	32.15	25.72	44.30	24.27	24.77
a 48 horas (288 Pasos)	RMSE [W/m <sup>2</sup> ]	157.56	118.63	122.2208	119.5134	247.62
	PICP (%)	94.78	88.14	88.20	90.88	89.76
	PINAW (%)	41.39	21.95	28.07	25.61	45.54

De la tabla 6.2 se aprecia que el modelo que obtiene un PICP relativamente alto manteniendo un PINAW bajo es el LSTM-QD. El modelo difuso también tiene un desempeño pero en las predicciones a 48 horas tiene un ancho casi del doble que el LSTM-QD.

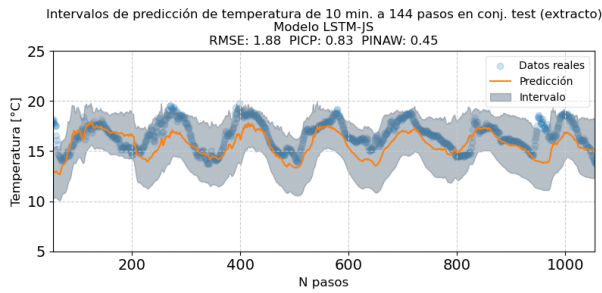
A continuación se muestra la comparación de la predicción de intervalos de los distintos modelos para el caso de la temperatura de 10 minutos a 24 horas (144 pasos), evaluados en el conjunto test:



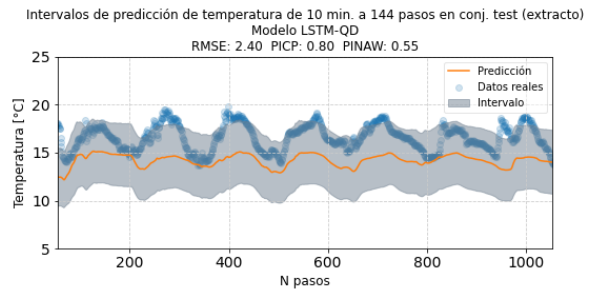
(a) PM - JointSupervision



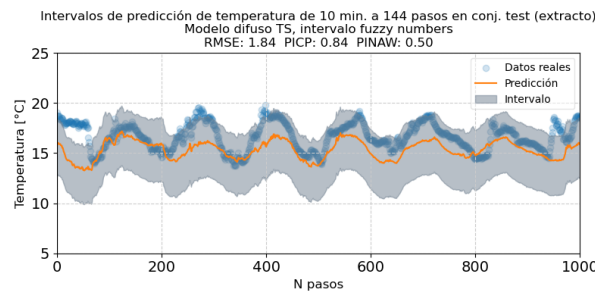
(b) PM - Quality Driven



(c) LSTM - JointSupervision



(d) LSTM - Quality Driven



(e) Fuzzy Numbers (Referencia)

Figura 6.2: Intervalos de predicción para los datos de temperatura de 10m a 144 horas (PM: Perceptrón multicapa).

De la figura 6.2 se observa que en general los intervalos están desplazados hacia abajo con respecto a los datos reales. Esto se debe a que los datos de entrenamiento pertenecen a los meses de primavera mientras que los de test a los de verano, por lo que presentan dinámicas distintas. Este problema proviene de la falta de datos de temperatura de 10 minutos.

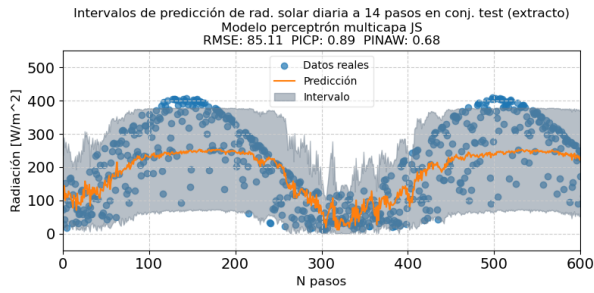
En la siguiente tabla se muestra los errores de predicción, porcentaje de cobertura del intervalo (PICP) y el ancho promedio normalizado del intervalo (PINAW) de los distintos modelos de predicción de intervalos. Para el caso de la temperatura de 10 minutos evaluados a 12 horas (72 pasos), a 24 horas (144 pasos) y 48 horas (288 pasos), evaluados en el conjunto test:

Tabla 6.3: Comparación entre los distintos modelos de intervalos en los datos de temperatura a 10 minutos, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno).

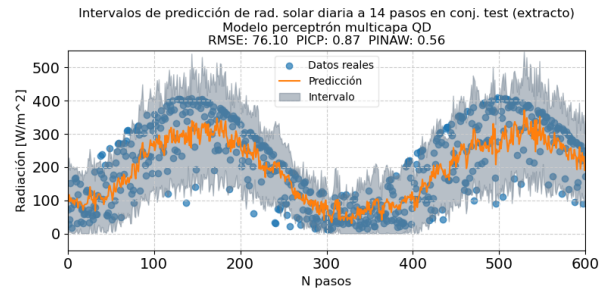
Horizonte de Predicción	Métricas	Temperatura a 10m				
		Tipo modelo	PM-JS	PM-QD	LSTM-JS	LSTM-QD
a 12 horas (72 Pasos)	RMSE [°C]	2.75	2.27	2.66	2.32	1.95
	PICP (%)	61.65	78.54	51.05	69.06	82.94
	PINAW (%)	44.74	48.95	38.73	46.61	50.45
a 24 horas (144 Pasos)	RMSE [°C]	2.09	2.11	1.8753	2.41	1.84
	PICP (%)	74.83	75.83	83.18	79.16	83.72
	PINAW (%)	41.06	43.52	44.89	54.49	49.87
a 48 horas (288 Pasos)	RMSE [°C]	2.10	1.89	2.15	2.31	2.06
	PICP (%)	83.26	88.61	75.31	75.38	85.20
	PINAW (%)	48.57	51.13	43.23	47.35	53.04

De la tabla 6.3 se tiene que el modelo con es el modelo difuso, obteniendo un PICP superior a 80 en todos los pasos, teniendo el mejor desempeño en error y PICP a 12 horas y 24 horas, mientras que el PM-QD tiene el mejor desempeño a 48 horas. Se elige el modelo difuso como el mejor. Todos los modelos obtienen un PINAW grande, además de que la predicción está desplazada hacia arriba. Esto se puede deber a que los datos de entrenamiento son de primavera y los datos de test pertenecen a datos de verano. Faltan entonces, más datos para aprender mejor la dinámica completa.

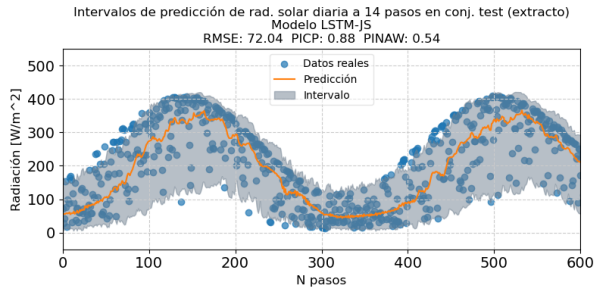
A continuación se muestra la comparación de la predicción de intervalos de los distintos modelos para el caso de la radiación solar diaria a 14 días (14 pasos), evaluados en el conjunto test:



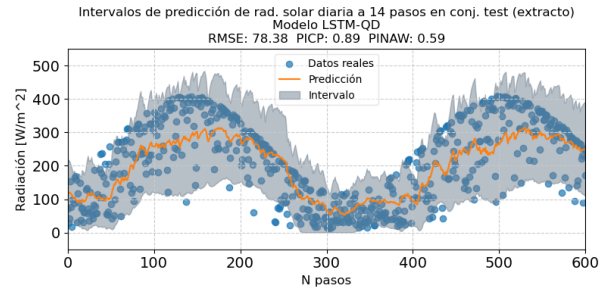
(a) PM - JointSupervision



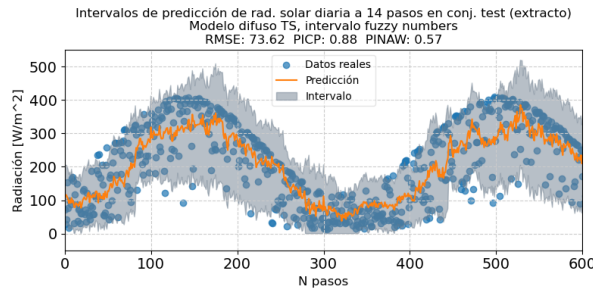
(b) PM - Quality Driven



(c) LSTM - JointSupervision



(d) LSTM - Quality Driven



(e) Fuzzy Numbers (Referencia)

Figura 6.3: Intervalos de predicción para los datos de radiación solar diaria a 14 días (PM: Perceptrón multicapa).

De la figura 6.3 se observa que el intervalo más suave que sigue de manera correcta la dinámica de los datos es el LSTM-JS, mientras que los demás tienen un ancho mayor y más variable. A pesar de esto, todos engloban a los datos de buena manera.

En la siguiente tabla se muestra los errores de predicción, porcentaje de cobertura del intervalo (PICP) y el ancho promedio normalizado del intervalo (PINAW) de los distintos modelos de predicción de intervalos. Para el caso de la radiación solar diaria evaluados a 1 día (1 paso), a 14 días (14 pasos) y 28 días (288 pasos), evaluados en el conjunto test:



Tabla 6.4: Comparación entre los distintos modelos de intervalos en los datos de radiación solar diaria, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno).

Horizonte de Predicción	Métricas	Radiación solar diaria					
		Tipo modelo	PM-JS	PM-QD	LSTM-JS	LSTM-QD	FN-TS
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]		70.66	70.97	66.38	70.84	69.00
	PICP (%)		87.74	85.04	90.29	89.76	86.87
	PINAW (%)		53.76	49.42	48.41	52.51	50.99
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]		85.11	76.10	70.45	76.69	72.40
	PICP (%)		88.81	86.99	89.08	88.54	87.99
	PINAW (%)		67.99	56.31	52.13	57.66	55.58
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]		83.03	78.49	69.17	71.17	76.48
	PICP (%)		89.10	87.35	90.10	86.19	88.68
	PINAW (%)		68.79	59.32	53.69	50.74	60.99

De la tabla 6.4 se aprecia que el ancho del intervalo es grande en todos los modelos. Esto se debe a que los datos tienen una variabilidad muy grande para días contiguos. La razón de esta variabilidad, como se explicó anteriormente, posiblemente se deba a la nubosidad existente. También se observa que los resultados son mucho más estables a distintos pasos variando poco. Lo que indica más confiabilidad en su uso. El modelo con mejor PICP y PINAW fue el LSTM-JS, aunque el PM-JS también obtuvo valores cercanos, pero un PINAW levemente mayor.

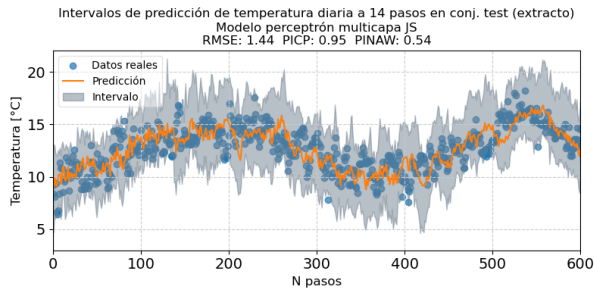
En la siguiente tabla se muestra los errores de predicción, porcentaje de cobertura del intervalo (PICP) y el ancho promedio normalizado del intervalo (PINAW) de los distintos modelos de predicción de intervalos. Para el caso de la temperatura diaria evaluados a 1 día (1 paso), a 14 días (14 pasos) y 28 días (288 pasos), evaluados en el conjunto test:

Tabla 6.5: Comparación entre los distintos modelos de intervalos en los datos de temperatura diaria, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno).

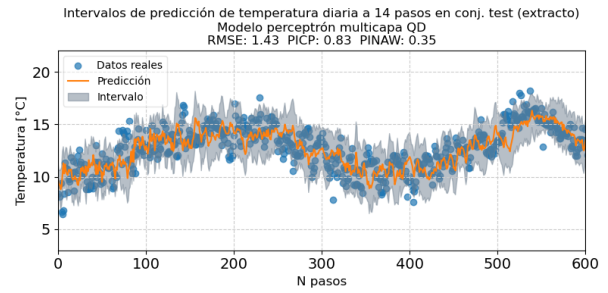
Horizonte de Predicción	Métricas	Temperatura diaria				
		Tipo modelo	PM-JS	PM-QD	LSTM-JS	LSTM-QD
a 1 día (1 Paso)	RMSE [°C]	1.06	1.06	1.09	1.09	1.03
	PICP (%)	95.84	86.50	93.88	85.34	95.42
	PINAW (%)	43.28	26.77	39.34	28.07	35.11
a 14 días (14 Pasos)	RMSE [°C]	1.44	1.43	1.51	1.51	1.41
	PICP (%)	95.09	83.45	92.30	82.89	93.62
	PINAW (%)	54.28	35.38	49.31	33.54	44.57
a 28 días (28 Pasos)	RMSE [°C]	1.48	1.44	1.62	1.62	1.49
	PICP (%)	93.50	87.47	94.79	82.84	94.29
	PINAW (%)	56.14	38.95	56.53	33.72	48.48

De la tabla 6.5 se aprecia que los modelos tienen un comportamiento estable para distintos número de pasos, lo que los hace más confiables. Se tiene que el modelo difuso obtiene el error mas bajo y un PICP alto. Mientras que modelo PM-QD tiene un PICP cercano al 90 %, lo que se considera bueno, pero además obteniendo un un ancho de cobertura 10 % menor. Por lo que se decide utilizar seleccionar este método como el mejor. Ya que finalmente lo que será usado a futuro sera los intervalos, por lo que prima que sea angosto.

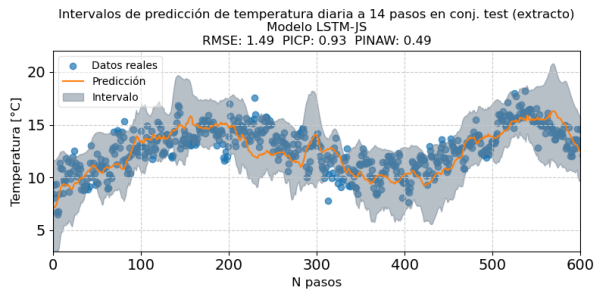
A continuación se muestra la comparación de la predicción de intervalos de los distintos modelos para el caso de la temperatura diaria a 14 días (14 pasos), evaluados en el conjunto test:



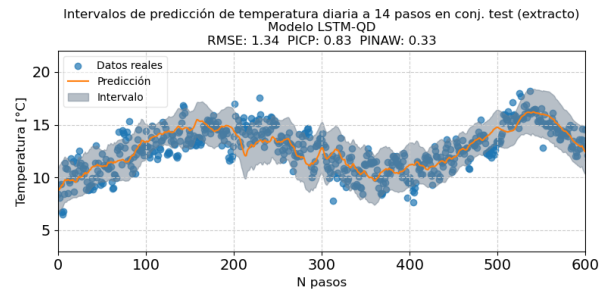
(a) PM - JointSupervision



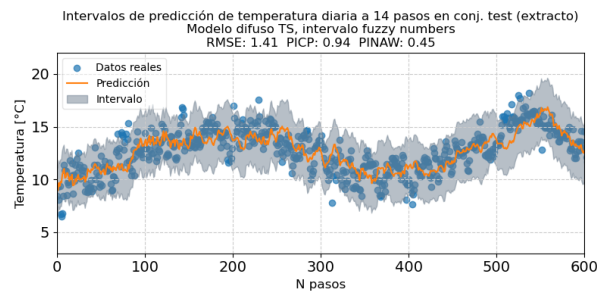
(b) PM - Quality Driven



(c) LSTM - JointSupervision



(d) LSTM - Quality Driven



(e) Fuzzy Numbers (Referencia)

Figura 6.4: Intervalos de predicción para los datos de temperatura diaria a 14 días (PM: Perceptrón multicapa).

De la figura 6.4 se observa que los intervalos siguen de buena manera a la dinámica, aunque algunos tienen un comportamiento muy ruidoso.

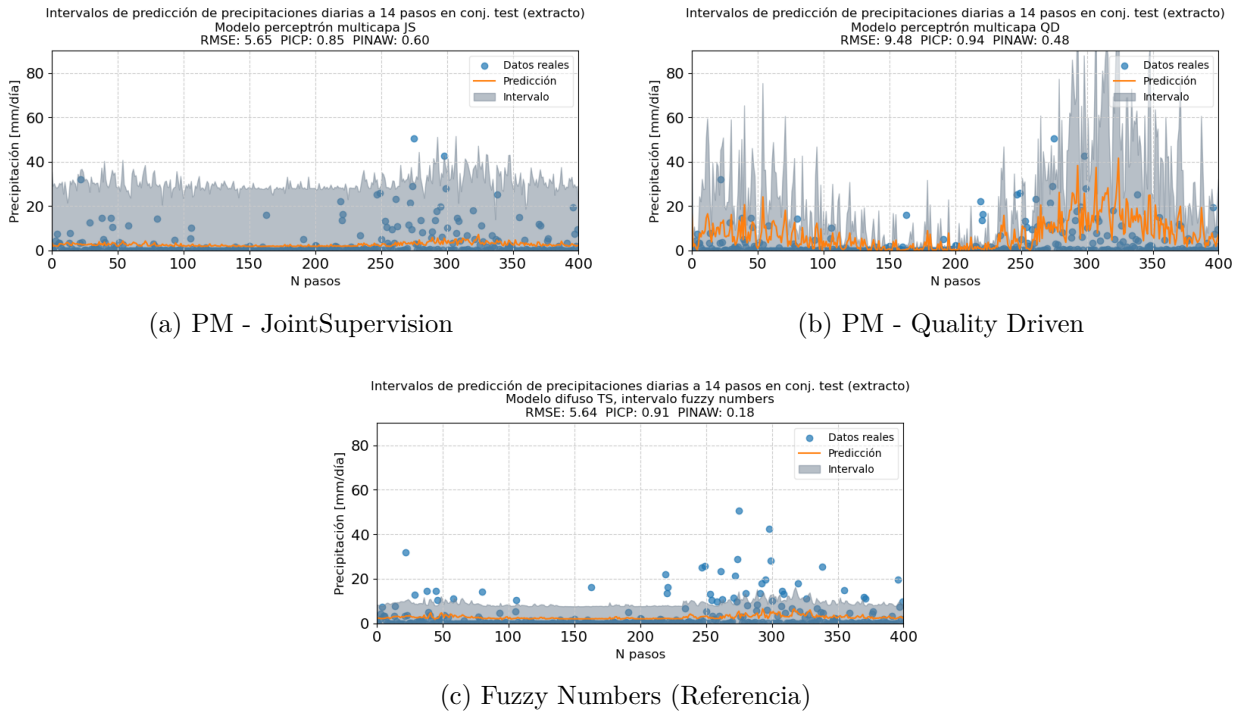
En la siguiente tabla se muestra los errores de predicción, porcentaje de cobertura del intervalo (PICP) y el ancho promedio normalizado del intervalo (PINAW) de los distintos modelos de predicción de intervalos. Para el caso de la precipitación diaria evaluados a 1 día (1 paso), a 14 días (14 pasos) y 28 días (288 pasos), evaluados en el conjunto test:

Tabla 6.6: Comparación entre los distintos modelos de intervalos en los datos de precipitación diaria, evaluados en su conjunto de test (PM: perceptrón multicapa, JS: Joint Supervision, QD:QualityDriven, FN: Fuzzy Numbers, TS: Takagi Sugeno).

Horizonte de Predicción	Métricas	Precipitación diaria		
		PM-JS	PM-QD	FN-TS
	<b>Tipo modelo</b>			
a 1 día (1 Paso)	RMSE [mm/día]	5.38	5.62	5.41
	PICP (%)	86.77	92.33	89.98
	PINAW (%)	59.33	17.36	17.27
a 14 días (14 Pasos)	RMSE [mm/día]	5.65	9.48	5.62
	PICP (%)	84.55	93.84	90.65
	PINAW (%)	60.14	48.23	17.76
a 28 días (28 Pasos)	RMSE [mm/día]	5.72	9.70	5.65
	PICP (%)	80.69	93.61	90.85
	PINAW (%)	66.37	41.39	19.15

De la tabla 6.6 se aprecia el mal desempeño de las distintas métricas. Este mal resultado era esperable debido a la naturaleza caótica de las predicciones y de que se intenta predecir a varios pasos. Y además de que en experimentos de intervalos, hay una correlación entre el resultado de predicción con los resultados de los intervalos. Como se obtuvieron malas predicciones en la sección anterior, era esperable que los intervalos también fuesen deficientes.

A continuación se muestra la comparación de la predicción de intervalos de los distintos modelos para el caso de la precipitación diaria a 14 días (14 pasos), evaluados en el conjunto test:



(a) PM - JointSupervision

(b) PM - Quality Driven

(c) Fuzzy Numbers (Referencia)

Figura 6.5: Intervalos de predicción para los datos de precipitación diaria a 14 días (PM: Perceptrón multicapa).

De las figuras en 6.5, se observa que todos los intervalos tienen un ancho muy grande, en donde JS y *Fuzzy Numbers* tiene un ancho casi constante, mientras que QD muy variable, con muchos falsos positivos. Existe un aumento en el ancho en las épocas de invierno. Pero esto no es aporte significativo de información ya que es sabido que en invierno es más probable que llueva.

Otro punto relevante, es porqué los modelos basados en la red LSTM dan mejores resultados que para la predicción del capítulo anterior. Primero que todo hay que recordar que para los modelos de predicción de intervalos, a diferencia de los modelos de predicción, se entrenan modelos distintos para cada paso distinto. Entonces no existe una retroalimentación de la salida sobre la entrada para predecir pasos futuros como en el caso de los modelos de predicción, quitando ese error.

### 6.3. Discusión

En este capítulo se lleva a cabo el entrenamiento de los modelos de intervalos de predicción de *Quality Driven* basados en el perceptrón multicapa y LSTM, así como los modelos de intervalos de predicción de *Joint Supervision* basados en el perceptrón multicapa y LSTM. La arquitectura utilizada en los modelos base se ha construido utilizando los hiperparámetros óptimos encontrados en el capítulo anterior. Además, se compara el desempeño de estos modelos con un modelo de intervalos de predicción difuso *Fuzzy Numbers*, basado en el modelo de predicción difuso *Takagi Sugeno* tipo 2.

Para llevar a cabo la comparación, se crea un modelo distinto para cada nuevo paso, excepto en el caso del modelo difuso. En los datos registrados cada 10 minutos, se realizan predicciones para un paso de 12, 24 y 48 horas. Para los datos diarios, las predicciones se

hacen a 1 día, 14 días y 28 días.

La comparación se realiza calculando el error de predicción utilizando las métricas de RMSE, el porcentaje de cobertura PICP y el ancho de intervalo PINAW.

Para los datos de radiación solar registrados cada 10 minutos, se obtiene que el modelo LSTM-QD entrega los mejores resultados debido a su bajo error, alto PICP y bajo PINAW en comparación con los demás modelos.

En el caso de la temperatura registrada cada 10 minutos, el modelo difuso obtiene el mejor resultado a 12 y 24 horas, mientras que el modelo PM-QD lo hace para la predicción a 48 horas. Por lo tanto, se inclina a utilizar el modelo difuso.

Para la radiación solar diaria, se selecciona como mejor modelo el LSTM-JS debido a que su error es bajo y el PICP se encuentra en un rango del 90 %, además de tener un menor PINAW en comparación con los demás modelos.

En el caso de la temperatura diaria, se tiene que el modelo difuso obtiene el error más bajo y un alto PICP. Por otro lado, el modelo PM-QD tiene un PICP cercano al 90 %, lo que se considera bueno, pero también obtiene un ancho de cobertura un 10 % menor. Por lo tanto, se decide utilizar este último modelo, ya que lo que se utilizará en el futuro son los intervalos y es importante que sean angostos.

En el caso de las precipitaciones diarias, es difícil establecer cuál tiene el mejor desempeño. Por un lado, el modelo PM-JS tiene un ancho de banda muy grande y casi constante a lo largo del tiempo, elevándose levemente en invierno. Del mismo modo, el modelo difuso también tiene este comportamiento con un ancho constante, pero con la diferencia de que es más angosto. En el caso de PM-QD, se tiene una variabilidad muy grande que no lo hace confiable.

En el caso de la precipitación, es la única variable en la cual los modelos no aportan información significativa acerca de la incertidumbre, por lo tanto, se debería plantear mejor su uso como cuantificador de incertidumbre.

# Capítulo 7

## Conclusiones y trabajo futuro

En este capítulo se describe las conclusiones y el trabajo futuro de esta memoria.

### 7.1. Conclusiones

En esta memoria se logró cumplir el objetivo de analizar los datos de la estación meteorológica de la comunidad y se encontraron diversas fuentes de datos, validándolas y comparándolas con los datos obtenidos de la estación de la comunidad. Esto amplió el número de datos disponibles para entrenar modelos, lo que servirá como base para entrenamientos futuros. También se logró diseñar e implementar algoritmos de predicción basados en redes neuronales, en donde se implementaron los modelos de perceptrón multicapa y LSTM, así como intervalos de predicción basados en modelos de redes neuronales. En particular, se implementaron cuatro modelos de intervalos de predicción: modelos de *Joint Supervision* basados en perceptrón multicapa, modelos de *Quality driven* basados en perceptrón multicapa, modelos de *Joint Supervision* basados en redes LSTM y modelos de *Quality driven* basados en redes LSTM. Se validó el desempeño de los distintos algoritmos mediante los resultados de un modelo difuso entrenado en los mismos conjuntos de datos.

Se concluye que las predicciones obtenidas son aceptables en la mayoría de los casos, excepto para las precipitaciones diarias, lo que sugiere que podrían ser utilizadas en el sistema de control agua-energía que se quiere implementar en la comunidad mapuche. Sin embargo, se necesita una mayor investigación para obtener valores más precisos para las precipitaciones diarias.

Con respecto a los datos, se concluye que es importante adquirirlos, pero no es suficiente para utilizarlos directamente. En el caso de los datos de radiación solar, temperatura y velocidad del viento, fue necesario compararlos con los datos registrados en la comunidad para seleccionar los más adecuados. En el caso de la precipitación, como no había registros en la comunidad, se necesitó la opinión de expertos para evaluar la fuente de datos mediante distintos análisis.

En cuanto a la predicción de intervalos, se obtuvieron resultados estables y con un porcentaje de cobertura relativamente alto en los casos de radiación diaria, temperatura diaria y radiación solar diaria. Sin embargo, su uso dependerá de la forma en que se utilice la

información proporcionada. La elección del modelo de predicción debe ser estudiada cuidadosamente teniendo en cuenta el objetivo de su uso y el número de datos disponibles para el entrenamiento del modelo.

En algunos resultados, se obtuvieron predicciones muy similares entre el perceptrón multicapa y el modelo difuso *Takagi Sugeno*, lo que se esperaba dado que ambos son aproximadores universales. Se encontró un *trade off* entre la temporalidad de los datos y la estocasticidad, ya que a mayor escala de tiempo, la dinámica es menos estocástica, pero se dispone de menos datos, lo que dificulta el aprendizaje del modelo. Mientras que, a menor escala temporal, existe una mayor cantidad de datos, pero la estocasticidad aumenta, lo que hace que el modelo tenga que aprender de una señal más compleja.

Por último, se destaca que el método JS obtuvo mejores resultados en comparación con el QD en algunos conjuntos de datos, pero el tiempo de entrenamiento del primero es mucho mayor ya que debe entrenar tantos modelos como  $\lambda$  se prueben, mientras que en QD solo se debe entrenar un modelo.

## 7.2. Trabajo Futuro

Para mejorar las predicciones de las distintas variables climatológicas, es recomendable considerar la utilización de modelos con entradas multivariadas que contemplen en la entrada más variables además de la que se quiere predecir. En particular, podrían ser otras variables climatológicas, lo que tiene sentido debido a la interacción que existe entre ellas y a que ya se cuenta con datos.

Además, se podría estudiar la posibilidad de utilizar servicios externos de pago que entreguen información de variables climatológicas adicionales. En el caso de la radiación solar, se podría considerar incorporar información de la hora del día de los datos, ya que estos tienen una marcada estacionalidad. Para esta misma variable, también se podría considerar la nubosidad a futuro, sacada de datos satelitales. Como se mencionó en el análisis, se observa que hay días en donde la radiación es muy baja comparada con los días contiguos, y se le atribuye este comportamiento principalmente a la nubosidad. Cabe destacar que en la literatura se indica que la predicción de nubosidad es fiable hasta 2 días, lo que la hace útil en el caso de la predicción a corto plazo, pero no tanto para el mediano plazo.

También se podría explorar la utilización de un modelo de red recurrente distinto a la LSTM, que tenga un número de parámetros más acorde al número de datos disponibles.

En relación a la temperatura diaria, se observa que los valores muy altos que se dan en algunos días de verano y los valores muy bajos que se dan en algunos días de invierno no se están captando bien en el modelo. Por lo tanto, en lugar de utilizar un único modelo, se podría emplear dos modelos distintos, uno para el período de primavera-verano y otro para el período de otoño-invierno, con el fin de mejorar las predicciones.

En resumen, para mejorar las predicciones de las variables climatológicas, se sugiere la utilización de modelos con entradas multivariadas, la consideración de información adicional



de servicios externos, la exploración de modelos de red recurrente distintos a la LSTM y la utilización de modelos separados para diferentes períodos de tiempo.

# Bibliografía

- [1] Boisier, J. P., Alvarez-Garretón, C., Cordero, R. R., Damiani, A., Gallardo, L., Garreaud, R. D., Lambert, F., Ramallo, C., Rojas, M., y Rondanelli, R., “Anthropogenic drying in central-southern Chile evidenced by long-term observations and climate model simulations,” *Elementa: Science of the Anthropocene*, vol. 6, 2018, [doi:10.1525/elementa.328](https://doi.org/10.1525/elementa.328). 74.
- [2] Rinaudo, J.-D. y Donoso, G., “State, market or community failure? untangling the determinants of groundwater depletion in copiapó (chile),” *International Journal of Water Resources Development*, vol. 35, no. 2, pp. 283–304, 2019, [doi:10.1080/07900627.2017.1417116](https://doi.org/10.1080/07900627.2017.1417116).
- [3] Garreaud, R. D., Boisier, J. P., Rondanelli, R., Montecinos, A., Sepúlveda, H. H., y Veloso-Aguila, D., “The central chile mega drought (2010–2018): A climate dynamics perspective,” *International Journal of Climatology*, vol. 40, no. 1, pp. 421–439, 2020, [doi:https://doi.org/10.1002/joc.6219](https://doi.org/10.1002/joc.6219).
- [4] Yu, Y., Si, X., Hu, C., y Zhang, J., “A review of recurrent neural networks: Lstm cells and network architectures,” *Neural Comput.*, vol. 31, p. 1235–1270, 2019, [doi:10.1162/neco\\_a\\_01199](https://doi.org/10.1162/neco_a_01199).
- [5] Rosenblatt, F., “The perceptron: a probabilistic model for information storage and organization in the brain.,” *Psychological review*, vol. 65, pp. 386–408, 1958.
- [6] Agahian, S. y Akan, T., “Battle royale optimizer for training multi-layer perceptron,” *Evolving Systems*, vol. 13, p. 563–575, [doi:10.1007/s12530-021-09401-5](https://doi.org/10.1007/s12530-021-09401-5).
- [7] Manning, T., Sleator, R., y Walsh, P., “Biologically inspired intelligent decision making,” *Bioengineered*, vol. 5, 2013, [doi:10.4161/bioe.26997](https://doi.org/10.4161/bioe.26997).
- [8] Werbos, “Backpropagation and neurocontrol: a review and prospectus,” en *International 1989 Joint Conference on Neural Networks*, pp. 209–216 vol.1, 1989, [doi:10.1109/IJCNN.1989.118583](https://doi.org/10.1109/IJCNN.1989.118583).
- [9] Werbos, P., “Backpropagation through time: what it does and how to do it,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990, [doi:10.1109/5.58337](https://doi.org/10.1109/5.58337).
- [10] Kabir, H. M. D., Khosravi, A., Hosen, M. A., y Nahavandi, S., “Neural network-based uncertainty quantification: A survey of methodologies and applications,” *IEEE Access*, vol. 6, pp. 36218–36234, 2018, [doi:10.1109/ACCESS.2018.2836917](https://doi.org/10.1109/ACCESS.2018.2836917).
- [11] Cartagena, O., Parra, S., Muñoz-Carpintero, D., Marín, L. G., y Sáez, D., “Review on fuzzy and neural prediction interval modelling for nonlinear dynamical systems,” *IEEE Access*, vol. 9, pp. 23357–23384, 2021, [doi:10.1109/ACCESS.2021.3056003](https://doi.org/10.1109/ACCESS.2021.3056003).
- [12] Khosravi, A., Nahavandi, S., Creighton, D., y Atiya, A. F., “Comprehensive review of

- neural network-based prediction intervals and new advances,” *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1341–1356, 2011, [doi:10.1109/TNN.2011.2162110](https://doi.org/10.1109/TNN.2011.2162110).
- [13] Khosravi, A., Nahavandi, S., Creighton, D., y Atiya, A. F., “Lower upper bound estimation method for construction of neural network-based prediction intervals,” *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 337–346, 2011, [doi:10.1109/TNN.2010.2096824](https://doi.org/10.1109/TNN.2010.2096824).
- [14] Hwang, J. T. G. y Ding, A. A., “Prediction intervals for artificial neural networks,” *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 748–757, 1997, [doi:10.1080/01621459.1997.10474027](https://doi.org/10.1080/01621459.1997.10474027).
- [15] Ungar, L., De, R., y Rosengarten, V., “Estimating prediction intervals for artificial neural networks,” 2000.
- [16] Denker, J. S. y LeCun, Y., “Transforming neural-net output levels to probability distributions,” en *Proceedings of the 3rd International Conference on Neural Information Processing Systems, NIPS’90*, (San Francisco, CA, USA), p. 853–859, Morgan Kaufmann Publishers Inc., 1990.
- [17] Torres, A. F., Walker, W. R., y McKee, M., “Forecasting daily potential evapotranspiration using machine learning and limited climatic data,” *Agricultural Water Management*, vol. 98, pp. 553–562, 2011, <https://ideas.repec.org/a/eee/agiwat/v98y2011i4p553-562.html>.
- [18] Nix, D. y Weigend, A., “Estimating the mean and variance of the target probability distribution,” en *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN’94)*, vol. 1, pp. 55–60 vol.1, 1994, [doi:10.1109/ICNN.1994.374138](https://doi.org/10.1109/ICNN.1994.374138).
- [19] Paass, G., “Assessing and improving neural network predictions by the bootstrap algorithm,” en *Advances in Neural Information Processing Systems (Hanson, S., Cowan, J., y Giles, C., eds.)*, vol. 5, Morgan-Kaufmann, 1992, <https://proceedings.neurips.cc/paper/1992/file/0ff8033cf9437c213ee13937b1c4c455-Paper.pdf>.
- [20] Zhou, Z.-H., Wu, J., y Tang, W., “Ensembling neural networks: Many could be better than all,” *Artificial Intelligence*, vol. 137, no. 1, pp. 239–263, 2002, [doi:https://doi.org/10.1016/S0004-3702\(02\)00190-X](https://doi.org/10.1016/S0004-3702(02)00190-X).
- [21] Marín, L. G., Sumner, M., Muñoz-Carpintero, D., Köbrich, D., Pholboon, S., Sáez, D., y Núñez, A., “Hierarchical energy management system for microgrid operation based on robust model predictive control,” *Energies*, vol. 12, no. 23, p. 4453, 2019.
- [22] Sáez, D., Ávila, F., Olivares, D., Cañizares, C., y Marín, L., “Fuzzy prediction interval models for forecasting renewable resources and loads in microgrids,” *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 548–556, 2015, [doi:10.1109/TSG.2014.2377178](https://doi.org/10.1109/TSG.2014.2377178).
- [23] Quan, H., Srinivasan, D., y Khosravi, A., “Short-term load and wind power forecasting using neural network-based prediction intervals,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 2, pp. 303–315, 2014, [doi:10.1109/TNNLS.2013.2276053](https://doi.org/10.1109/TNNLS.2013.2276053).
- [24] ak, R., Li, Y.-F., y Zio, E., “Estimation of prediction intervals of neural network models by a multi-objective genetic algorithm,” vol. 7, 2012, [doi:10.1142/9789814417747\\_0166](https://doi.org/10.1142/9789814417747_0166).
- [25] Pearce, T., Zaki, M., Brintrup, A., y Neely, A., “High-quality prediction intervals for deep learning: A distribution-free, ensembled approach,” 2018, [doi:10.48550/ARXIV.1808.08862](https://doi.org/10.48550/ARXIV.1808.08862).

- [26] Cruz, N., Marin, L. G., y Saez, D., “Neural network prediction interval based on joint supervision,” en 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–8, IEEE, 2018.
- [27] Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., y Nahavandi, S., “A review of uncertainty quantification in deep learning: Techniques, applications and challenges,” *Information Fusion*, vol. 76, pp. 243–297, 2021, doi:<https://doi.org/10.1016/j.inffus.2021.05.008>.
- [28] Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., y Nahavandi, S., “A review of uncertainty quantification in deep learning: Techniques, applications and challenges,” *Information Fusion*, vol. 76, pp. 243–297, 2021, doi:<https://doi.org/10.1016/j.inffus.2021.05.008>.
- [29] Ho, S., Xie, M., Tang, L., Xu, K., y Goh, T., “Neural network modeling with confidence bounds: a case study on the solder paste deposition process,” *IEEE Transactions on Electronics Packaging Manufacturing*, vol. 24, no. 4, pp. 323–332, 2001, doi:[10.1109/6104.980042](https://doi.org/10.1109/6104.980042).
- [30] Lu, T. y Viljanen, M., “Prediction of indoor temperature and relative humidity using neural network models: Model comparison,” *Neural Computing and Applications*, vol. 18, pp. 345–357, 2009, doi:[10.1007/s00521-008-0185-3](https://doi.org/10.1007/s00521-008-0185-3).
- [31] Kuo, Y. M., Liu, C. W., y Lin, K. H., “Evaluation of the ability of an artificial neural network model to assess the variation of groundwater quality in an area of blackfoot disease in taiwan,” *Water Res.*, vol. 38, no. 1, pp. 148–158, 2004.
- [32] Khosravi, A., Mazloumi, E., Nahavandi, S., Creighton, D., y van Lint, J. W. C., “Prediction intervals to account for uncertainties in travel time prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 537–547, 2011, doi:[10.1109/TITS.2011.2106209](https://doi.org/10.1109/TITS.2011.2106209).
- [33] Khosravi, A., Nahavandi, S., y Creighton, D., “Load forecasting and neural networks: A prediction interval-based perspective,” en *Computational Intelligence in Power Engineering (Studies in Computational Intelligence*, vol. 302, pp. 131–150,, Berlin, Germany: Springer-Verlag, doi:[10.1007/978-3-642-14013-6\\_5](https://doi.org/10.1007/978-3-642-14013-6_5).
- [34] Hwang, J.-N. y Little, E., “Real time recurrent neural networks for time series prediction and confidence estimation,” en *Proceedings of International Conference on Neural Networks (ICNN’96)*, vol. 4, pp. 1889–1894 vol.4, 1996, doi:[10.1109/ICNN.1996.549189](https://doi.org/10.1109/ICNN.1996.549189).
- [35] Rigamonti, M., Baraldi, P., Zio, E., Roychoudhury, I., Goebel, K., y Poll, S., “Ensemble of optimized echo state networks for remaining useful life prediction,” *Neurocomputing*, vol. 281, pp. 121–138, 2018, doi:<https://doi.org/10.1016/j.neucom.2017.11.062>.
- [36] Yao, W., Zeng, Z., y Lian, C., “Generating probabilistic predictions using mean-variance estimation and echo state network,” *Neurocomputing*, vol. 219, pp. 536–547, 2017, doi:<https://doi.org/10.1016/j.neucom.2016.09.064>.
- [37] Khosravi, A. y Nahavandi, S., “An optimized mean variance estimation method for uncertainty quantification of wind power forecasts,” *International Journal of Electrical*

- Power Energy Systems, vol. 61, p. 446–454, 2014, [doi:10.1016/j.ijepes.2014.03.060](https://doi.org/10.1016/j.ijepes.2014.03.060).
- [38] Pinson, P. y Kariniotakis, G., “On-line assessment of prediction risk for wind power production forecasts,” *Wind Energy*, vol. 7, no. 2, pp. 119–132, 2004, [doi:https://doi.org/10.1002/we.114](https://doi.org/10.1002/we.114).
- [39] Talebizadeh, M. y Moridnejad, A., “Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ann and anfis models,” *Expert Systems with Applications*, vol. 38, no. 4, pp. 4126–4135, 2011, [doi:https://doi.org/10.1016/j.eswa.2010.09.075](https://doi.org/10.1016/j.eswa.2010.09.075).
- [40] Mazloumi, E., Rose, G., Currie, G., y Moridpour, S., “Prediction intervals to account for uncertainties in neural network predictions: Methodology and application in bus travel time prediction,” *Engineering Applications of Artificial Intelligence*, vol. 24, no. 3, pp. 534–542, 2011, [doi:https://doi.org/10.1016/j.engappai.2010.11.004](https://doi.org/10.1016/j.engappai.2010.11.004).
- [41] Zio, E., “A study of the bootstrap method for estimating the accuracy of artificial neural networks in predicting nuclear transient processes,” *IEEE Transactions on Nuclear Science*, vol. 53, no. 3, pp. 1460–1478, 2006, [doi:10.1109/TNS.2006.871662](https://doi.org/10.1109/TNS.2006.871662).
- [42] Baraldi, P., Compare, M., Saucó, S., y Zio, E., “Ensemble neural network-based particle filtering for prognostics,” *Mechanical Systems and Signal Processing*, vol. 41, no. 1, pp. 288–300, 2013, [doi:https://doi.org/10.1016/j.ymsp.2013.07.010](https://doi.org/10.1016/j.ymsp.2013.07.010).
- [43] Conduit, B., Jones, N., Stone, H., y Conduit, G., “Design of a nickel-base superalloy using a neural network,” *Materials Design*, vol. 131, pp. 358–365, 2017, [doi:https://doi.org/10.1016/j.matdes.2017.06.007](https://doi.org/10.1016/j.matdes.2017.06.007).
- [44] Ristovski, K., Vucetic, S., y Obradovic, Z., “Uncertainty analysis of neural-network-based aerosol retrieval,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 409–414, 2012, [doi:10.1109/TGRS.2011.2166120](https://doi.org/10.1109/TGRS.2011.2166120).
- [45] Khosravi, A., Nahavandi, S., y Creighton, D., “Quantifying uncertainties of neural network-based electricity price forecasts,” *Applied Energy*, vol. 112, pp. 120–129, 2013, [doi:https://doi.org/10.1016/j.apenergy.2013.05.075](https://doi.org/10.1016/j.apenergy.2013.05.075).
- [46] Khosravi, A., Nahavandi, S., Srinivasan, D., y Khosravi, R., “Constructing optimal prediction intervals by using neural networks and bootstrap method,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 8, pp. 1810–1815, 2015, [doi:10.1109/TNNLS.2014.2354418](https://doi.org/10.1109/TNNLS.2014.2354418).
- [47] Marín, L. G., Valencia, F., y Sáez, D., “Prediction interval based on type-2 fuzzy systems for wind power generation and loads in microgrid control design,” en *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 328–335, 2016, [doi:10.1109/FUZZ-IEEE.2016.7737705](https://doi.org/10.1109/FUZZ-IEEE.2016.7737705).
- [48] Shi, Z., Liang, H., y Dinavahi, V., “Direct interval forecast of uncertain wind power based on recurrent neural networks,” *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1177–1187, 2018, [doi:10.1109/TSTE.2017.2774195](https://doi.org/10.1109/TSTE.2017.2774195).
- [49] Li, R. y Jin, Y., “A wind speed interval prediction system based on multi-objective optimization for machine learning method,” *Applied Energy*, vol. 228, pp. 2207–2220, 2018, [doi:10.1016/j.apenergy.2018.07.032](https://doi.org/10.1016/j.apenergy.2018.07.032).
- [50] ak, R., Li, Y.-F., Vitelli, V., y Zio, E., “Multi-objective genetic algorithm optimization of a neural network for estimating wind speed prediction intervals,” *rep. tec.*, 2013.

- [51] Galván, I. M., Valls, J. M., Cervantes, A., y Aler, R., “Multi-objective evolutionary optimization of prediction intervals for solar energy forecasting with neural networks,” *Information Sciences*, vol. 418-419, pp. 363–382, 2017, doi:<https://doi.org/10.1016/j.ins.2017.08.039>.
- [52] Ni, Q., Zhuang, S., Sheng, H., Kang, G., y Xiao, J., “An ensemble prediction intervals approach for short-term pv power forecasting,” *Solar Energy*, vol. 155, pp. 1072–1083, 2017, doi:<https://doi.org/10.1016/j.solener.2017.07.052>.
- [53] Quan, H., Srinivasan, D., y Khosravi, A., “Uncertainty handling using neural network-based prediction intervals for electrical load forecasting,” *Energy*, vol. 73, pp. 916–925, 2014, doi:<https://doi.org/10.1016/j.energy.2014.06.104>.
- [54] Wang, J., Gao, Y., y Chen, X., “A novel hybrid interval prediction approach based on modified lower upper bound estimation in combination with multi-objective salp swarm algorithm for short-term load forecasting,” *Energies*, vol. 11, no. 6, p. 1561, 2018.
- [55] Rana, M., Koprinska, I., Khosravi, A., y Agelidis, V. G., “Prediction intervals for electricity load forecasting using neural networks,” en *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2013, doi:[10.1109/IJCNN.2013.6706839](https://doi.org/10.1109/IJCNN.2013.6706839).
- [56] Hosen, M. A., Khosravi, A., Nahavandi, S., y Creighton, D., “Improving the quality of prediction intervals through optimal aggregation,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, pp. 4420–4429, 2015, doi:[10.1109/TIE.2014.2383994](https://doi.org/10.1109/TIE.2014.2383994).
- [57] Hosen, M. A., Khosravi, A., Nahavandi, S., y Creighton, D., “Prediction interval-based neural network modelling of polystyrene polymerization reactor – a new perspective of data-based modelling,” *Chemical Engineering Research and Design*, vol. 92, no. 11, pp. 2041–2051, 2014, doi:<https://doi.org/10.1016/j.cherd.2014.02.016>.
- [58] Kavousi-Fard, A., Khosravi, A., y Nahavandi, S., “Reactive power compensation in electric arc furnaces using prediction intervals,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 7, pp. 5295–5304, 2017, doi:[10.1109/TIE.2017.2677345](https://doi.org/10.1109/TIE.2017.2677345).
- [59] Quan, H., Srinivasan, D., y Khosravi, A., “Particle swarm optimization for construction of neural network-based prediction intervals,” *Neurocomputing*, vol. 127, pp. 172–180, 2014, doi:<https://doi.org/10.1016/j.neucom.2013.08.020>. *Advances in Intelligent Systems*.
- [60] Cruz, N., Marin, L., y Saez, D., “Prediction intervals with lstm networks trained by joint supervision,” en *2019 International Joint Conference on Neural Networks (IJCNN)*, doi:[10.1109/ijcnn.2019.8852215](https://doi.org/10.1109/ijcnn.2019.8852215).
- [61] Cruz, N., Marin, L., y Saez, D., “Neural network prediction interval based on joint supervision,” en *2018 International Joint Conference on Neural Networks (IJCNN)*, doi:[10.1109/ijcnn.2018.8489264](https://doi.org/10.1109/ijcnn.2018.8489264).
- [62] Gal, Y. y Ghahramani, Z., “Dropout as a bayesian approximation: Representing model uncertainty in deep learning,” en *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, ICML’16*, p. 1050–1059, JMLR.org, 2016.
- [63] Lakshminarayanan, B., Pritzel, A., y Blundell, C., “Simple and scalable predictive uncertainty estimation using deep ensembles,” 2016, doi:[10.48550/ARXIV.1612.01474](https://doi.org/10.48550/ARXIV.1612.01474).
- [64] Detlefsen, N. S., Jørgensen, M., y Hauberg, S., “Reliable training and estimation of

- variance networks,” 2019, [doi:10.48550/ARXIV.1906.03260](https://doi.org/10.48550/ARXIV.1906.03260).
- [65] Patel, K. y Waslander, S., “Accurate prediction and uncertainty estimation using decoupled prediction interval networks,” 2022, [doi:10.48550/ARXIV.2202.09664](https://doi.org/10.48550/ARXIV.2202.09664).
- [66] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., y Polosukhin, I., “Attention is all you need,” 2017, [doi:10.48550/ARXIV.1706.03762](https://doi.org/10.48550/ARXIV.1706.03762).
- [67] Mathonsi, T. y Zyl, T. L. v., “Prediction interval construction for multivariate point forecasts using deep learning,” en 2020 7th International Conference on Soft Computing Machine Intelligence (ISCFMI), pp. 88–95, 2020, [doi:10.1109/ISCFMI51676.2020.9311603](https://doi.org/10.1109/ISCFMI51676.2020.9311603).
- [68] Liu, S., Ji, H., y Wang, M., “Nonpooling convolutional neural network forecasting for seasonal time series with trends,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 8, pp. 2879–2888,, [doi:10.1109/TNNLS.2019.2934110](https://doi.org/10.1109/TNNLS.2019.2934110).
- [69] “Reportes fluviométricos, información oficial hidrometeorológica y de calidad de aguas en línea, ministerio de obras públicas.” <https://snia.mop.gob.cl/BNAConsultas/reportes>. Accessed: 2022-06-13.
- [70] “Reportes fluviométricos, información oficial hidrometeorológica y de calidad de aguas en línea, ministerio de obras públicas.” <https://snia.mop.gob.cl/BNAConsultas/reportes>. Accessed: 2022-06-13.
- [71] Molina, A., Falvey, M., y Rondanelli, R., “A solar radiation database for chile.,” *Scientific Reports*, vol. 7, no. 1, pp. 14823–14823, 2017.
- [72] Muñoz, R. C., Falvey, M. J., Arancibia, M., Astudillo, V. I., Elgueta, J., Ibarra, M., Santana, C., y Vásquez, C., “Wind energy exploration over the atacama desert: A numerical model–guided observational program,” *Bulletin of the American Meteorological Society*, vol. 99, no. 10, pp. 2079 – 2092, 2018, [doi:10.1175/BAMS-D-17-0019.1](https://doi.org/10.1175/BAMS-D-17-0019.1).
- [73] Muñoz Sabater, J., “Era5-land hourly data from 1981 to present. copernicus climate change service (c3s) climate data store (cds.”, [doi:10.24381/cds.e2161bac](https://doi.org/10.24381/cds.e2161bac). Accessed on < 25-04-2022 >).
- [74] Muñoz Sabater, J., “Era5-land hourly data from 1950 to 1980. copernicus climate change service (c3s) climate data store (cds.”, [doi:10.24381/cds.e2161bac](https://doi.org/10.24381/cds.e2161bac). Accessed on < 25-04-2022 >).
- [75] “Producto grillado cr2met, center for climate and resilience research (cr2).” [https://www.cr2.cl/datos-productos-grillados/?cp\\_cr2met](https://www.cr2.cl/datos-productos-grillados/?cp_cr2met). Accessed: 2022-06-13.
- [76] “Estación climatológica tranapunte.” <https://climatologia.meteochile.gob.cl/application/informacion/fichaDeEstacion/380054/>. Accessed: 2022-06-13.
- [77] “Estación climatológica quiripio.” <https://climatologia.meteochile.gob.cl/application/informacion/fichaDeEstacion/380051>. Accessed: 2022-06-13.
- [78] Endo, A., Cartagena, O., Ocaranza, J., Sáez, D., y Muñoz, C., “Fuzzy and neural prediction intervals for robust control of a greenhouse,” en 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–8, 2022, [doi:10.1109/FUZZ-IEEE55066.2022.9882701](https://doi.org/10.1109/FUZZ-IEEE55066.2022.9882701).
- [79] Takagi, T. y Sugeno, M., “Fuzzy identification of systems and its applications to modeling and control,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 1, pp. 116–132, 1985, [doi:10.1109/TSMC.1985.6313399](https://doi.org/10.1109/TSMC.1985.6313399).

- [80] Marín, L. G., Cruz, N., Sáez, D., Sumner, M., y Núñez, A., “Prediction interval methodology based on fuzzy numbers and its extension to fuzzy systems and neural networks,” *Expert Systems with Applications*, vol. 119, pp. 128–141, 2019, doi:<https://doi.org/10.1016/j.eswa.2018.10.043>.



# Anexos

## Anexo A. Resultados de entrenamiento perceptrón multicapa

En esta sección se muestra los resultados del error que se obtiene en la optimización inicial del número de neuronas en la capa oculta del modelo de perceptrón multicapa en cada uno de los conjuntos.

En la figura A.1 se muestra la comparación del RMSE obtenido para distintos número de neuronas en la capa oculta en el conjunto de radiación solar registrada cada 10 minutos.

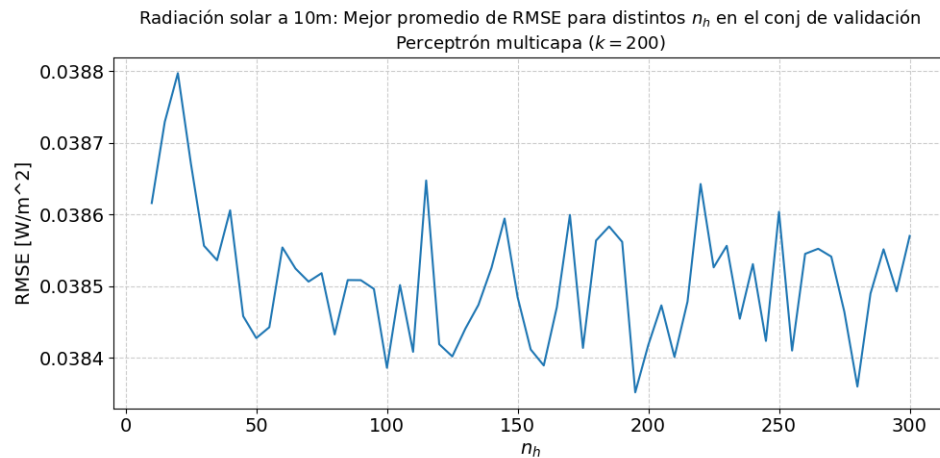


Figura A.1: Promedio de RMSE de la validación cruzada en los datos de radiación solar a 10 minutos, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores.

De la figura A.1 se observa que el mínimo error se encuentra cercano a las 200 neuronas. Pero este número es considerablemente grande. Siguiendo la tendencia que sigue el error, existe una meseta en donde el error ya no tiende a bajar, sino que para alguna arquitectura en particular puede que baje pero no para las cercanas. Esto es un síntoma de un sobreajuste a los datos. Por tanto se toma el mínimo desde el momento en que se llega a la meseta del error, que en este caso es 100.

En la figura A.2 se muestra la comparación del RMSE obtenido para distintos número de neuronas en la capa oculta en el conjunto de temperatura registrada cada 10 minutos.

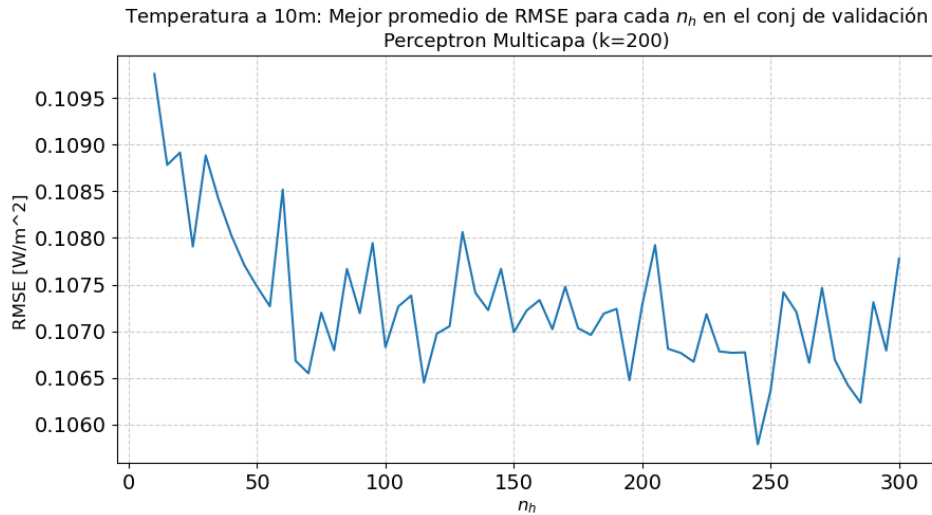


Figura A.2: Promedio de RMSE de la validación cruzada en los datos de temperatura a 10 minutos, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores.

De la figura A.2 se observa que el mínimo error se encuentra cercano a las 250 neuronas. Pero este número es considerablemente grande. Siguiendo la tendencia que sigue el error, existe una meseta en donde el error ya no tiende a bajar, sino que para alguna arquitectura en particular puede que baje pero no para las cercanas. Esto es un síntoma de un sobreajuste a los datos. Por tanto se toma el mínimo desde el momento en que se llega a la meseta del error.

En la figura A.3 se muestra la comparación del RMSE obtenido para distintos número de neuronas en la capa oculta en el conjunto de radiación diaria.

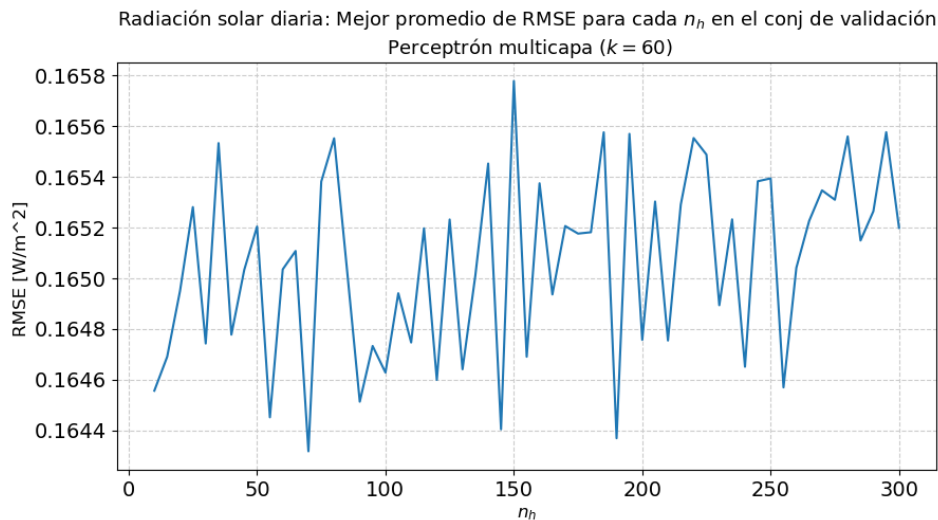


Figura A.3: Promedio de RMSE de la validación cruzada en los datos de radiación solar diaria, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores.

De la figura A.3 se observa un comportamiento distinto a los anteriores. Se obtiene un mínimo cerca de las 70 neuronas y luego sube. Esto quiere decir que se está sobreajustando.

En la figura A.4 se muestra la comparación del RMSE obtenido para distintos número de neuronas en la capa oculta en el conjunto de temperatura diaria.

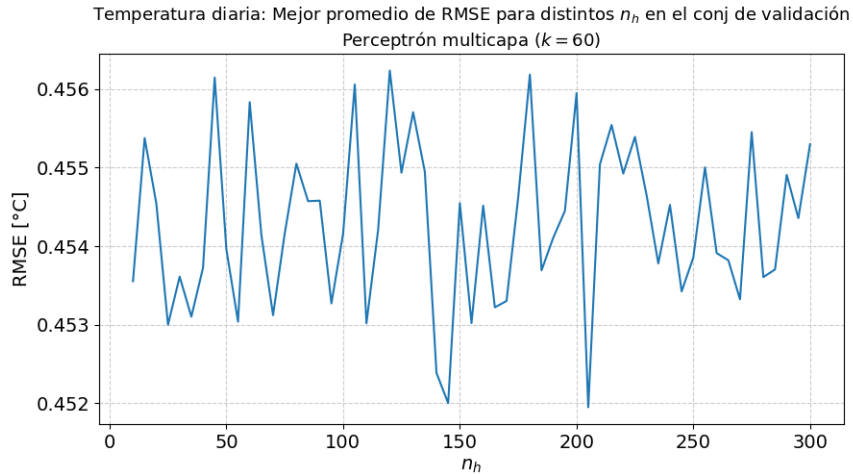


Figura A.4: Promedio de RMSE de la validación cruzada en los datos de temperatura diaria, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta y un número fijo de regresores.

De la figura A.3 se observa un comportamiento distinto a los anteriores. Se obtiene un mínimo cerca de las 70 neuronas y luego sube. Esto quiere decir que se está sobreajustando. Se elige el valor mínimo antes de que empiece a subir el error.

En la figura A.5 se muestra la comparación del RMSE obtenido para distintos número de neuronas en la capa oculta en el conjunto de precipitaciones diarias.

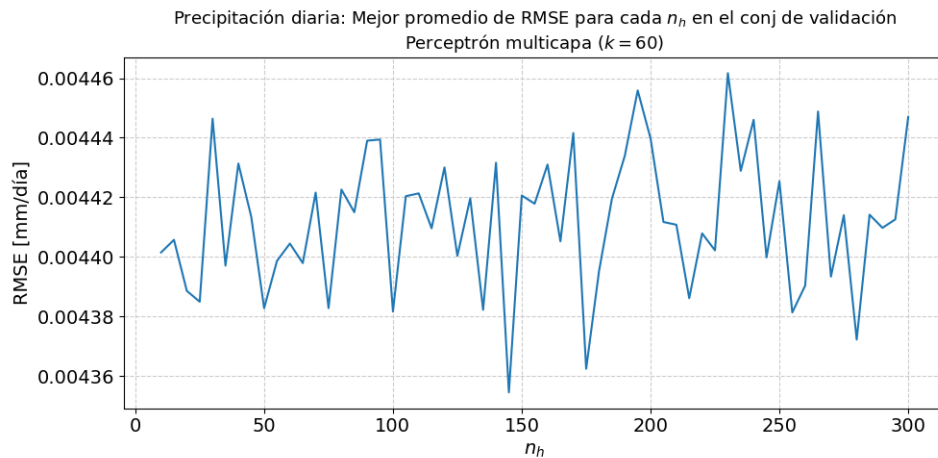


Figura A.5: Promedio de RMSE de la validación cruzada en los datos de precipitación diaria, evaluados en el conjunto de validación para distintas configuraciones con distintos número de neuronas en la capa oculta.

De la figura A.5 se observa que el mínimo error se encuentra cercano a las 150 neuronas. Después de esto empieza a aumentar su valor de error.

## Anexo B. Análisis de sensibilidad

A continuación, se muestran los resultados del análisis de sensibilidades que se le realiza al modelo inicial para cada uno de los conjunto de datos. En el caso de los datos diarios se fijan el número de regresores en 60 y en los datos cada 10 minutos se fijan 200 regresores.

En la figura B.1 se muestra los gradiente absolutos promedios de cada regresor, para la temperatura registrada cada 10 minutos.

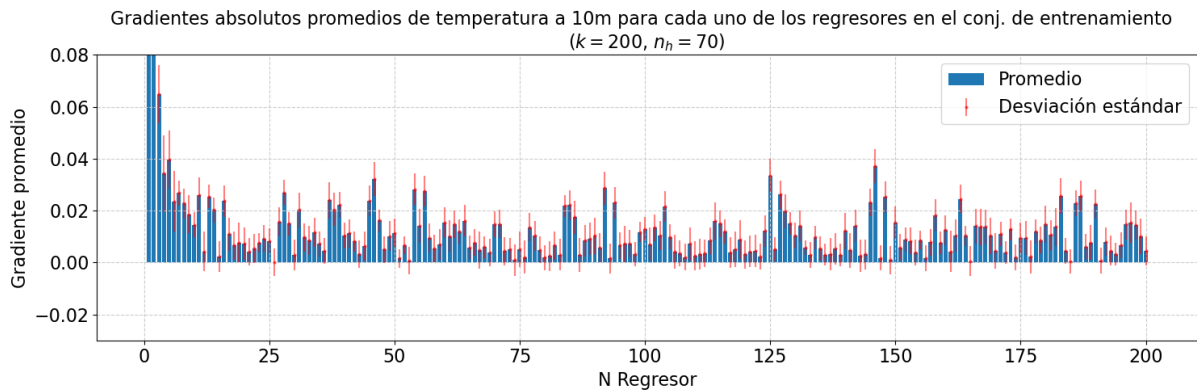


Figura B.1: Gradientes absolutos promedio de temperatura cada 10 minutos para cada uno de los regresores evaluados en el conjunto de entrenamiento.

En la figura B.2 se muestra los gradiente absolutos promedios de cada regresor, para la temperatura registrada cada 10 minutos.

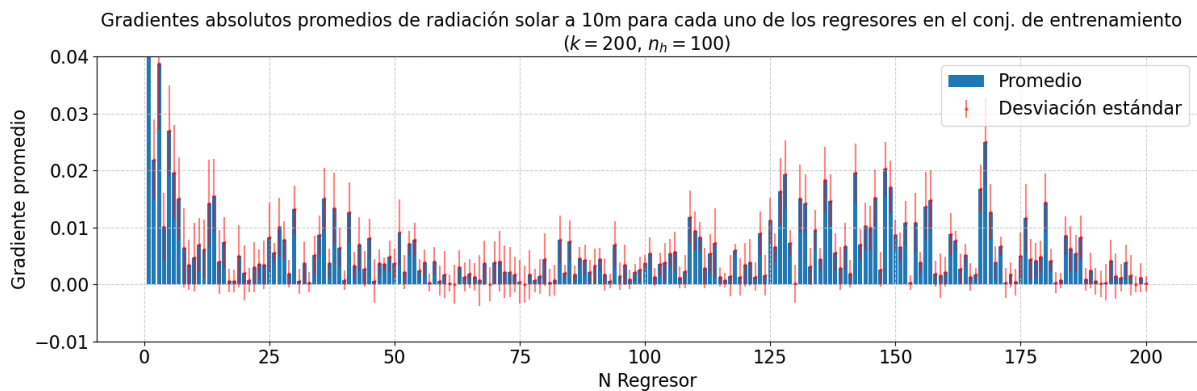


Figura B.2: Gradientes absolutos promedio de radiación solar cada 10 minutos para cada uno de los regresores evaluados en el conjunto de entrenamiento.

En la figura B.3 se muestra los gradiente absolutos promedios de cada regresor, para la temperatura diaria.

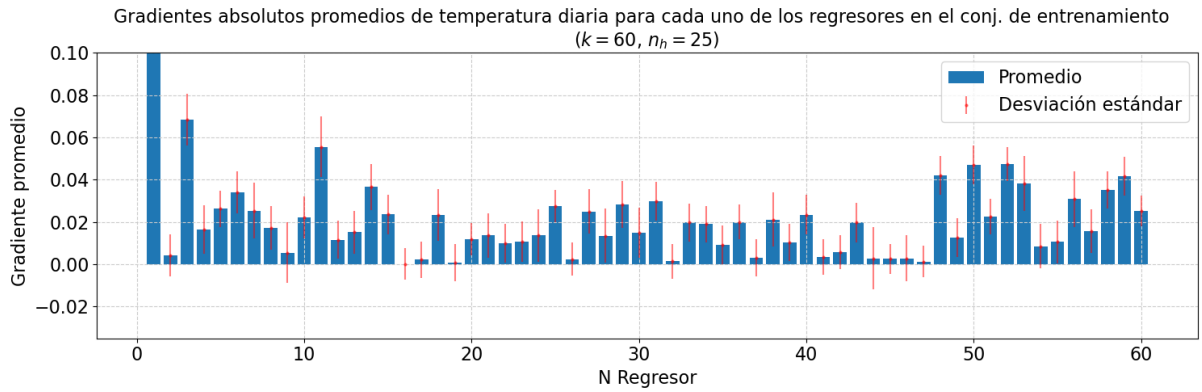


Figura B.3: Gradientes absolutos promedio de temperatura diaria para cada uno de los regresores evaluados en el conjunto de entrenamiento.

En la figura B.4 se muestra los gradiente absolutos promedios de cada regresor, para la radiación solar diaria.

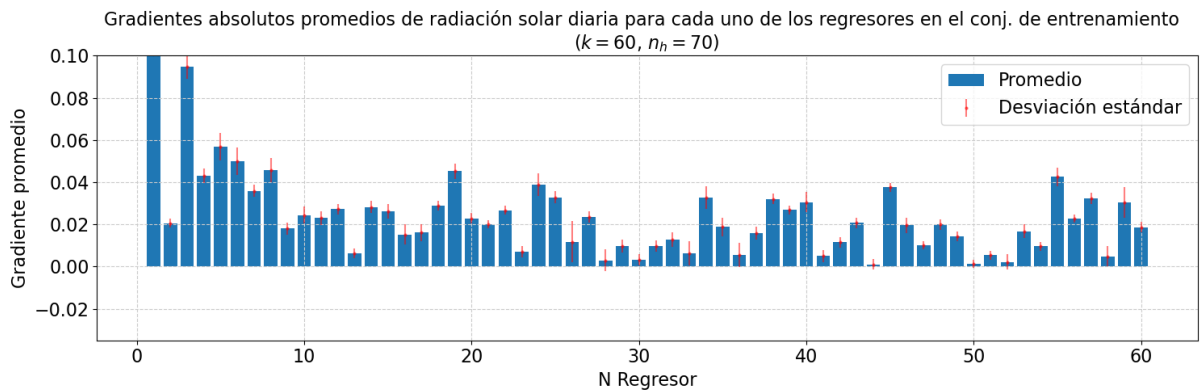


Figura B.4: Gradientes absolutos promedio de radiación solar diaria para cada uno de los regresores evaluados en el conjunto de entrenamiento.

En la figura B.5 se muestra los gradiente absolutos promedios de cada regresor, para la precipitación diaria.

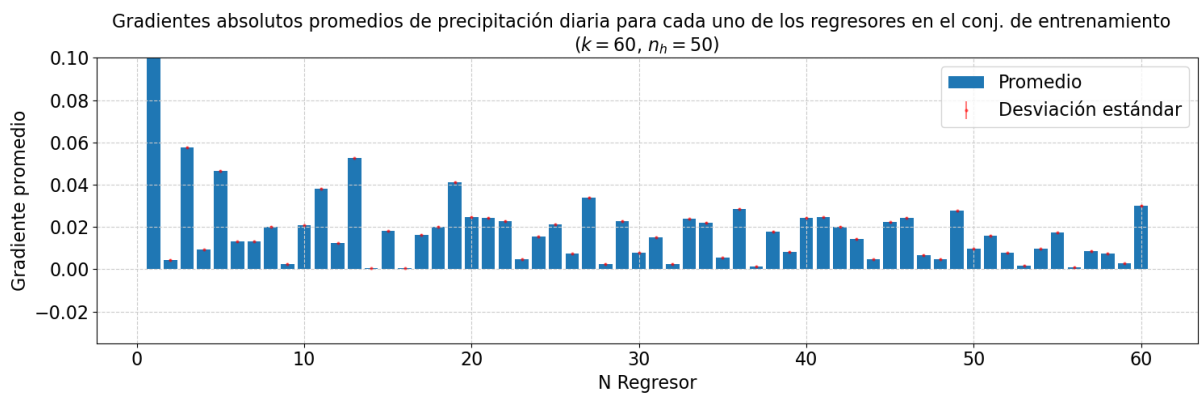


Figura B.5: Gradientes absolutos promedio de precipitación diaria para cada uno de los regresores evaluados en el conjunto de entrenamiento.

En todas los análisis de sensibilidad realizados, mostrados en las figuras B.4, B.3, B.5, se puede observar que la magnitud es muy alta para los regresores más cercanos al tiempo actual. Esto es esperable, debido a que el valor de estas variables, exceptuando las precipitaciones, por lo general no varían en extremo en el corto plazo. Esto es esperable debido que los datos más próximos están correlacionados con el dato actual, debido a la estacionalidad que representan.

Por otra parte, en los datos a corto plazo, se puede observar de las figuras B.2 y B.1 que existe una magnitud grande al rededor del regresor número 144. Esto se debe a que este regresor representa el dato registrado un día atrás, el cual tiene una fuerte correlación con el dato actual debido a una estacionalidad diaria que presentan estos datos.

## Anexo C. Resultado de intervalos

En la tabla C.1 se muestran los resultados de los modelos de intervalos de predicción basados en el perceptrón multicapa además del perceptron multicapa solo, en el conjunto de radiación solar registrada cada 10 minutos. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 12 horas, 24 horas y 48 horas, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.1: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de radiación solar a 10 minutos evaluado en los conjuntos de test .

Radiación solar registrada cada 10 minutos				
Horizonte de Predicción		Perceptron Multicapa		
Conjunto		train		
Tipo modelo		MP-QD	MP-JS	FC
a 12 horas (72 Pasos)	RMSE	219.41	249.55	145.12
	PICP (%)	85.05	64.82	-
	PINAW (%)	38.96	26.87	-
a 24 horas (144 Pasos)	RMSE	136.82	141.10	148.22
	PICP (%)	81.95	86.19	-
	PINAW (%)	15.39	21.03	-
a 48 horas (288 Pasos)	RMSE	144.26	169.02	172.29
	PICP (%)	74.25	87.16	-
	PINAW (%)	13.27	26.04	-
Conjunto		val		
Tipo modelo		MP-QD	MP-JS	FC
a 12 horas (72 Pasos)	RMSE	174.53	147.39	125.01
	PICP (%)	94.87	89.23	-
	PINAW (%)	42.97	26.48	-
a 24 horas (144 Pasos)	RMSE	117.79	125.11	131.63
	PICP (%)	84.22	90.37	-
	PINAW (%)	17.52	23.81	-
a 48 horas (288 Pasos)	RMSE	117.95	153.55	151.15
	PICP (%)	76.69	88.68	-
	PINAW (%)	14.89	29.06	-
Conjunto		test		
Tipo modelo		MP-QD	MP-JS	FC
a 12 horas (72 Pasos)	RMSE	239.68	332.80	119.53
	PICP (%)	78.47	53.56	-
	PINAW (%)	64.59	52.60	-
a 24 horas (144 Pasos)	RMSE	99.49	111.81	125.84
	PICP (%)	94.64	94.64	-
	PINAW (%)	25.72	32.16	-
a 48 horas (288 Pasos)	RMSE	118.63	157.56	180.21
	PICP (%)	88.15	94.78	-
	PINAW (%)	21.95	41.39	-
Numero parámetros de modelos		1712	1758	1666
Estadísticas dataset		# Datos	min	max
		48836	0.0025	1196.3

En la tabla C.2 se muestran los resultados de los modelos de intervalos de predicción basados en la red LSTM además de la red LSTM sola, en el conjunto de radiación solar registrada cada 10 minutos. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 12 horas, 24 horas y 48 horas, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.2: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de radiación solar a 10 minutos evaluado en los conjuntos de test .

Radiación solar registrada cada 10 minutos				
Horizonte de Predicción		LSTM		
Conjunto		train		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 12 horas (72 Pasos)	RMSE [°C]	173.65	248.16	294.53
	PICP (%)	90.55	78.69	-
	PINAW (%)	28.02	58.85	-
a 24 horas (144 Pasos)	RMSE [°C]	147.38	174.07	356.01
	PICP (%)	86.98	85.22	-
	PINAW (%)	18.34	25.80	-
a 48 horas (288 Pasos)	RMSE [°C]	157.45	152.88	295.14
	PICP (%)	83.11	85.46	-
	PINAW (%)	17.28	22.12	-
Conjunto		val		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 12 horas (72 Pasos)	RMSE [°C]	156.30	213.45	328.59
	PICP (%)	92.58	89.99	-
	PINAW (%)	29.89	60.60	-
a 24 horas (144 Pasos)	RMSE [°C]	128.76	133.36	398.69
	PICP (%)	89.95	90.58	-
	PINAW (%)	22.33	26.01	-
a 48 horas (288 Pasos)	RMSE [°C]	130.41	128.47	323.47
	PICP (%)	89.43	89.70	-
	PINAW (%)	21.92	27.09	-
Conjunto		test		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 12 horas (72 Pasos)	RMSE [°C]	197.04	260.60	175.95
	PICP (%)	88.62	52.32	-
	PINAW (%)	47.26	93.82	-
a 24 horas (144 Pasos)	RMSE [°C]	108.82	198.85	217.76
	PICP (%)	91.34	80.96	-
	PINAW (%)	24.27	44.30	-
a 48 horas (288 Pasos)	RMSE [°C]	118.43	121.26	164.45
	PICP (%)	91.05	88.31	-
	PINAW (%)	25.61	28.07	-
Numero parametros de modelos		13106	13163	13049
Estadísticas dataset		# Datos	min	max
		48836	0.0025	1196.3

En la tabla C.3 se muestran los resultados de los modelos de intervalos de predicción basados en el perceptrón multicapa además del perceptron multicapa solo, en el conjunto de



temperatura registrada cada 10 minutos. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 12 horas, 24 horas y 48 horas, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.3: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de temperatura a 10 minutos evaluado en los conjuntos de entrenamiento, validación y test.

Temperatura registrada cada 10 minutos				
Horizonte de Predicción	Métricas	Perceptron Multicapa		
Conjunto		train		
Tipo modelo		MP-QD	MP-JS	MP
a 12 horas (72 Pasos)	RMSE [°C]	1.52	1.45	2.63
	PICP (%)	92.88	87.77	-
	PINAW (%)	41.72	31.94	-
a 24 horas (144 Pasos)	RMSE [°C]	1.49	1.49	2.83
	PICP (%)	93.06	81.81	-
	PINAW (%)	41.57	28.97	-
a 48 horas (288 Pasos)	RMSE [°C]	1.54	1.59	3.18
	PICP (%)	92.96	87.73	-
	PINAW (%)	42.46	34.39	-
Conjunto		val		
Tipo modelo		MP-QD	MP-JS	MP
a 12 horas (72 Pasos)	RMSE [°C]	1.92	2.09	2.07
	PICP (%)	86.81	80.00	-
	PINAW (%)	48.38	44.12	-
a 24 horas (144 Pasos)	RMSE [°C]	1.95	1.98	1.81
	PICP (%)	81.89	79.19	-
	PINAW (%)	44.93	39.90	-
a 48 horas (288 Pasos)	RMSE [°C]	2.03	2.18	1.89
	PICP (%)	81.44	75.18	-
	PINAW (%)	48.72	43.26	-
Conjunto		test		
Tipo modelo		MP-QD	MP-JS	MP
a 12 horas (72 Pasos)	RMSE [°C]	2.27	2.75	1.77
	PICP (%)	78.54	61.65	-
	PINAW (%)	48.95	44.75	-
a 24 horas (144 Pasos)	RMSE [°C]	2.11	2.09	1.81
	PICP (%)	75.83	74.83	-
	PINAW (%)	43.52	41.06	-
a 48 horas (288 Pasos)	RMSE [°C]	1.89	2.10	1.95
	PICP (%)	88.61	83.26	-
	PINAW (%)	51.13	48.58	-
Numero parametros de modelos		1237	1333	1141
Estadísticas dataset		# Datos	min	max
		13485	6.78	21.85

En la tabla C.4 se muestran los resultados de los modelos de intervalos de predicción basados en la red LSTM además de la red LSTM sola, en el conjunto de temperatura registrada

cada 10 minutos. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 12 horas, 24 horas y 48 horas, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.4: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de temperatura a 10 minutos evaluado en los conjuntos de entrenamiento, validación y test.

Temperatura registrada cada 10 minutos				
Horizonte de Predicción	Métricas	LSTM		
Conjunto		train		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 12 horas (72 Pasos)	RMSE [°C]	1.36	1.44	2.11
	PICP (%)	92.89	80.63	-
	PINAW (%)	37.41	26.72	-
a 24 horas (144 Pasos)	RMSE [°C]	1.49	1.54	2.18
	PICP (%)	94.51	88.83	-
	PINAW (%)	41.43	34.80	-
a 48 horas (288 Pasos)	RMSE [°C]	1.61	1.59	2.12
	PICP (%)	89.71	84.17	-
	PINAW (%)	39.92	31.28	-
Conjunto		val		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 12 horas (72 Pasos)	RMSE [°C]	2.04	2.22	2.74
	PICP (%)	80.83	74.21	-
	PINAW (%)	47.16	38.17	-
a 24 horas (144 Pasos)	RMSE [°C]	2.19	1.82	2.50
	PICP (%)	83.19	83.86	-
	PINAW (%)	52.88	44.26	-
a 48 horas (288 Pasos)	RMSE [°C]	2.11	2.07	2.52
	PICP (%)	78.49	75.38	-
	PINAW (%)	46.66	42.00	-
Conjunto		test		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 12 horas (72 Pasos)	RMSE [°C]	2.32	2.70	3.03
	PICP (%)	69.06	51.09	-
	PINAW (%)	46.61	38.73	-
a 24 horas (144 Pasos)	RMSE [°C]	2.41	1.91	2.58
	PICP (%)	79.16	81.94	-
	PINAW (%)	54.49	44.90	-
a 48 horas (288 Pasos)	RMSE [°C]	2.31	2.23	2.57
	PICP (%)	75.38	73.72	-
	PINAW (%)	47.35	43.23	-
Numero parametros de modelos		64766	64893	64639
Estadísticas dataset		# Datos	min	max
		13485	6.78	21.85

En la tabla C.5 se muestran los resultados de los modelos de intervalos de predicción basados en el perceptrón multicapa además del perceptron multicapa solo, en el conjunto de radiación solar diaria. Se usan las métricas de RMSE, PICP y PINAW para las predicciones

a 1 día, 14 días y 28 días, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.5: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de radiación solar diaria evaluado en los conjuntos de entrenamiento, validación y test.

Radiación solar diaria				
Horizonte de Predicción		Perceptron Multicapa		
Conjunto		train		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]	68.71	68.37	66.60
	PICP (%)	87.73	89.96	-
	PINAW (%)	48.17	53.64	-
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]	75.60	80.73	73.85
	PICP (%)	87.93	91.50	-
	PINAW (%)	54.19	67.86	-
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]	79.00	81.20	78.90
	PICP (%)	87.32	90.98	-
	PINAW (%)	57.27	68.99	-
Conjunto		val		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]	71.94	71.41	69.30
	PICP (%)	84.49	90.34	-
	PINAW (%)	48.26	54.09	-
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]	82.25	86.56	78.35
	PICP (%)	80.50	90.99	-
	PINAW (%)	55.23	69.29	-
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]	84.88	86.66	83.78
	PICP (%)	81.58	90.73	-
	PINAW (%)	57.85	70.29	-
Conjunto		test		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]	70.97	70.66	68.13
	PICP (%)	85.04	87.74	-
	PINAW (%)	49.42	53.76	-
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]	76.10	85.11	73.63
	PICP (%)	86.99	88.81	-
	PINAW (%)	56.31	67.99	-
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]	78.49	83.03	78.81
	PICP (%)	87.35	89.10	-
	PINAW (%)	59.32	68.79	-
Numero parametros de modelos		1542	1598	1486
Estadísticas dataset		# Datos	min	max
		4744	11.35	415.03

En la tabla C.6 se muestran los resultados de los modelos de intervalos de predicción basados en la red LSTM además de la red LSTM sola, en el conjunto de radiación solar diaria. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 1 día, 14 días y 28 días, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.6: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de radiación solar diaria evaluado en los conjuntos de entrenamiento, validación y test.

Radiación solar diaria				
Horizonte de Predicción		LSTM		
Conjunto		train		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]	70.71	66.79	114.05
	PICP (%)	90.48	91.22	-
	PINAW (%)	53.90	48.91	-
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]	76.89	71.68	196.73
	PICP (%)	89.73	92.73	-
	PINAW (%)	59.00	52.98	-
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]	78.34	69.12	178.14
	PICP (%)	85.85	94.32	-
	PINAW (%)	52.93	54.32	-
Conjunto		val		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]	72.62	69.10	109.21
	PICP (%)	89.35	90.19	-
	PINAW (%)	54.63	49.91	-
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]	80.64	74.56	183.95
	PICP (%)	89.30	89.63	-
	PINAW (%)	60.70	54.06	-
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]	80.84	75.59	168.69
	PICP (%)	82.74	89.25	-
	PINAW (%)	52.90	54.92	-
Conjunto		test		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 1 día (1 Paso)	RMSE [W/m <sup>2</sup> ]	70.84	66.38	117.94
	PICP (%)	89.76	90.29	-
	PINAW (%)	52.51	48.41	-
a 14 días (14 Pasos)	RMSE [W/m <sup>2</sup> ]	76.69	70.45	199.85
	PICP (%)	88.54	89.08	-
	PINAW (%)	57.66	52.13	-
a 28 días (28 Pasos)	RMSE [W/m <sup>2</sup> ]	71.17	69.17	179.18
	PICP (%)	86.20	90.00	-
	PINAW (%)	50.74	53.69	-
Numero parametros de modelos		18086	18153	18019
Estadísticas dataset		# Datos	min	max
		4744	11.35	415.03

En la tabla C.7 se muestran los resultados de los modelos de intervalos de predicción basados en el perceptrón multicapa además del perceptron multicapa solo, en el conjunto de temperatura diaria. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 1 día, 14 días y 28 días, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.7: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de temperatura diaria evaluado en los conjuntos de entrenamiento, validación y test.

Temperatura diria				
Horizonte de Predicción		Perceptron Multicapa		
Conjunto		train		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [°C]	1.33	1.25	1.25
	PICP (%)	81.61	98.78	-
	PINAW (%)	13.81	24.56	-
a 14 días (14 Pasos)	RMSE	1.71	1.58	1.68
	PICP (%)	80.48	98.02	-
	PINAW (%)	18.29	29.13	-
a 28 días (28 Pasos)	RMSE	1.80	1.64	1.92
	PICP (%)	81.28	97.39	-
	PINAW (%)	19.23	28.81	-
Conjunto		val		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [°C]	1.48	1.44	1.38
	PICP (%)	66.74	90.79	-
	PINAW (%)	17.82	32.37	-
a 14 días (14 Pasos)	RMSE [°C]	1.86	1.84	1.75
	PICP (%)	70.01	89.85	
	PINAW (%)	23.46	38.27	
a 28 días (28 Pasos)	RMSE	1.81	1.76	1.83
	PICP (%)	73.70	89.92	
	PINAW (%)	24.83	37.62	
Conjunto		test		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [°C]	1.06	1.06	1.01
	PICP (%)	86.50	95.84	
	PINAW (%)	26.77	43.28	
a 14 días (14 Pasos)	RMSE [°C]	1.43	1.44	1.47
	PICP (%)	83.45	95.09	
	PINAW (%)	35.38	54.28	
a 28 días (28 Pasos)	RMSE [°C]	1.44	1.48	1.66
	PICP (%)	87.47	93.50	
	PINAW (%)	38.95	56.14	
Numero parametros de modelos		1442	1488	1396
Estadísticas dataset		# Datos	min	max
		4748	2.3	23.9

En la tabla C.8 se muestran los resultados de los modelos de intervalos de predicción basados en la red LSTM además de la red LSTM sola, en el conjunto de temperatura diaria. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 1 día, 14 días y 28 días, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.8: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo LSTM, entrenados en el conjunto de temperatura diaria evaluado en los conjuntos de entrenamiento, validación y test.

Temperatura diria				
Horizonte de Predicción		LSTM		
Conjunto		train		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 1 día (1 Paso)	RMSE [°C]	1.30	1.25	1.76
	PICP (%)	86.90	96.00	-
	PINAW (%)	16.80	24.16	-
a 14 días (14 Pasos)	RMSE [°C]	1.61	1.49	4.03
	PICP (%)	84.30	96.75	-
	PINAW (%)	19.06	28.24	-
a 28 días (28 Pasos)	RMSE [°C]	1.63	1.72	4.42
	PICP (%)	83.80	94.68	-
	PINAW (%)	19.46	31.29	-
Conjunto		val		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 1 día (1 Paso)	RMSE [°C]	1.42	1.44	1.90
	PICP (%)	81.56	90.73	-
	PINAW (%)	22.99	30.96	-
a 14 días (14 Pasos)	RMSE [°C]	1.68	1.78	4.18
	PICP (%)	78.31	88.78	-
	PINAW (%)	25.13	36.09	-
a 28 días (28 Pasos)	RMSE [°C]	1.87	1.99	4.43
	PICP (%)	73.64	88.83	-
	PINAW (%)	25.38	39.55	-
Conjunto		test		
Tipo modelo		LSTM-QD	LSTM-JS	LSTM
a 1 día (1 Paso)	RMSE [°C]	1.07	1.09	1.52
	PICP (%)	85.34	93.88	-
	PINAW (%)	28.07	39.34	-
a 14 días (14 Pasos)	RMSE [°C]	1.38	1.51	4.11
	PICP (%)	82.89	92.30	-
	PINAW (%)	33.54	49.32	-
a 28 días (28 Pasos)	RMSE [°C]	1.47	1.62	4.43
	PICP (%)	82.84	94.79	-
	PINAW (%)	33.72	56.53	-
Numero parametros de modelos		7136	7178	7094
Estadísticas dataset		# Datos	min	max
		4748	2.3	23.9

En la tabla C.9 se muestran los resultados de los modelos de intervalos de predicción basados en el perceptrón multicapa además del perceptron multicapa solo, en el conjunto de precipitaciones diarias. Se usan las métricas de RMSE, PICP y PINAW para las predicciones a 1 día, 14 días y 28 días, evaluados en los conjunto de entrenamiento, validación y test.

Tabla C.9: Resultados de los modelos de predicción y modelos de intervalos neuronales basados en el modelo de perceptrón multicapa, entrenados en el conjunto de precipitación diaria evaluado en los conjuntos de entrenamiento, validación y test.

Precipitación diaria				
Horizonte de Predicción	Métricas	Perceptrón Multicapa		
Conjunto		train		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [mm/día]	6.44	6.22	6.29
	PICP (%)	91.72	76.47	-
	PINAW (%)	9.58	31.63	-
a 14 días (14 Pasos)	RMSE [mm/día]	10.58	6.55	6.57
	PICP (%)	94.32	79.06	-
	PINAW (%)	27.08	31.69	-
a 28 días (28 Pasos)	RMSE [mm/día]	10.81	6.58	6.58
	PICP (%)	93.69	72.37	-
	PINAW (%)	23.15	35.04	-
Conjunto		val		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [mm/día]	7.08	6.83	6.59
	PICP (%)	91.11	90.86	-
	PINAW (%)	15.49	48.95	-
a 14 días (14 Pasos)	RMSE [mm/día]	11.02	7.04	6.96
	PICP (%)	94.73	90.71	-
	PINAW (%)	43.58	48.09	-
a 28 días (28 Pasos)	RMSE [mm/día]	10.85	7.08	7.06
	PICP (%)	94.64	88.76	-
	PINAW (%)	37.00	53.30	-
Conjunto		test		
Tipo modelo		MP-QD	MP-JS	MP
a 1 día (1 Paso)	RMSE [mm/día]	5.62	5.38	5.46
	PICP (%)	92.34	86.77	-
	PINAW (%)	17.36	59.32	-
a 14 días (14 Pasos)	RMSE [mm/día]	9.48	5.65	5.67
	PICP (%)	93.84	84.54	-
	PINAW (%)	48.24	60.13	-
a 28 días (28 Pasos)	RMSE [mm/día]	9.70	5.72	5.77
	PICP (%)	93.61	80.69	-
	PINAW (%)	41.38	66.36	-
Numero de parámetros de modelos		197	213	108
Estadísticas dataset		<b># Datos</b>	<b>min</b>	<b>max</b>
		4349	2.70	96.80

Se puede observar de todas las tablas que existe una tendencia a que a mayor horizonte temporal, mayor el error. Esto se debe a que a medida que las variables climatológicas son de carácter caótico, por lo que mientras menor distancia temporal se tenga, se tendrá mayor información del futuro próximo.

En el caso del modelo de predicción solo, se tiene un mayor error que en los casos de intervalos de predicción. Esto se debe que el modelo de predicción solo predice de manera

futura mediante la recurrencia, es decir usar las predicciones como entrada para luego predecir el siguiente paso. Mientras que se implementa un modelo distinto para cada paso requerido en el caso de intervalos de predicción.