



Facultad de Ingeniería  
Escuela de Ingeniería Informática

## **ESTIMACIÓN DE POTENCIAL DE RECOLECCIÓN DE AGUA DE ROCÍO MEDIANTE MODELOS Y APRENDIZAJE AUTOMÁTICO**

Por

**Rodrigo Montenegro**

Trabajo realizado para optar al Título de  
**INGENIERO EN INFORMÁTICA**  
Prof. Guía: Jean-Gabriel Minonzio  
Noviembre 2023

# Resumen

Este trabajo de título investigativo aborda el problema de la estimación de la cantidad de agua de acuerdo al fenómeno del rocío, usando dos enfoques para predecir la cantidad de agua recolectada. El primer enfoque corresponde a un modelo semi-analítico desarrollado por Daniel Beysens, y el segundo enfoque, es un modelo basado en una Red Neuronal Artificial (RNA), también denominada MultiLayer Perceptron.

El objetivo general de este trabajo es comprobar cuál enfoque tiene mayor desempeño en la predicción de la cantidad de agua recolectada por día y acumulado, y así poder generar una solución de apoyo en los lugares que tengan la problemática de escasez hídrica y la falta de acceso al agua, tal como ocurre en diversas comunas de Chile, por ejemplo en Petorca, Copiapo, Vallenar y la Serena. Los datos utilizados corresponden a datos climáticos de la zona de Combarbalá, los cuales fueron recolectados durante un período de dos años, en el contexto del proyecto Fondecyt de Iniciación N.º 11140863 (2014).

En la etapa de experimentación, se probaron dos entrenamientos en función a arquitectura de las RNA, la primera corresponde a Levenger-Marquart y la segunda a Adamax. Los experimentos desarrollados utilizaron los datos de la región de Coquimbo, indicando que es posible predecir de forma correcta la medición experimental usando el modelo de Beysens y de las RNA. Para concluir, fue posible implementar el modelo de D. Beysens y una red neuronal artificial para predecir la salida correspondiente al flujo de recolección de agua de rocío de datos promedio diario (PD) y 6 AM. Se ejecutó el experimento diez veces con cada enfoque, de modo de obtener diez conjuntos de datos de entrenamiento de pruebas aleatorias, al igual que los de validación. El análisis de correlación se hizo por día y acumulado, en donde finalizaremos con la tabla del resumen de resultados, descubriendo el modelo idóneo de recolección de rocío de la zona del Norte de Chile.

# Índice general

<b>Resumen</b>	<b>III</b>
<b>1. Introducción</b>	<b>4</b>
<b>2. Marco Conceptual y Estado del Arte</b>	<b>10</b>
2.1. Marco Conceptual . . . . .	10
2.1.1. Definición de términos Básicos conceptuales: . . . . .	11
2.1.2. Variables de estudios . . . . .	19
2.2. Estado del Arte . . . . .	22
<b>3. Definición del Problema</b>	<b>29</b>
3.1. Formulación del Problema . . . . .	29
3.1.1. Hipótesis . . . . .	30
3.1.2. Preguntas de Hipótesis . . . . .	30
3.2. Solución Propuesta . . . . .	30
3.2.1. Importancia e Impacto de la solución . . . . .	31
3.3. Objetivos . . . . .	32
3.3.1. Objetivo General . . . . .	32
3.3.2. Objetivos Específicos . . . . .	32
3.4. Metodología . . . . .	32
3.4.1. Enfoque de investigación . . . . .	32
3.4.2. Fases de investigación . . . . .	33
<b>4. Diseño de la solución</b>	<b>43</b>
4.1. Solución del problema . . . . .	43
4.2. Técnicas de análisis . . . . .	45
<b>5. Experimentación</b>	<b>46</b>
5.1. Diseños de experimentos . . . . .	46
5.2. Datasets y Casos de estudio . . . . .	50
5.3. Interpretación de resultados . . . . .	56

<b>6. Implementación</b>	<b>58</b>
6.1. Software utilizado . . . . .	58
6.2. Hardware utilizado . . . . .	61
6.3. Lenguajes de programación . . . . .	62
6.4. Estrategia de implementación . . . . .	64
6.5. Visualización . . . . .	66
<b>7. Conclusiones</b>	<b>81</b>
<b>8. Anexo 1</b>	<b>84</b>
8.1. Implantación . . . . .	84
8.1.1. Requerimientos . . . . .	84
8.1.2. Preparación de Ambiente . . . . .	86
8.2. Manual de usuario . . . . .	88
Anexo 2 . . . . .	95
8.3. Procedimiento de la captura y correlación de la variable de nubosidad utilizando la página web Worldview, proporcionado por la NASA . . . . .	95
Anexo 3 . . . . .	101
Anexo 4 . . . . .	110
8.4. Códigos de implementación . . . . .	110
<b>Bibliografía</b>	<b>116</b>

# Índice de figuras

2.1. Neurona natural, referenciada desde la página web [1]. . . . .	12
2.2. Arquitectura RNA "MLP" . . . . .	13
2.3. Neurona operacional. . . . .	14
2.4. Modelo matemático de D. Beysens. . . . .	18
2.5. Captada por J.-G. Minonzio, de agua del rocío matutino por medio de la tubería de la Universidad de Valparaíso (Noviembre 2021, 09:00 AM). . . . .	19
2.6. Evento de niebla, en Monte Kinabalu Borneo, [2]. . . . .	20
2.7. Punto del rocío en las plantas, obtenida desde [3]. . . . .	21
2.8. Medición de nubosidad en octas, fuente de referencia [4]. . . . .	22
3.1. Diagrama de Metodología del proyecto . . . . .	33
3.2. Diagrama de obtención de datos del proyecto FONDECYT de número 11140863. . . . .	34
3.3. Diagrama de generación, datos instantáneos. . . . .	36
3.4. Diagrama de generación, datos instantáneos, elaboración propia. . . . .	37
3.5. Diagrama de arquitectura RNA "MLP". . . . .	38
3.6. Diagramas de modelo de la RNA. . . . .	39
3.7. Diagrama de modelo matemático realizado por D. Beysens . . . . .	39
4.1. Diagrama de la solución del problema . . . . .	44
5.1. Etapas de la experimentación. . . . .	46
5.2. Diagrama cliente-servidor. . . . .	47
5.3. Funcionamiento del contenedor de Docker. . . . .	48
5.4. Diseño de la experimentación. . . . .	48
5.5. Colector de rocío, proyecto FONDECYT de N.º 11140863, situado en Combarbalá. . . . .	50
5.6. Imagen de la ubicación de Combarbalá, fuente de referencia [5]. . . . .	51
5.7. PLC utilizado en el proyecto FONDECYT de N.º 11140863. . . . .	52
5.8. Parámetros promedio diario (PD), proyecto FONDECYT de N.º 11140863.	52
5.9. Parámetros en horarios (6 AM), proyecto FONDECYT de N.º 11140863. .	53
5.10. Etapas de casos de estudios. . . . .	55

6.1.	Diagrama de Componente del sistema. . . . .	60
6.2.	Esquema Hardware. . . . .	61
6.3.	Imagen del logo de Python . . . . .	63
6.4.	Imagen del logo de SQL . . . . .	63
6.5.	Esquema de estrategia de implementación. . . . .	65
8.1.	Ejecución Docker network y Docker compose. . . . .	87
8.2.	Listado de aplicaciones con el comando Docker compose ps. . . . .	88
8.3.	Jupyter Notebook, IP y puerto. . . . .	88
8.4.	Archivo de datos SSH Keys. . . . .	89
8.5.	Conexión GitHub y Jupyter Notebook. . . . .	90
8.6.	Conexión GitHub y Jupyter Notebook. . . . .	91
8.7.	Conexión GitHub y Jupyter Notebook. . . . .	91
8.8.	Archivos del repositorio para efectuar los experimentos. . . . .	92
8.9.	Experimento número 10. . . . .	93
8.10.	Imagen de conexión GitHub y Jupyter notebook. . . . .	94
8.11.	Imagen de la página Web la NASA. . . . .	96
8.12.	Diagrama para obtener la nubosidad desde la página web Woldview. . . . .	97
8.13.	Imagen o pixel del color asociado a la nubosidad, ubicado en Combarbala. .	98
8.14.	Diagrama del script. . . . .	99
8.15.	Imagen de correlación de nubosidad-script y nubosidad-manual, de un $R^2$ igual a 0.933. . . . .	100

# Índice de tablas

1.1.	Tabla de conclusión de modelos. . . . .	8
2.1.	Gráfico Sigmoide. . . . .	16
2.2.	Gráfico Hiperbólica. . . . .	16
2.3.	Gráfico Relu. . . . .	16
2.4.	Tabla de descripción de métodos de activación [6]. . . . .	16
2.5.	Tabla de Algoritmos de Optimización y descripción. . . . .	17
2.6.	Tabla de resumen de literatura . . . . .	28
3.1.	Resultados Día a día . . . . .	41
3.2.	Resultados Acumulados . . . . .	41
3.3.	Resultados PD y 6 AM . . . . .	41
3.4.	Resumen de resultados de los casos de estudio. . . . .	42
5.1.	Conclusión de modelos. . . . .	53
6.1.	Tabla de modelos D.Beysens . . . . .	67
6.2.	Tabla de modelos RNA . . . . .	68
6.3.	Tabla de modelos acumulados D. Beysens . . . . .	69
6.4.	Tabla de modelos acumulados RNA . . . . .	70
6.5.	Tabla de modelos horario PD . . . . .	72
6.6.	Tabla de modelos horario 6 AM . . . . .	73
6.7.	Tabla de descripción Casos de estudio 1 y 3 “Día a día PD”. . . . .	75
6.8.	Tabla de descripción Casos de estudio 2 y 4 “Acumulado PD”. . . . .	76
6.9.	Tabla de descripción Casos de estudio 5 y 7, Día a día en horario 6 AM. . . . .	77
6.10.	Tabla de descripción Casos de estudio 1 y 3 “Día a día PD” . . . . .	78
6.11.	Resumen de los parámetros de correlación (promedio y desviación estándar de las 10 pruebas), resultados PD. . . . .	79
6.12.	Resumen de los parámetros de correlación (promedio y desviación estándar de las 10 pruebas), resultados 6 AM. . . . .	80
7.1.	Conclusión de modelos . . . . .	82

8.1.	Requerimientos mínimos . . . . .	84
8.2.	Requerimiento recomendado . . . . .	85
8.3.	Requerimiento Utilizados. . . . .	85
8.4.	Tabla de modelo D. Beysens, experimento número 1 . . . . .	102
8.5.	Tabla de modelo RNA, experimento número 1 . . . . .	103
8.6.	Tabla de Acumulado RNA, experimento número 1 . . . . .	104
8.7.	Tabla de Acumulado D. Beysens, experimento número 1 . . . . .	105
8.8.	Tabla de descripción Casos de estudio 1 y 3 “Día a día PD” con las 10 pruebas	106
8.9.	Tabla de descripción Casos de estudio 2 y 4 “Acumulado PD” . . . . .	107
8.10.	Tabla de descripción Casos de estudio 5 y 7 Día a día en horario 6 AM . . . . .	108
8.11.	Tabla de descripción Casos de estudio 1 y 3 “Día a día PD” con las 10 pruebas	109

# Glosario

**Algoritmo:** Representación de instrucciones consistentes y lógicas que buscan en emplear una solución respecto a un problema.

**Artificial:** Es un instrumento o algo no natural, hecho por el ser humano.

**Aprendizaje:** Adquisición de conocimiento por medio de estudio o experiencias. Proceso de los RNA de modificar pesos entre elementos de procesamiento según el error. El Proceso en los seres humanos de adquirir conocimientos y utilizarlos para resolver problemas.

**Análisis:** Estudio de los requerimientos dados por el usuario o cliente por medio de información de los procesos que tiene fin desarrollar un software.

**Capa o nivel:** Agrupación de neuronas en una red neural cuya entrada proviene por medio de la misma fuente y en la capa final las salidas son dirigidas a otras neuronas.

**Conexión:** Elemento que une dos neuronas en una red neuronal, permitiendo el paso de la salida de la neurona origen a la de destino.

**MSE:** Se define el error cuadrático medio como el valor esperado del cuadrado de la diferencia entre el valor medido y el parámetro que trata de estimar.

**Función Activación:** Función aplicada por la neurona a la suma ponderada de las entradas para producir el valor de salida.

**Inteligencia:** Facultad de comprender y conocer.

**Learning Rate:** Tasa de aprendizaje es un parámetro de ajuste en un algoritmo de optimización que determina el tamaño del paso en cada iteración mientras se mueve hacia un mínimo de una función de pérdida.

**Neurona:** Es una célula del sistema nervioso central que posee la capacidad de recibir y decodificar información en forma de señales eléctricas y químicas, transmitiéndolas a otras células.

**Neurona Artificial:** Las neuronas artificiales son microchips que pretenden imitar las funciones de las neuronas biológicas.

**Red:** Conjunto de elementos conectados para cumplir un objetivo.

**RNA:** Red de neuronas artificiales compuestas de un gran número interconectado de elementos que trabajaban en cooperación para resolver problemas específicos.

**Sistema:** Aglomerado de objetos interrelacionados que buscan cumplir un objetivo.

**MLP:** Perceptrón multicapa que corresponde a un RNA.

**Error medio:** Porcentaje medio de inexactitud o equivocación al realizar una operación matemática

**Pendiente:** En matemáticas, la definición de la pendiente se usa para describir la inclinación y dirección de rectas.

**IA:** La inteligencia artificial consiste en la búsqueda de sistema y procesamiento de datos que nos permitan imitar la inteligencia humana, realizando tareas que requieran aprendizaje y solución del problema.

**Coeficiente:** El coeficiente de posición, es el número que señala el punto donde la recta interceptará al eje de las ordenadas y.

**Épocas:** Este es el número de veces que se ejecutarán los algoritmos. En cada ciclo (época) todos los datos de entrenamiento pasan por la red neuronal para que esta aprenda sobre ellos.

**Meteorología:** Es una ciencia encargada del estudio de la atmósfera y los fenómenos que en ella se puedan producir, tales como: el viento, la temperatura, la humedad, la presión del aire, entre otros.

**Datos Training:** Son los datos de entrenamiento que usamos para entrenar un modelo de RNA.

**Datos Testing:** Son los datos de prueba que nos reservamos para comprobar o validar el modelo de RNA que hemos generado en el entrenamiento.

**Rocío:** Es el vapor de agua que con la frialdad de la noche se condensa en la atmósfera en muy pequeñas gotas, las cuales aparecen luego sobre la superficie de la tierra o sobre las plantas.

**NaN:** Proviene del inglés abreviado como "not a number", por lo que son datos no numéricos.

**CW:** Del abreviado "water collection", lo que significa agua colectada o recolectada, también denominada como:  $(\frac{dh}{dt})$ .

**Script:** En informática, un script es secuencia de comandos.

**PD:** Se refieren a la información o datos que se recopila y se presenta en forma de promedio para un día.

**6 AM:** Esta expresión o termino se emplea para describir los datos recopilados a las 6 de la madrugada específicamente.

# **Capítulo 1**

## **Introducción**

Con la actual amenaza desenfrenada del cambio climático y la escasez hídrica a nivel mundial, siendo unos de los mayores desafíos de nuestra época, surge la necesidad de acceder a fuentes suplementarias de agua, tal y como es el rocío, la desalinización, la niebla, las lluvias entre otros fenómenos. Esto cobra una relevancia aún mayor cuando consideramos que más del 70 por ciento de los recursos hídricos del mundo se utilizan para producir alimentos [7]. En este contexto de escasez hídrica a nivel global, donde cada vez son más las ubicaciones afectadas por este problema, se generan diversos desafíos que tienen un impacto significativo en el abastecimiento de alimentos y los precios de estos, además de agravar los efectos del calentamiento global, actividades industriales y el crecimiento de la población. “En América Latina y el Caribe, los niños, niñas y adolescentes viven en zonas de alta vulnerabilidad al acceso al agua, lo que significa que no disponen de agua suficiente para satisfacer sus necesidades diarias” [8]. Dado esta situación actual ocasionada en el medio ambiente, buscamos implementar nuevos métodos y tecnologías para predecir la obtención de agua dulce diaria por medio de fuentes no convencionales, como es el agua del mar y la desalación del agua salobre, reutilización de aguas residuales y captación de agua en la atmósfera, como la lluvia, niebla y rocío. El aumento constante en los costos de las energías en estas últimas décadas, han tenido un impacto negativo en procesos intensivos que demandan o genera grandes cantidades de energía. Nace el interés sobre la recolección de agua de rocío, ya que es un sistema pasivo el cual no requiere energía adicional.

En nuestro trabajo de título, nos centraremos en el estudio del fenómeno del rocío, el cual se forma de manera natural durante las noches y madrugadas en muchas regiones del mundo. Representa una solución complementaria a esta problemática, aunque no se trata de una fuente hídrica tradicional, el rocío puede desempeñar un papel crucial al contribuir al almacenamiento de agua si aprovechamos esta fuente de manera consistente y diaria. Entre los beneficios de captar el agua de rocío, se destaca su capacidad para reducir significativamente los costos financieros y evitar la necesidad de llevar a cabo procedimientos costosos en plantas de tratamiento. Desde una perspectiva económica, esta

técnica es particularmente relevante, ya que no requiere el consumo de energía [9], tampoco la asignación de vastas extensiones de terreno, en contraposición a la desalinización y otros métodos no convencionales para obtener agua. Además, aporta humedad a las plantas, reduce la evaporación, beneficio estético y salud de las plantas. Por lo tanto, su aprovechamiento puede brindar una fuente adicional y sostenible de recurso hídrico, contribuyendo así a abordar la escasez de agua de manera más efectiva.

Llevamos a cabo una revisión exhaustiva de la literatura enfocada en la evaluación de múltiples modelos matemáticos empleados en la predicción de la recolección de agua de rocío. Entre los diversos modelos que analizamos, identificamos dos que demostraron una sólida capacidad para relacionar efectivamente los parámetros de entrada con la variable que pronostica la cantidad de agua de rocío recolectada. El primer modelo fue desarrollado por D. Beysens [10], se presenta como un enfoque semi-teórico que se apoya en los principios de balances de energía y masa. Este modelo también incorpora algunos parámetros empíricos, los cuales se derivan a partir de datos experimentales recopilados en diversas ubicaciones alrededor del mundo. Por otro lado, los modelos de redes neuronales, abreviados como RNA, son de naturaleza completamente empírica. Estos modelos se enmarcan en la estadística multivariante y se caracterizan por la totalidad de sus parámetros que se basan en observaciones y datos empíricos, en particular este modelo utiliza el algoritmo Levenberg-Marquardt [11], también conocido como el método de mínimos cuadrados amortiguados.

Una de las principales limitaciones es que los modelos teóricos y semi-teóricos tienen el problema de la baja precisión para períodos de tiempo de un día o incluso una hora, otra limitación como se menciona en el documento de D. Beysens “*Estimación del rocío mundial*” es la precisión limitada en algunas ubicaciones, Aunque la fórmula proporciona una estimación general del potencial de rocío, su precisión puede verse afectada por la falta de datos precisos de radiación y transferencia de calor. Además, los datos meteorológicos utilizados en la fórmula pueden no estar disponibles en todas las ubicaciones o pueden tener ciertas limitaciones en términos de su representatividad [10]. A diferencia Las redes neuronales tienen la ventaja de poder relacionar múltiples variables en grandes cantidades de datos, incluyendo relaciones altamente no lineales. No obstante, Si los datos de entrenamiento provienen de un lugar específico o de un conjunto limitado de ubicaciones geográficas, el modelo se especializará en esas áreas. Esto significa que sus predicciones serán más precisas en esas ubicaciones específicas, pero menos precisas en lugares que no se parezcan a los datos de entrenamiento (problema de extrapolabilidad geográfica del modelo de red neuronal para esta aplicación en particular). Adicional en el texto “*Predicción de la generación solar a partir del clima Pronóstico mediante aprendizaje automático*” Se menciona que los modelos de RNA pueden tener resultados deficientes en la predicción a largo plazo. Esto se debe a que la precisión de los modelos depende de las condiciones iniciales y estas pueden cambiar con el tiempo [12].

Dado el vasto progreso tecnológico en el ámbito de las herramientas y software informático en estos últimos años, surge la necesidad de imitar e implementar los modelos

arduos, que se presentan en la literatura y que se sustentan en tecnologías actuales. En el contexto del complejo modelamiento de modelos matemáticos estimativos, como la Red Neuronal Artificial (RNA) y el modelo propuesto por D. Beysens, es imperativo considerar la elección de las herramientas tecnológicas adecuadas. Estas elecciones no solo simplifican el proceso, sino que también fortalecen la efectividad y la seguridad de nuestro enfoque. En primer lugar, destacamos el lenguaje de programación Python, una elección sabia debido a su capacidad de aprender y utilizar. Python ofrece una sintaxis legible y una amplia gama de bibliotecas especializadas que simplifican la implementación de algoritmos matemáticos complejos. Esto no solo aumenta la eficiencia de nuestro trabajo, sino que también nos brinda una comprensión más profunda de los modelos que estamos desarrollando. A fin de gestionar y ver nuestros modelos, gráficos y los resultados derivados de ellos, recurrimos al potente entorno Jupyter Notebook. Este software es una herramienta vital que facilita la comunicación y presentación de nuestros hallazgos. Al proporcionar un entorno interactivo y la capacidad de documentar el proceso de modelado. En el ámbito de la gestión de datos, nuestra elección recae en el sistema de gestión de bases de datos MySQL. Almacena con seguridad nuestros modelos y resultados, lo que es fundamental para mantener la integridad de nuestros datos sensibles. En la era actual de amenazas ciberneticas, la seguridad de las conexiones y la protección de datos sensibles son cruciales. Es por eso que implementamos Nginx, un servidor web y proxy inverso altamente confiable, para salvaguardar nuestras comunicaciones y datos. Por último el software Docker nos permite ejecutar nuestro proyecto en un entorno aislado conocido como contenedor, este enfoque simplifica la implementación del conjunto de nuestras tecnologías que fueron mencionadas anteriormente.

La propuesta de solución es suplementaria o parcial, dado que el fenómeno del rocío proporciona un complemento de agua dulce en áreas semiáridas y áridas [13]. Esto se debe a que la cantidad de rocío disponible suele ser limitada en comparación con fuentes de agua más consistentes y significativas, como la lluvia o ríos, etc. El rocío se forma principalmente durante las noches, cuando las temperaturas desciden y el aire se enfria, lo que conduce a la condensación del vapor de agua presente en el aire en forma de pequeñas gotas sobre las superficies expuestas, como las plantas. Debido a su dependencia de las condiciones climáticas y a su producción variable, el agua de rocío no puede satisfacer de manera constante o confiable las necesidades hídricas, lo que la convierte en una fuente secundaria o suplementaria en comparación con otras fuentes de agua más predecibles y controlables. Para finalizar este fenómeno tiene un gran impacto en los consiguientes sectores, en ámbito económico, social, ambiental e informático.

La solución del trabajo de título se desarrolla en una primera instancia en la incorporación de las bases de datos a nuestras herramientas informáticas mencionadas en el párrafo anterior, estos datos fueron recopilados en el proyecto FONDECYT 11140863, fue realizado por la Universidad de La Serena y bajo la dirección del profesor Danilo Carvajal. Una vez importada la bases de datos de recolección de agua de rocío, realizamos una limpieza de las bases de datos suprimiendo algunas filas y columnas que para el experimento no son

redundantes, posteriormente eliminamos los datos que no utilizamos en esta investigación, algunos de estos datos son denominados NaN (*not a number*). Después, generamos dos separaciones equilibradas de datos, la primera consiste en dividir el rocío pequeño, medio y alto. la segunda segmentación es para el correcto ingreso a los modelos experimentales empleados, consiste en dos tercios (2/3) de datos de “entrenamiento” y un tercio (1/3) de estos datos para validar nuestros modelos denominados “pruebas”. Estas pruebas consistirán en diez iteraciones con diferentes datos definidos aleatorios, lo que permitirá comprender el comportamiento los modelos y el resultados de las correlaciones. El primer modelo es el propuesto por el científico de alto conocimiento relacionado con el fenómeno del rocío Daniel Beysens [10], a partir de una fórmula para predecir la recolección de agua de rocío medido en  $L/m^2/día$ , esta fórmula contiene los siguientes parámetros de entrada en función de estimar la recolección del flujo de agua de rocío, estos son: la temperatura del aire ( $T_a$ ), temperatura del rocío ( $T_d$ ) o Humedad relativa ( $RH$ ), Velocidad del viento ( $V$ ) y por último la nubosidad ( $N$ ). Tras la planificación y ejecución de las correlaciones, se emplea una acumulación o sumatoria en función de los resultados de correlaciones experimentales, tal como se describe en el documento de D. Beysens [10]. El segundo modelo es el propuesto en el documento de Lekouch [11], el cual aplican un modelo de arquitectura basado en una RNA denominada MLP, que contiene 3 capas, la primera capa es nombrada entrada y posee 4 neuronas, sumado a ello tenemos una capa oculta de 5 neuronas y por último una capa de salida de 1 neurona. Las variables de entradas son: Temperatura del aire ( $T$ ), Humedad relativa ( $HR$ ), Velocidad del viento ( $V$ ) y por última la nubosidad ( $N$ ). Al igual que el modelo de D. Beysens descrito anteriormente, se emplea una acumulación de resultados por medio de las correlaciones diarias, concluyendo con la comparación de ambos modelos a través de una tabla de resumen con el propósito de obtener los resultados de los modelos idóneos en el norte de Chile, estas estimaciones representaran la cantidad del flujo de recolección de agua de rocío medida en  $L/m^2/día$ .

Para finalizar nuestra investigación se realizó una tabla que contiene los datos, métodos y análisis de resultados correspondiente a las correlaciones de coeficiente de determinación  $R^2$  y pendientes  $m$ . Con el objetivo de concluir el desempeño de cada caso, los cuales se distinguen en la tabla 1 y pueden ser: bueno, regular, marginal y malo.

Datos	Métodos	Correlación	#	Desempeño	$R^2 / m$
PD	Beysens	Día a día	1	Bueno, regular, marginal y malo	$R^2 / m$
		Acumulado	2	Bueno, regular, marginal y malo	$R^2 / m$
	RNA	Día a día	3	Bueno, regular, marginal y malo	$R^2 / m$
		Acumulado	4	Bueno, regular, marginal y malo	$R^2 / m$
6 AM	Beysens	Día a día	5	Bueno, regular, marginal y malo	$R^2 / m$
		Acumulado	6	Bueno, regular, marginal y malo	$R^2 / m$
	RNA	Día a día	7	Bueno, regular, marginal y malo	$R^2 / m$
		Acumulado	8	Bueno, regular, marginal y malo	$R^2 / m$

Tabla 1.1: Tabla de conclusión de modelos.

La tabla previa reviste una importancia fundamental, ya que se considera la conclusión de los modelos. Esta tabla nos permitió analizar y distinguir los resultados estimados ideales o perfectos en función de los gráficos, las correlaciones y los casos de estudio.

Las contribuciones principales provinieron del investigador D. Beysens, de nacionalidad francesa, reconocido como un pionero en el campo de la predicción meteorológica. El modelo matemático presentado en el documento [10] desempeñó un papel fundamental en el desarrollo de este estudio, al servir como una referencia crucial para la comparación con las predicciones generadas por las Redes Neuronales Artificiales (RNA). Además, destacamos la labor de los investigadores D. Carvajal y J.-G. Minonzio, quienes se han destacado como prominentes exponentes en el ámbito de la investigación bajo el título “*Recolección de agua de rocío integrada en techos en Combarbalá, Chile*” [13]. Mientras que la recolección de agua de rocío se ha experimentado a pequeña escala en distintas ubicaciones alrededor del mundo, es crucial resaltar que este documento representa una contribución única en su género. En este trabajo, se llevaron a cabo experimentos a gran escala que, hasta la fecha, carecían de documentación. El objetivo central de esta investigación consistió en evaluar un año completo de recolección de agua de rocío en el norte de Chile, específicamente en un techo galvanizado de 36 metros cuadrados.

El documento se estructura en siete capítulos, el primer capítulo se dedica a la introducción, el segundo al marco conceptual con los principales conceptos y definiciones, seguido por la recopilación de obras de arte, Por consiguiente el capítulo 3, se encontrará la definición del problema y la hipótesis donde se establece la formulación bajo el contexto actual, seguido se propone una solución suplementaria o parcial para la recolección del rocío, esta propuesta está relacionada con el modelo de D. Beysens [10] y la red neuronal artificial [11], con la ayuda de los datos experimentales y estos modelos mencionados anteriormente nos permite predecir y anticipar la cantidad de agua de rocío diaria o acumulativa. Posteriormente, se definen la importancia e impacto de la solución, los objetivos a cumplir y la metodología a utilizar. En el capítulo 4 establece el diseño de la solución, donde se explica como se implementara al igual que se modelara la red neuronal y el modelo matemático, a través de un diagrama de solución del problema para finalizar se explicara las técnicas para

el análisis de datos. En el capítulo 5 se definen las etapas de experimentos igualmente los casos de estudio y por último se interpretan los resultados. El capítulo 6 establece la implementación, El cual tiene como objetivo principal explicar el software, así como el hardware utilizado, los lenguajes de programación, la estrategia de implementación y por último la visualización de los resultados. Para concluir con el capítulo 7, el cual se designa como las conclusiones, establece a grandes rasgos el proceso del trabajo realizado, la implementación de los resultados, el cumplimiento del objetivo, la hipótesis, las dificultades presentadas durante el desarrollo del informe, además de la mirada personal hacia el futuro.

# **Capítulo 2**

## **Marco Conceptual y Estado del Arte**

### **2.1. Marco Conceptual**

Esta sección se fundamenta en la condimentación teórica de las variables de estudio, tales como los parámetros ambientales de las RNA derivadas de la documentación [11] y el modelo matemático elaborado por D. Beysens en el documento [10] Pasando luego por la revisión de la literatura, donde se evalúan varios estudios realizados con anterioridad en el ámbito, y extremadamente relacionados con esta investigación, de igual forma se empleará la definición de términos básicos, donde se aclara cualquier confusión que pueda presentar el lector, y finalmente la definición conceptual y operacional de las variables, establecidas de la siguiente forma.

→ **Definición de términos básicos conceptuales.**

- Inteligencia Artificial (IA).
- Redes Neuronales Artificiales (RNA).
- Modelo matemático estimativo D. Beysens.

→ **Variables de estudios o parámetros medioambientales.**

- Flujo de Agua de Rocío ( $\frac{dh}{dt}$ ).
- Humedad Relativa ( $HR$ ).
- Temperatura aire ( $Ta$ )
- Velocidad del viento ( $V$ ).
- Nubosidad ( $N$ ).
- Temperatura del punto de rocío ( $Td$ ).

### **2.1.1. Definición de términos Básicos conceptuales:**

- **Inteligencia Artificial:** Consisten en la búsqueda y desarrollo de un sistema que contenga un procesamiento de datos que sean capaces de imitar la red neural de inteligencia humana (IA), permitiéndonos resolver ciertos problemas que requieran un aprendizaje, con el fin de una estimación, como resultados. El propósito final es reproducir todas las facetas de la inteligencia humana. Se destaca la red neuronal artificial(RNA), ya que han sido utilizadas con éxito en la predicción de patrones en datos de series temporales de lluvia, y que han demostrado ser más precisas que otros modelos en el pronóstico de parámetros meteorológicos a largo plazo a nivel de distrito. Además, se mencionan diversas aplicaciones de la IA, como la simulación del cambio climático, la predicción de inundaciones y la estimación de la temperatura y la humedad relativa. A partir de 1986, el panorama basado en las RNA fue alentador respecto a las investigaciones y desarrollo de las redes neuronales basado en el contexto de la década de los 90, en la actualidad son numerosos los trabajos investigativos que se realizan y publican cada año, en conjunto de nuevas tecnologías que surgen en el presente [12].
- **Red Neuronal Artificial (RNA):** Se desarrollan bajo la estructura básica de una neurona natural como se conecta mediante dendritas, cada neurona contiene infinitas entradas nombradas dendritas, que condicionan el estado de su única salida, el axón, este puede ir conectado a una dendrita de otra neurona mediante sinapsis química, los cuales podemos visualizar en la siguiente figura 2.1. A demás la RNA son una técnica matemática eficiente y robusta de regresión no lineal, capaz de realizar diversas capacidades como el reconocimiento de patrones, la clasificación y las predicciones, como se menciona en documento de Lekouch [11].

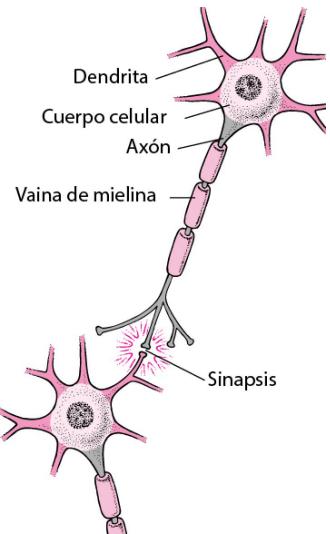


Figura 2.1: Neurona natural, referenciada desde la página web [1].

- **Modelo red neuronal MLP:** El Modelo MLP es utilizado para abordar problemas reales del ámbito del medio ambiente entre otras, es de similar estructura de la red neuronal cerebral humana como se menciona "la arquitectura del Perceptrón Multicapa debido a su amplio uso en el ámbito medioambiental. Esta arquitectura utiliza una conexión de retroalimentación entre las capas, lo que significa que la información se transmite de una capa a la siguiente a través de funciones de activación, pesos y sesgos" [11]. La neurona contiene una gran habilidad para aprender a partir de datos en primer momento, en entrenamientos y luego de validación, permitiéndonos resolver problemas de estimación de acuerdo a variables independientes y relaciones no lineales. Las redes neuronales son modelos creados al ordenar operaciones matemáticas siguiendo una determinada estructura, estas han sido ampliamente utilizadas en el dominio ambiental con fenómenos como; el Rocío, la niebla, las heladas, entre otros. La forma común de representar la estructura de una red neuronal es mediante el uso de capas (layers), formadas a su vez por neuronas (unidades, units o neurons). Cada neurona realiza una operación sencilla y está conectada a las neuronas de la capa anterior y de la capa siguiente mediante pesos, cuya función es regular la información que se propaga de una neurona a otra. La imagen 2.2 de arquitectura MLP, contiene ‘n’ neuronas en la capa de entradas de acuerdo a los parámetros independientes, ‘m’ neuronas en la capa oculta y por último la capa de salida asociada a una última neurona que nos permite la estimación de acuerdo al modelo efectuado.

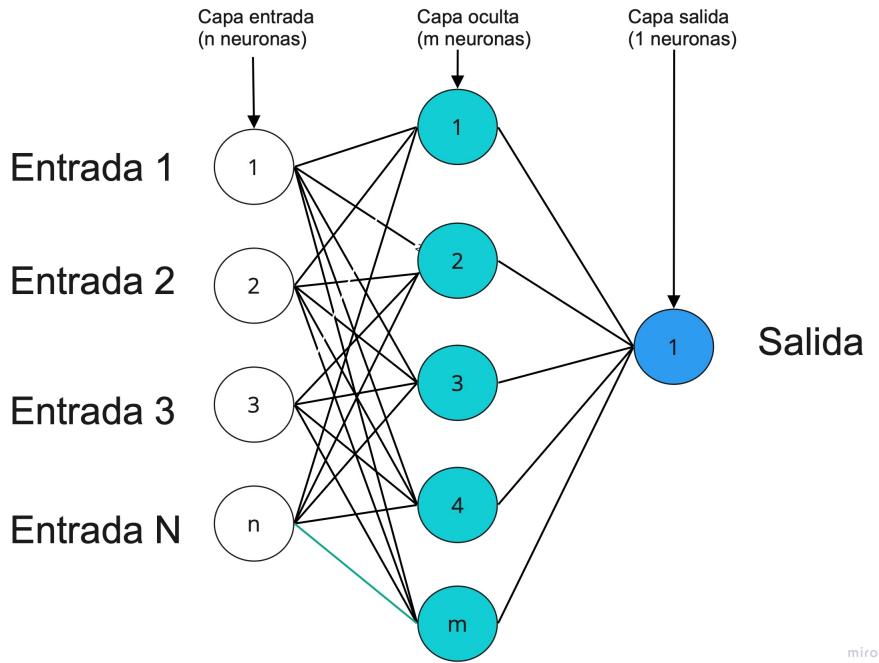


Figura 2.2: Arquitectura RNA "MLP"

**Estructura de la red:** La estructura de la red son múltiples neuronas interactuando por ende, la representación gráfica será a través de capas(layer), ya que nos permite la forma ordenada de la representación de acuerdo al modelo de la RNA, tal como en la figura 2.2, podemos divisar que contiene una primera capa de la red neuronal (color blanco) se conoce como capa de entrada (input layer) y recibe los datos en bruto, es decir, el valor de los predictores. La capa intermedia (color turquesa), conocida como capa oculta (hidden layer), recibe los valores de la capa de entrada, ponderados por los pesos (flechas negras). La última capa, llamada capa de salida (output layer), combina los valores que salen de la capa intermedia para generar la predicción, esto ocurre luego de la división de datos como se menciona en el documento [11].

**Neurona:** Es la unidad funcional de los modelos de redes. Dentro de cada neurona, ocurren simplemente dos operaciones: la suma ponderada de sus entradas y la aplicación de una función de activación, se puede evidenciar en la figura 2.3, es una neurona operacional, como se expresa en la página web de Machine Learning con Python [6].

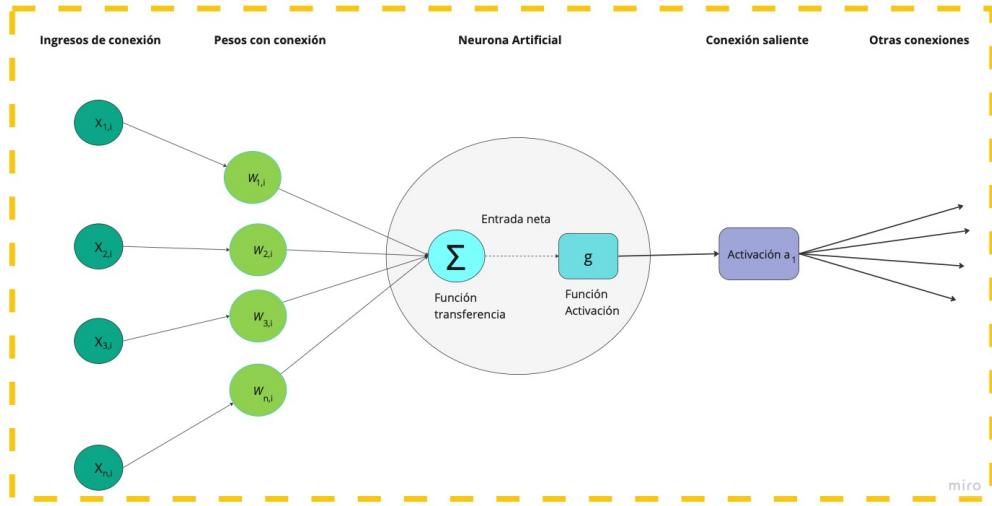


Figura 2.3: Neurona operacional.

Como funciona la neurona estructural, se multiplica cada valor de entrada “ $x_i$ ” por su peso asociado “ $w_i$ ”. Este es el valor neto de entrada a la neurona. A continuación, este valor se pasa por una función, conocida como función de activación, que transforma el valor neto de entrada en un valor de salida [6].

El valor neto de entrada a una neurona es la suma de los valores que le llegan, ponderados por el peso de las conexiones.

$$\text{entrada} = \sum_{i=1}^n x_i w_i + b$$

En lugar de utilizar la sumatoria, esta operación suele representarse como el producto matricial, donde X representa el vector de los valores de entrada y W el vector de pesos.

$$\text{entrada} = XW + b$$

A este valor se le aplica una función de activación ( $g$ ) que lo transforma en lo que se conoce como valor de activación ( $a$ ), que es lo que finalmente sale de la neurona.

$$a = g(\text{entrada}) = g(XW + b)$$

Para la capa de entrada, donde únicamente se quiere incorporar el valor de los predictores, la función de activación es la unidad, es decir, sale lo mismo que entra. En

la capa de salida, la función de activación utilizada suele ser la identidad o sigmoide para problemas de regresión relacionados con la estimación del rocío [6].

**Función de activación:** La función de activación se encarga de devolver una salida a partir de un valor de entrada, normalmente 0 en la gran mayoría el conjunto de valores de salida en un rango determinado como  $(0,1)$  o  $(-1,1)$  “Se buscan funciones que las derivadas sean simples, para minimizar con ello el coste computacional” [14]. A continuación, se presentarán las características fundamentales de las funciones de activación que se mencionarán desde la página web “Redes neuronales con Python” [6], estas serán enumeradas en la tabla siguiente.

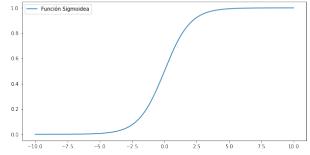
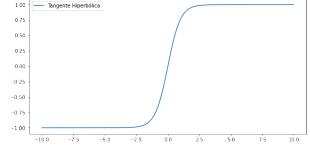
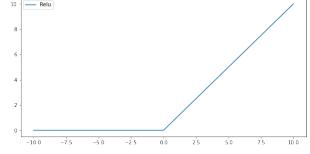
Métodos de activación	Características	Formula	Gráficos
Sigmoide	La función sigmoide transforma valores en el rango de (-inf, +inf) a valores en el rango (0, 1).	$Sigmoid(X) = \frac{1}{1 + exp(-x)}$	 <p>Tabla 2.1: Gráfico Sigmoidal.</p>
Tanh	Tanh es como un sigmoide a escala y tiene problemas similares con los gradientes que la función sigmoide original. El ajuste de los pesos suprimirá los problemas del modelo con los gradientes de desaparición.	$Tanh(x) = \frac{1 - exp(-2x)}{1 + exp(-2x)}$	 <p>Tabla 2.2: Gráfico Hiperbólica.</p>
Relu	Las funciones de Relu ayudan a lograr una convergencia rápida, por lo que el modelo se entrena rápidamente. También es más rápido calcular derivadas en una función esencialmente lineal.	$Relu(x) = max(x, 0)$	 <p>Tabla 2.3: Gráfico Relu.</p>

Tabla 2.4: Tabla de descripción de métodos de activación [6].

**Función de costo:** La función loss o función de costo, también llamada función de pérdida, es la encargada de cuantificar la distancia entre el valor real y el valor predicho por la red, en otras palabras, mide cuánto se equivoca la red al realizar predicciones [6].

**Error cuadrático medio:** El error cuadrático medio (mean squared error, MSE) es con diferencia la función de coste más utilizada en problemas de regresión. Para una determinada observación, el error cuadrático se calcula como la diferencia al cuadrado entre el valor predicho y el valor real [11].

**Entrenamiento:** El entrenamiento de una red neuronal artificial (RNA) implica ajustar los pesos y sesgos de la red para que pueda realizar predicciones precisas. Durante el entrenamiento, la red se expone a un conjunto de datos de entrada y se ajustan los pesos y sesgos en función de la diferencia entre las salidas reales y las salidas predichas por la red. El objetivo es minimizar el error entre las salidas reales y las predichas [12].

**Backpropagation:** Es el algoritmo que permite cuantificar la influencia que tiene cada peso de la red en sus predicciones. En el caso de las redes, la derivada parcial del error respecto a un parámetro peso mide cuanta “responsabilidad” ha tenido ese parámetro en el error cometido [6].

Existen varios tipos de algoritmos de optimización para una red neuronal, nos centramos en dos tipos conocidos, los cuales son; Levenberg-Marquardt y Adamax, que serán detallados a continuación en la siguiente tabla.

Algoritmo de Optimización	Descripción
Levenberg-Marquardt (LM)	Es utilizado para la minimización de la función de error, es el Algoritmo de Levenberg-Marquardt, con retro propagación para el cálculo del gradiente del error [11].
Adamax	El algoritmo utilizado es Adamax que es una variante de Adam basada en la norma del infinito, además Adamax es superior a Adam.

Tabla 2.5: Tabla de Algoritmos de Optimización y descripción.

- **Modelo matemático D. Beysens:** El modelo de Beysens es un modelo semi-teórico basado en balances de energía y masa, con algunos parámetros empíricos obtenidos a partir de datos experimentales de distintos lugares del mundo. El modelo matemático propuesto presenta una fórmula analítica para estimar la producción de rocío, en el caso de que sea negativo el resultado será denominado como '0', Tal cual como se presenta la fórmula y los parámetros atmosféricos que son considerados en el documento de D. Beysens [10].

$$\frac{dh}{dt} = \begin{cases} (0.37 \cdot (1 + 0.204323H - 0.0238893H^2 - (18.0132 - 1.04963H + 0.21891H^2) \cdot 10^{-2}Td)) \cdot \left(\frac{T_d + 273.15}{285}\right)^4 \cdot (1 - \frac{N}{8}) \\ + (0.06(T_d - T_a)) \cdot (1 + 100 \cdot (1 - e^{-\frac{(V)}{V_0}})^{20}), & \text{si es positivo} \\ 0, & \text{si es negativo} \end{cases} \quad (2.1)$$

El modelo está compuesto por las siguientes variables de entradas medioambientales, estas son;

1. Td: Dew temperatura C°.
2. H: Altura m.
3. N: Nubosidad Octa.
4. V: Velocidad de viento m/s.
5. Ta: Temperatura Aire C°.

A partir de la fórmula obtenemos el parámetro de salida o la variable a medir que es la variable de salida es el flujo de rocío estimativo por una noche en mm, denominado ( $\frac{dh}{dt}$ ). Estas Variables serán representadas a continuación en la siguiente figura 2.4, que interpreta el modelo matemático de D. Beysens.

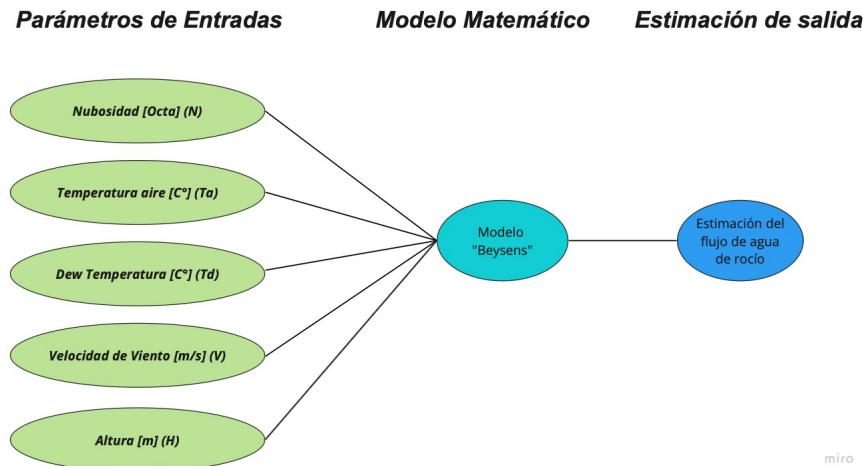


Figura 2.4: Modelo matemático de D. Beysens.

## 2.1.2. Variables de estudios

**Rocío:** Se menciona en el texto *"Frecuencia de rocío en los EE.UU. De una red de radiómetros in situ"*, El rocío se forma principalmente en la parte superior del dosel de los pastizales debido a un enfriamiento radiativo más efectivo y dentro del dosel del bosque debido a una mayor humedad relativa, especialmente durante las noches de cielo despejado y viento en calma. Durante estas condiciones, el suelo pierde calor, enfriando la capa de aire en contacto con él, lo que da lugar a la formación de pequeñas gotas de agua conocidas como rocío [15].

La siguiente fotografía 2.5 fue capturada alrededor de las 09:00 AM. En la imagen se observa un día normal en Valparaíso, cerca de la costa, con alta nubosidad. Podemos notar, a través del sistema de canaletas y los techos del edificio Hucke de la Universidad de Valparaíso, la salida de agua de rocío debido a la humedad relativa presente en el ambiente.



Figura 2.5: Captada por J.-G. Minonzio, de agua del rocío matutino por medio de la tubería de la Universidad de Valparaíso (Noviembre 2021, 09:00 AM).

**HR:** Es la cantidad de vapor de agua que se tiene respecto a la cantidad necesaria para llegar al punto de saturación expresada en porcentaje, un aire saturado tiene una humedad relativa del 100 por ciento, indica que el aire ya no puede almacenar

más vapor y desde ese momento el vapor adicional se convertirá en agua líquida o cristales de hielo, como se verifica en la posterior imagen 2.6, perteneciente al bosque en Monte Kibalu.

También la humedad relativa depende de la temperatura ambiente; una temperatura más baja conduce a una humedad relativa más alta, como se menciona en el siguiente documento [13].



Figura 2.6: Evento de niebla, en Monte Kinabalu Borneo, [2].

#### **Punto del rocío ( $T_d$ ):**

Es la temperatura a la cual debe enfriarse el aire para que se sature de vapor y comience a condensarse como niebla, entre más se acerque a la temperatura al punto de rocío mayor será la probabilidad de condensación.

La temperatura del punto de rocío es, como menciona en el documento “la temperatura a la que se debe enfriar el aire para alcanzar la saturación (suponiendo que la presión del aire permanezca igual). Cuando la temperatura se enfriá hasta el punto de rocío, se puede producir niebla o rocío y la humedad relativa llega al 100 por ciento” [16], como se visualiza en la siguiente imagen 2.7, el rocío es representativo y se queda alojado en la superficie de las hojas de las plantas.



Figura 2.7: Punto del rocío en las plantas, obtenida desde [3].

La ecuación de  $Td$  o temperatura de rocío contiene las siguientes variables;  $Ta$  que es la variable denominada temperatura del aire,  $a$  ser igual al valor 17.27,  $b$  de valor 237.7,  $HR$  perteneciente a la variable de Humedad relativa. Por medio de estos parámetros es posible calcular el punto de rocío [16].

$$Td = \frac{b * \left( \frac{a * Ta}{(a + Ta) + ln(HR)} \right)}{a - \left( \frac{a * Ta}{(a + Ta) + ln(HR)} \right)}$$

**Temperatura del aire ( $Ta$ ):** Se define como una magnitud física descriptiva que mide el estado materia y la forma como esta intercambia la energía térmica del ambiente. Se menciona en el sitio web nombrado Meteoblue [17], que la temperatura del aire es la característica del clima que nos sentimos más directamente con el Calor y frío; fresco, al igual que pesado; caliente y helado. Adicional, como se menciona en la página web INIA [18], la temperatura del ambiente se encuentra en un equilibrio, “La temperatura indica en valores numéricos el nivel de energía interna que se encuentra en un lugar en ese momento. Esta energía interna se encuentra en equilibrio entre el sistema (plantas, animales, etc.).”

**Velocidad del Viento ( $V$ ):** Hace referencia a cuán rápido se mueve el aire en un lugar específico. En este contexto, ”vector con dirección y magnitud” significa que la velocidad del viento se representa como una cantidad con información sobre la dirección en la que se desplaza el viento y cuán rápido lo hace.

Cuando hablamos de ”factores que inciden en la velocidad del viento”, nos referimos a que las condiciones climáticas no sean la adecuada, se estima en el documento ”*Recolección de agua de rocío integrada en el techo en Combarbalá Chile*”, La mayoría de los eventos de rocío ocurren para velocidades del viento inferiores a 4,0

m/s, ya que velocidades más altas pueden detener la formación de rocío al aumentar la transferencia de calor por convección a la superficie del condensador [13].

**Nubosidad ( $N$ ):** Es la fracción del cielo cubierta por nubes observada en un lugar determinado, como se visualiza en la página web [4], estas dividen la bóveda celeste en ocho partes y la nubosidad se mide en octas. Como se logra visualizar en la posterior imagen 2.8, la nubosidad sitúa desde 0/8 que indica un cielo completamente despejado hasta 8/8 para un cielo completamente cubierto. Nos centraremos en este último en el cual encontramos el fenómeno nombrado rocío. La nubosidad total  $N$  (octas) como en la siguiente imagen, que es una caracterización relativa simple del efecto de enfriamiento radiativo entre la superficie de condensación y la fracción de cielo despejado, como se describe en el siguiente artículo [11].

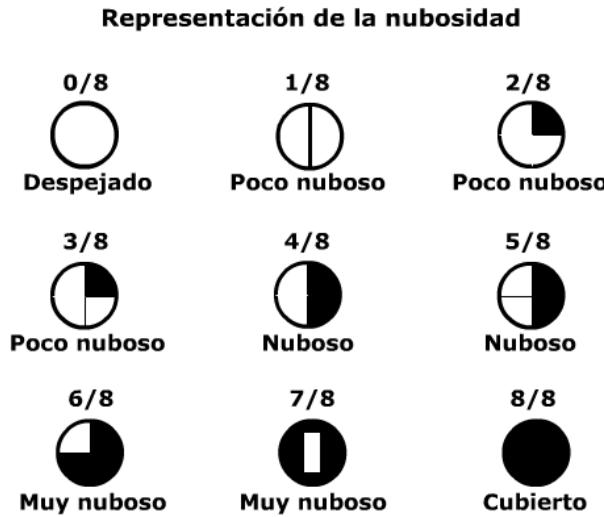


Figura 2.8: Medición de nubosidad en octas, fuente de referencia [4].

## 2.2. Estado del Arte

En el proceso de elaboración del estado del arte, una herramienta esencial para el estudio y análisis de la problemática que estamos abordando, nos dedicamos a recopilar información de siete fuentes literarias distintas. Estas fuentes abarcan un período que va desde 2012 hasta 2018 y se centran en la investigación relacionada con la recolección y estimación del rocío a través de modelos matemáticos para su predicción. Esta recopilación literaria se convierte en un recurso fundamental para nuestra investigación de tesis.

Una vez revisados cada uno de los artículos, procedimos a resumir los hallazgos más relevantes relacionados con el fenómeno del rocío. A continuación, presentaremos estos resúmenes.

1. La primera literatura fue realizada en Massachusetts, EE.UU en el año 2011 por un grupo de investigadores conformado por: N. Sharma, P. Sharma, D. Irwin, P. Shenoy, el cual es nombrado “*Predicción de la generación solar a partir del clima Pronóstico mediante aprendizaje automático*” El artículo presenta un estudio sobre la generación de modelos de predicción para la energía solar utilizando técnicas de aprendizaje automático. El objetivo es abordar el desafío de predecir la generación de energía renovable, que es intermitente e incontrolable. Se utilizan pronósticos meteorológicos del Servicio Meteorológico Nacional (NWS) como datos de entrada para desarrollar modelos de predicción específicos del sitio. Se comparan diferentes técnicas de regresión, incluyendo mínimos cuadrados lineales y máquinas de vectores de soporte, y se evalúa su precisión utilizando datos históricos y lecturas de intensidad solar. Los resultados muestran que los modelos basados en SVM son un 27 porciento más preciso que los modelos existentes basados en pronósticos del tiempo. El estudio destaca la importancia de generar modelos de predicción automatizados para escalar la generación distribuida de energías renovables en la red [19].
2. El segundo documento es realizado por los siguientes investigadores: Gyanesh Shrivastava, Sanjeev Karmakar, Manoj Kumar Kowar, Pulak Guhathakurta en (2012), en su trabajo nombrado ”*Aplicación de redes neuronales artificiales en la predicción del tiempo y una revisión exhaustiva de la literatura*” es una revisión exhaustiva de la literatura sobre el uso de redes neuronales artificiales (RNA) en la predicción de patrones de lluvia monzónica a largo plazo. Los autores mencionan que las RNA son capaces de manejar problemas de no linealidad mejor que las técnicas estadísticas convencionales existentes.

El objetivo del estudio es ampliar y evaluar las RNA para la predicción de patrones de lluvia monzónica. Los autores revisan la literatura desde 1923 hasta 2012 y encuentran que las arquitecturas de RNA como BPN y RBFN son las más establecidas para pronosticar el comportamiento caótico de las lluvias monzónicas.

Se mencionan varios estudios que han utilizado RNA para predecir parámetros meteorológicos como la temperatura, la escorrentía de lluvia y las precipitaciones. Estos estudios han encontrado que las RNA son más eficientes que los modelos estadísticos en la predicción de las lluvias monzónicas a largo plazo.

Además, se mencionan aplicaciones de las RNA en la predicción del nivel de marea, el modelado de lluvia y escorrentía, y la clasificación de la variabilidad de las precipitaciones [12].

3. Un grupo de científicos compuesto por D. Beysens, I. Lekouch, M. Muselli, A. Mongruel y B. Kabbachi en el año 2012, trabajaron en una investigación titulada "*Recolección de rocío, niebla y lluvia en azoteas en el suroeste de Marruecos y modelado estimativo de rocío usando redes neuronales*", el estudio se centró en la recolección de rocío, niebla y lluvia en los tejados del suroeste de Marruecos y su predicción utilizando una red neuronal artificial (RNA). Se recopilaron datos meteorológicos fiables de Mirleft, una región costera de Marruecos, y se utilizaron como entrada para la RNA. El objetivo era estimar la producción de rocío en función de parámetros como la temperatura del aire y del punto de rocío, la velocidad del viento y la nubosidad. Se utilizó una RNA con algoritmo de entrenamiento levenger-Marquardt de tres capas, una inicial de 1 con una capa oculta de 5 neuronas y una capa de salida correspondiente al rendimiento de rocío previsto. Se realizaron pruebas con diferentes números de neuronas ocultas y se encontró que 5 neuronas proporcionaban los mejores resultados en términos de precisión y generalización. El conjunto de datos se dividió en un conjunto de entrenamiento (240 días) y un conjunto de validación (125 días). Durante el entrenamiento, se optimizaron las ponderaciones y sesgos de la RNA para minimizar el error cuadrático medio (MSE). Se encontró una buena correlación entre las predicciones de la RNA y las mediciones reales de la producción de rocío, con un coeficiente de determinación ( $R$  cuadrado) de 0,9194. Los resultados mostraron que la producción de rocío estaba fuertemente correlacionada con los parámetros meteorológicos mencionados anteriormente. Se encontraron dos regiones principales con altos rendimientos de rocío: al oeste de las montañas del Atlas y cerca de las regiones costeras. Por otro lado, se encontraron bajos rendimientos de rocío al este de las montañas del Atlas. En cuanto a la recolección de agua de rocío para satisfacer las necesidades humanas, se determinó que se requerían superficies de recolección grandes, aunque razonables. Estas superficies podrían ser cubiertas parcial o totalmente por techos familiares si cuentan con una infraestructura adecuada, como canaletas, tuberías y un tanque de almacenamiento [11].
4. El artículo es nombrado "*Estimación de la producción de rocío en todo el mundo a partir de unos pocos datos meteorológicos*" fue realizado por D. Beysens en el año 2016, el documento presenta una fórmula analítica simple para estimar la producción de rocío en diferentes lugares del mundo. La fórmula se basa en experimentos de laboratorio y utiliza datos meteorológicos comunes, como la cobertura de nubes, la velocidad del viento, la temperatura del aire y el punto de rocío. El documento muestra una buena concordancia entre los valores calculados utilizando la fórmula propuesta y los valores experimentales recolectados en 10 sitios diferentes. Además, se realizó un análisis de correlación que demostró una fuerte correlación entre los rendimientos de rocío experimentales y calculados, respaldado por un alto coeficiente de determinación  $R$  cuadrado y acumulados. Estos resultados indican que la fórmula analítica propuesta

es una herramienta precisa y útil para estimar el potencial de rocío en diferentes ubicaciones del mundo. La capacidad de predecir la producción de rocío puede ser especialmente beneficiosa en regiones áridas y subtropicales donde el acceso al agua limpia es limitado o inexistente. En conclusión, la fórmula analítica propuesta en el documento proporciona una estimación confiable y precisa de la producción de rocío en diferentes lugares del mundo. Esto puede ayudar a abordar los desafíos de escasez de agua y brindar acceso a agua dulce limpia y de bajo costo en regiones donde es más necesario[10].

5. D. Carvajal, J.-G. Minonzio, E. Casanga, J. Muñoz, A. Aravena, Montecinos y D. Beysens(2018) en su trabajo de investigación titulado "*Recolección de agua de rocío integrada en el techo en Combarbalá Chile*" El documento presenta un estudio sobre la captación de agua de rocío en Combarbalá Chile, evaluó de la captación de agua de rocío integrada en el techo. Se utilizó un techo de acero galvanizado de  $36\text{ m}^2$  inclinado a 15 grados, cubierto con una pintura de alta emisividad que contenía minerales de aluminosilicato para aumentar la producción de rocío. Se registraron los rendimientos de agua de rocío diariamente durante un año. Se encontró que la tasa de recolección promedio diaria de rocío fue de 1,9 L/día, con un máximo de 15 L/día. La producción máxima de rocío está fuertemente correlacionada con la humedad relativa y débilmente con la temperatura del aire y la velocidad del viento. Se estima que el agua de rocío podría contribuir aproximadamente al 8,2 del agua total recogida en un año, considerando también la recolección de lluvia. El sitio de estudio, El Chañar, se encuentra en una zona semiárida y las precipitaciones de lluvia son erráticas, afectadas por el fenómeno de El Niño-Oscilación del Sur (ENSO). Durante el período de estudio, se observó una deposición anual de rocío nocturno. El estudio también menciona que los techos a dos aguas fabricados con láminas de acero galvanizado corrugado son ampliamente utilizados en la zona. Se instalaron canalones de recogida y depósitos de almacenamiento en ambos lados del techo para la recolección de agua de rocío. Además, se destaca que los techos de acero pueden reducir la carga microbiológica del agua recolectada. Sin embargo, no se realizaron análisis de la calidad del agua de rocío en este estudio. En conclusión, el estudio demuestra que la captación de agua de rocío integrada en el techo puede ser una fuente adicional de agua en zonas semiáridas como Combarbalá, Chile. La producción de rocío está influenciada por factores como la humedad relativa, la temperatura del aire y la velocidad del viento. Los techos de acero galvanizado corrugado son una opción común para la captación de agua de rocío en la zona[13].
6. La última literatura realizada por D. Beysens, M. Berkelhammer, F. Ritter, nombrada como "*Frecuencia de rocío en los EE.UU. De una red de radiómetros in situ*". se enfoca en la producción y duración del rocío en diferentes ecosistemas, y cómo esto puede afectar el equilibrio hídrico de los mismos. El estudio detalla cómo la humedad relativa,

la temperatura, la velocidad del viento y otros factores influyen en la producción de rocío en pastizales y bosques. Se discuten diferentes modelos utilizados para estimar la producción de rocío, incluyendo modelos semiempíricos y modelos basados en la teoría física, como la ecuación de Penman-Monteith. También se mencionan los desafíos de medir la producción de rocío en diferentes ecosistemas y se discuten las limitaciones de los modelos utilizados. En general, el archivo proporciona información detallada sobre la producción de rocío en diferentes ecosistemas y cómo esto puede tener un impacto en el equilibrio hídrico de los mismos [15].

Tras describir cada literatura, se realizó un resumen, el cual podemos divisar en la siguiente tabla denominada “Estudio referente a distintas literaturas del ámbito del rocío”.

Estudio referente a distintas literaturas del ámbito del rocío					
Artículos	Eventos	Entradas	Modelos	Métodos	Resultado
N. Sharma, P. Sharma, D. Irwin, P. Shenoy [19] (2011) “ <i>Predicción de la generación solar a partir del clima: Pronóstico mediante aprendizaje automático</i> ” .	Energía solar	Nubosidad, Humedad Relativa, Precipitación, Temperatura, Punto de rocío, Velocidad.	Modelo RNA SVM	Pruebas del modelo mediante Regresión lineal mínimo cuadrado.	Los resultados muestran que los modelos de predicción basados en SVM construidos a partir de siete métricas de pronóstico del tiempo distintas son un 27 % más precisos para nuestro sitio que los modelos existentes.
Gyanesh S, Sanjeev K, Manoj K Kowar, Pulak G [12] (2012) “ <i>Aplicación de redes neuronales artificiales en la predicción del tiempo y una revisión exhaustiva de la literatura</i> ”.	Meteorológicos	Meteorológicos	Modelo de RNA de RBFN y BPN.	Descripción y resultados con base en estimación.	Basándose en los resultados favorables, se nos facilita el uso de estas metodologías o tipo de modelos, ya que podría ser aplicado en predicciones meteorológicas.
D. Beysens, I. Lekouch, M. Muselli, A. Mongruel y B. Kabbachi [11](2012). “ <i>Recolección de rocío, niebla y lluvia en azoteas en el suroeste de Marruecos y modelado estimativos de rocío usando redes neuronales</i> ”	Rocío, Niebla, lluvia.	Temperatura, Humedad Relativa, Velocidad, Nubosidad.	RNA MLP.	Descripción de prototipos de condensadores, registro, además de datos meteorológicos. Arquitectura, entradas y salida de las RNA.	Se desarrolla el modelo MLP en la zona del suroeste de Marruecos, en donde se realizó el Acumulado diario del modelo de la RNA, obteniendo correlaciones de R superiores 0.9.
D. Beysens [10] (2016) ” <i>Estimación de la producción de rocío en todo el mundo a partir de unos pocos datos meteorológicos</i> ”	Rocío	Cobertura de nubes, Velocidad del viento, Temperatura aire, Temperatura del rocío, Humedad Relativa.	Fórmula de Beysens	Se generan pruebas del modelo matemático en correlaciones de regresiones lineales respecto a los resultados calculados.	Se desarrolla una comparación entre valores experimentales de medición y calculados, mediante el modelo de la correlación día a día proveniente de la fórmula de Beysens y el modelo acumulativo.
D. Carvajal, J. Minonzio, E. Casanga, J. Muñoz, A. Aravena, S. Montecinos y D. Beysens (2018) [13] ” <i>Recolección de agua de rocío integrada en el techo en Combarbalá Chile</i> ”	Rocío	Humedad relativa, Temperatura del aire, Velocidad del viento, Nubosidad, Presión atmosférica,	No expone modelos empíricos basados en mediciones realizadas en el experimento específico en Combarbalá.	La metodología implementa la instalación y lugar de ubicación, medición dia a dia.	Los resultados muestran que a pesar de la baja humedad relativa nocturna durante todo el año es de un promedio de 48 porciento La recolección del rocío más estable alcanza valores máximos de 0,44 mm ( $L \text{ m}^{-2} \text{ night}^{-1}$ ) en la noche.

D. Beysens, M. Berkelhamer, F. Ritter (2018) [15]. "Frecuencia de rocío en los EE.UU. De una red de radiómetros in situ".	Rocío	Temperatura, Humedad relativa, Presión, Velocidad del viento, Presión.	Modelo de los modelos empíricos y modelos basados en un balance de energía.	Condensador radiativo pasivo (RDC); Se menciona que el RDC suele ser una placa plana inclinada, aislada térmicamente del suelo, de la cual se recolectan gotas de rocío bajo flujo de gravedad.	El documento discute la importancia de comprender la formación de rocío en los ecosistemas naturales y cómo las redes neuronales pueden ser utilizadas para analizar y predecir este fenómeno.
---	-------	--	---	---	--

Tabla 2.6: Tabla de resumen de literatura

Estos artículos evidencian las teorías propuestas a la recolección del fenómeno de agua de rocío, uno de los autores famosos y de aporte reiterativo en la literatura en función al fenómeno del rocío, es el nombrado científico D. Beysens, el cual propone la fórmula de predicción de la recolección de agua y la recolección total basándonos en el acumulado de cada día Beysens. En función a las literaturas estudiadas, La RNA más representativa que se logra destacar es la que entrena con el entrenamiento del algoritmo levenger-Marquardt (LM) y la característica principal es que contiene múltiple capas y logra estimar correctamente este tipo de fenómeno.

# **Capítulo 3**

## **Definición del Problema**

### **3.1. Formulación del Problema**

El problema de investigación se enfoca en la modelización y cuantificación de los volúmenes de agua de rocío que se pueden obtener a partir de la colección de los datos meteorológicos obtenidos desde un recolector de rocío en el proyecto Fondecyt de Iniciación N.º 11140863 (2014), corresponde a un proyecto de investigación financiado por el Fondo Nacional de Desarrollo Científico y Tecnológico de Chile. Se destaca la importancia de analizar estos datos y mejorar la precisión de las estimaciones utilizando técnicas automáticas y herramientas informáticas. En particular, se busca comparar los modelos de estimación basados en redes neuronales artificiales (RNA) [11] y el modelo propuesto por D. Beysens en este contexto [10]. La investigación se basa en un enfoque cuantitativo y experimental.

El uso de (RNA) para estimar la recolección de agua de rocío se justifica por varias razones fundamentales. Los modelos teóricos y semi-teóricos presentan limitaciones en términos de precisión, especialmente cuando se trata de períodos de tiempo cortos, como un día o incluso una hora. Estas limitaciones están relacionadas con la incapacidad de los modelos teóricos para capturar relaciones altamente no lineales entre múltiples variables meteorológicas que influyen en la recolección de agua de rocío. Las RNA ofrecen una ventaja significativa al poder relacionar de manera efectiva múltiples variables y manejar grandes cantidades de datos, permitiendo la identificación de relaciones complejas que los modelos teóricos no pueden captar. Esta capacidad es crucial para mejorar la precisión de las estimaciones de recolección de agua de rocío, especialmente en intervalos de tiempo más cortos. Es importante mencionar que las mediciones se realizarán en horarios específicos, como promedio diario “PD” y a las “6 AM”, para respaldar la recopilación de datos y la evaluación de los modelos. Sin embargo, también es esencial abordar el desafío de la extrapolabilidad geográfica, ya que la aplicación de modelos de RNA en un área geográfica específica puede no ser directamente transferible a otras ubicaciones debido a las variaciones geográficas y climáticas. Por lo tanto, la elección de utilizar RNA en este contexto se basa en

la necesidad de mejorar la precisión de las estimaciones de agua de rocío y en la capacidad de las RNA para abordar relaciones complejas y no lineales entre las variables meteorológicas, mientras se considera cuidadosamente la extrapolabilidad geográfica para garantizar su aplicabilidad en la ubicación de Combarbalá, Chile.

### **3.1.1. Hipótesis**

Al emplear redes neuronales artificiales con los mismos conjuntos de datos que el modelo desarrollado por Daniel Beysens, anticipamos lograr un mayor desempeño en la tarea de predecir la recolección de agua de Rocío. Nuestra expectativa es que luego de llevar a cabo 10 pruebas secuenciales, los resultados generados por el modelo basado en RNA exhiban una correlación superior en comparación a los resultados obtenidos por medio del modelos de D. Beysens.

### **3.1.2. Preguntas de Hipótesis**

- ¿Los modelos poseen una mayor predicción de promedio por día o por hora?
- ¿El resultado acumulado de la correlación  $R^2$  de la RNA, es mayor que el propuesto por D. Beysens?
- ¿El resultado del  $R^2$  del modelo de la RNA, es mayor que el propuesto por D. Beysens?
- ¿Cuál de las RNA obtiene mayores resultados de  $R^2$ ?
- ¿Existe una diferencia significativa en la precisión de la predicción de la recolección de agua de rocío entre los modelos de redes neuronales artificiales y el modelo propuesto por D. Beysens?
- ¿Los modelos basados en redes neuronales artificiales son más estables en términos de predicción de la recolección de agua de rocío que el modelo propuesto por D. Beysens?

## **3.2. Solución Propuesta**

La opción propuesta es de forma parcial o suplementaria en lugares que padecen de escasez hídrica, tal y como es establecida en el punto anterior 3.1. Se podrá aplicar la opción de usar los datos de Combarbalá del proyecto FONDECYT 11140863 ejecutado por la Universidad de La Serena y bajo la dirección del profesor Danilo Carvajal, para implantar nuevos modelos estimativos para este fenómeno con el enfoque informáticos en la ubicación de Combarbalá mediante el proyecto mencionado anteriormente. El lenguaje

de programación es Python de data análisis y Machine Learning, Python nos permite obtener una curva de aprendizaje que es rápida, ya que analiza gran cantidad de datos, además visualizar resultados, desarrollar gráficos, automatización y modelos matemáticos y, por otro lado, Machine Learning que, a través de algoritmos, dota a los ordenadores de la capacidad de identificar patrones en datos masivos y elaborar predicciones (análisis predictivo).[6]. De acuerdo a la literatura desarrollaremos los modelos matemáticos y de resultados acumulativos, precisamente con los datos del proyecto FONDECYT de número 11140863 del año 2016 en el lugar de Combarbalá, a cargo del profesor D. Carbajal. Se emplean 10 experimentos obteniendo resultados de estos para analizar, además de debatir de acuerdo a los gráficos y tablas, que determinan las desviaciones estándar y media aritmética de los resultados  $R^2$ ,  $m$ ,  $c$ ,  $SE$ . Posteriormente, se buscará realizar una tabla de resumen y conclusión de resultados para encontrar el modelo idóneo.

### **3.2.1. Importancia e Impacto de la solución**

El impacto económico hace referencia a la recolección del rocío, ya que para captarlo no se necesitan una gran cantidad de capital, abarata costos propios tales como la compra de agua, puesto que el agua recaudada puede ser utilizada en diversos fines como; uso humano, ganadería, higiene, uso ambiental, entre otros usos.

La importancia social que entrega la elaboración del sistema de captación de rocío, es la generación de nuevas oportunidades para adquirir agua, que es una fuente vital para el ser humano, más en lugares donde el acceso al agua es restringido, un ejemplo es la zona árida o semiárida de Petorca al interior de la V región, donde los camiones aljibes reparten cierta cantidad de agua a la semana por persona, limitando las actividades humanas tales como; La higiene, la elaboración de alimento, el consumo humano, riego, etc. Reduciendo la calidad de vida de las personas.

La importancia ambiental está relacionada con la generación de alimento para el consumo de animales, evitando la mortandad que azota las zonas áridas o semiáridas, arrasando con actividades típicas locales como el pastoreo, o el denominado arreo de ganado, la elaboración de productos lácteos de forma artesanal por ejemplo quesos, leches, entre otras actividades. Por otro lado, también se puede implementar para el riego controlado de cultivos (riego por goteo). Lo que permite concluir que la recolección del agua de rocío genera un aporte de forma parcial o complementaria. Por medio de este método, en lugares con escasez hídrica, el agua se puede utilizar en riego, alimento para animales, limpieza y por último el uso doméstico.

Por último, la importancia informática está relacionada con los pocos trabajos realizados en el área investigativa con datos experimentales de Combarbalá basándonos en los modelos estimativos por medio de modelos de RNA, que será una guía para los futuros investigadores.

## **3.3. Objetivos**

### **3.3.1. Objetivo General**

El objetivo principal de esta tesis es comparar el modelo de D. Beysens con una Red Neuronal Artificial (RNA) mediante técnicas de aprendizaje automático, con el propósito de estimar de manera precisa el flujo de agua de rocío en el norte de Chile.

### **3.3.2. Objetivos Específicos**

- Desarrollar los modelos matemáticos estimativos con datos de horario 6 AM y de datos promedio diario (PD).
- Implementar la correlación diaria y la acumulación de ambos modelos en función de datos experimentales además de los datos calculados.
- Guardar modelos y resultados.
- Analizar y comparar los resultados de los modelos estimativos.

## **3.4. Metodología**

### **3.4.1. Enfoque de investigación**

El enfoque de investigación del proyecto será de tipo cuantitativo, ya que evaluara distintas variables numéricas como son: la temperatura, humedad relativa, nubosidad, velocidad del viento y recolección de rocío. A partir de estas variables se realizará el análisis de resultados de la predicción de los modelos efectuados para determinar el que mejor se desempeña en la zona de estudio.

El tipo de investigación para realizar el proyecto será de forma experimental, debido a que nuestras variables o características son parámetros ambientales, esto nos permitirá emplear nuestros modelos matemáticos de predicción de recolección de rocío, permitiendo analizar los resultados aplicados con el fin de la obtención de la correlación de cada modelo estimativo. En la metodología experimental de este trabajo de investigación, se ha optado por utilizar el coeficiente de determinación ( $R^2$ ) en lugar del Error Cuadrático Medio (MSE) como la métrica principal de evaluación. Esto se debe a que el modelo de Beysens no se calcula como una Red Neuronal Artificial (RNA), en la cual el MSE es una métrica comúnmente utilizada. En el caso del modelo de Beysens y su aplicación en un enfoque experimental,  $R^2$  se elige como la métrica principal debido a su capacidad para medir la calidad del ajuste del modelo a los datos y cuánta variabilidad del resultado es explicada por ese modelo.  $R^2$  también se interpreta como el cuadrado del coeficiente de correlación ( $r$ ), lo

que permite evaluar la fuerza y la dirección de la relación entre las variables en el modelo. La elección de  $R^2$  se justifica en función de su relevancia para la naturaleza del modelos de D. Beysens y la RNA.

### 3.4.2. Fases de investigación

Las fases de investigación se realizaron de acuerdo a la metodología, las cuales son de tipo cascada, estas consisten en desarrollar el proyecto de origen investigativo de forma secuencial y sincrónica en proporción al tiempo del trabajo de título, hasta obedecer los objetivos específicos de cada fase propuestas en la investigación.

Para cumplir y desarrollar nuestro propósito investigativo se generarán 6 etapas en forma de cascada y ágil, tal como se puede visualizar en la figura 3.1, las etapas son; Importación de bases de datos, medición de datos 6 AM, limpiezas y separación de datos, modelos “Beysens y RNA”, Correlación ”Día a Día y acumulado”, por último se generara un análisis y discusión de resultados.

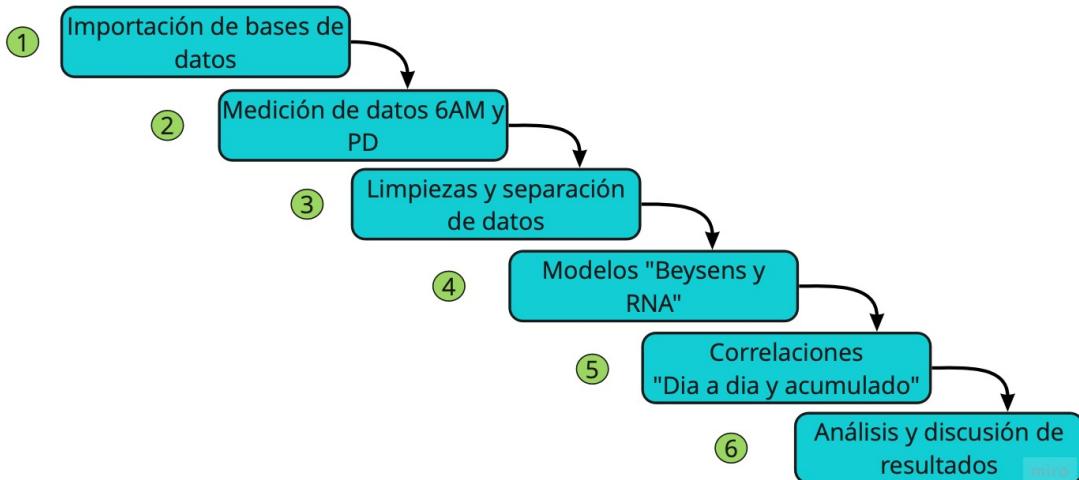


Figura 3.1: Diagrama de Metodología del proyecto

A continuación describiremos minuciosamente cada etapa del diagrama de la investigación:

1. La primera etapa corresponde a la importación de bases de datos, las cuales son obtenidas desde la ubicación de Combarbalá por medio del proyecto FONDECYT de número 11140863 dirigido por D. Carvajal, donde los datos se recolectan de acuerdo a dos formas, la primera es extraída de datos experimentales exactamente

de la ubicación de la estación de Combarbalá, los parámetros son; La temperatura, la velocidad del viento, la humedad relativa y el flujo del rocío. La segunda es de datos Geo espacial de la página Woldview de la NASA, donde fueron obtenidos por medio de imágenes capturadas por el satélite Terra de noche, el cual utilizaron estas capturas para transformarlas en una cantidad numérica de nubosidad de 0 a 100 por ciento. Para terminar, todos los datos son guardados en dos data sets en formato CSV de Excel, la primera se consideran datos promedio diario y la segunda bases de datos en bruto, como se indica en la siguiente Figura 3.2.

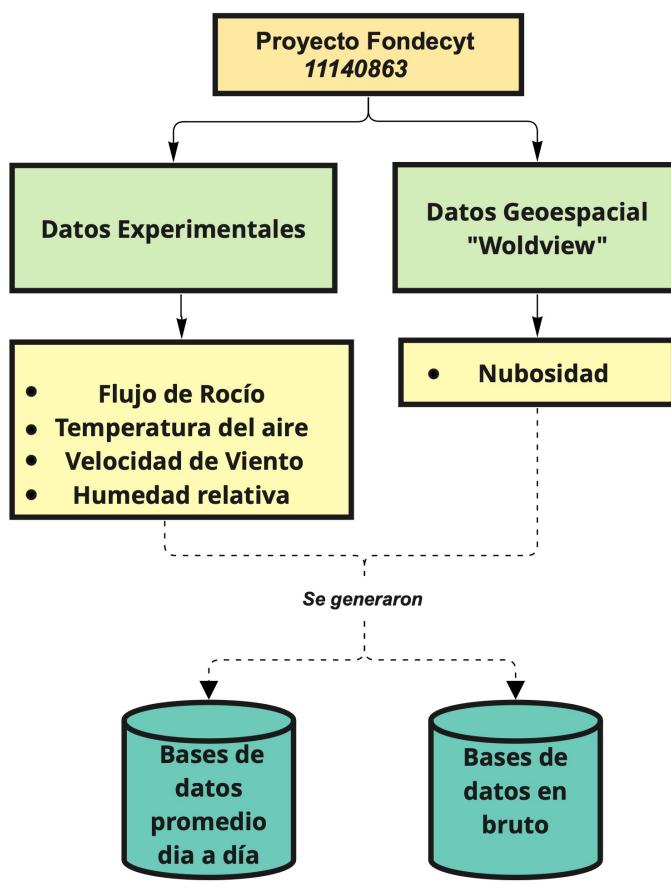


Figura 3.2: Diagrama de obtención de datos del proyecto FONDECYT de número 11140863.

2. En la segunda etapa de la metodología del proyecto, se obtienen datos de un periodo de 2 años, con alrededor de 500 días de medición de datos PD y brutos, los cuales filtraremos a las 6 AM, el desarrollo de estas etapas se llevara a cabo por medio de siguiente diagrama 3.3. La descripción del diagrama inicia con el filtro de datos brutos

en horarios de las 6 AM, en consideración de los factores climáticos tales como: Temperatura, Humedad Relativa, Velocidad del Viento y Nubosidad. Por medio de estos datos a partir de modelos matemáticos podemos estimar la recolección del flujo de agua de rocío. Luego en primera ocasión filtramos los datos desde el horario de las 12 PM hasta las 6 AM, ya que en este horario obtenemos el flujo de agua recolectada de rocío de toda la noche, en un segundo lugar se realiza un cambio de unidad de medida de mm a mL con el fin de obtener el agua acumulada diaria, para aquello la variable anteriormente es multiplicada por 1.6 y dividido en 0.3 como se visualiza en la siguiente fórmula.

$$\text{agua\_colectada\_acumulada [mL]} = \text{total\_rainfall [mm]} \times \frac{1.6 \text{ mL}}{0.3 \text{ mm}} \quad (3.1)$$

En tercer lugar, se persigue generar los datos instantáneos, son en horarios de las 6 AM mencionados por el profesor D. Carvajal, mediante la unidad de 'mL', correspondiente a la al día actual menos el día anterior, tal como se describe en la fórmula siguiente.

$$\text{Valor\_instantáneo [mL]} = \text{Acumulado\_día\_siguiente [mL]} - \text{Acumulado\_día\_anterior [mL]} \quad (3.2)$$

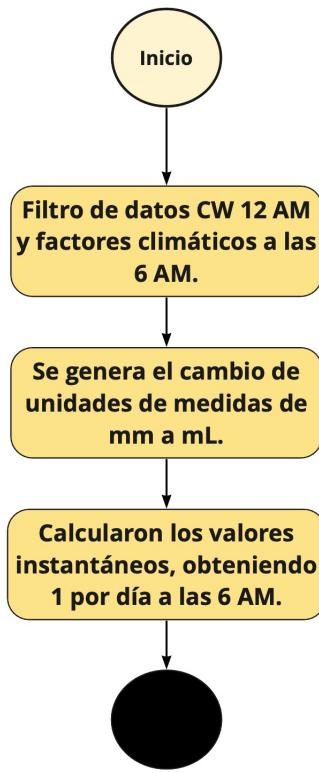


Figura 3.3: Diagrama de generación, datos instantáneos.

3. La tercera etapa corresponde a la limpieza de datos, el cual implica la eliminación de las filas que contengan datos en NaN, String o que sea vacío. Una vez realizado este procedimiento esperamos obtener datos de calidad que representen el fenómeno de rocío. Para cada modelo, se efectuó un equilibrio de datos el cual tiene como objetivo dividir en 3 partes equivalentes de 1/3 de datos de recolección de agua en; rocío pequeño, medio y alto. Tras finalizar el equilibrio de datos de cada sección se desarrolló una separación de datos de forma aleatoria, el cual fue sugerido en el documento de Lekouch, que corresponde a 2/3 de datos de entrenamiento, son los datos que ingresan por primera circunstancia, como entrada a la RNA, La segunda separación corresponde a 1/3 restante denominado con el nombre de pruebas, el cual valida nuestra RNA y el modelo de Beysens, tal y como se visualiza en la siguiente figura 3.4.

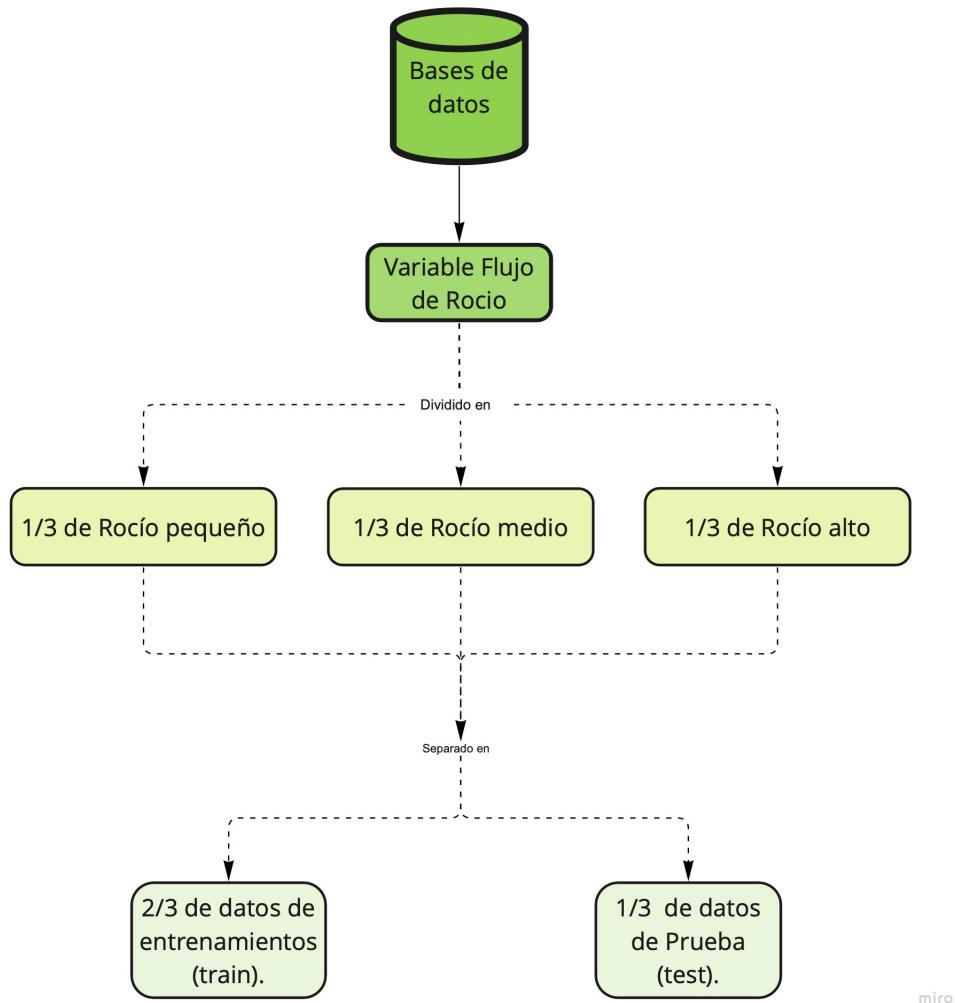


Figura 3.4: Diagrama de generación, datos instantáneos, elaboración propia.

4. La cuarta etapa se refiere a la discusión de los modelos matemáticos los cuales son extraídos desde la literatura, estos logran estimar, a partir de los parámetros ambientales, para determinar el flujo de agua de rocío.

El investigador D. Beysens en el año 2016 formuló el modelo matemático que logra predecir correctamente, al igual que el modelo de RNA sugerido en el documento de Lekouch en el año 2012. En ambos modelos, luego de estimar las correlaciones diarias, se realiza una medición de acumulado de cada resultado de correlación de cada modelo con el fin de interpretar y analizar el modelo idóneo en función a los datos obtenidos de la localidad de Combarbalá.

A continuación se describirán minuciosamente, los modelos mencionados, estos son:

- a) **Modelo de la Red Neuronal Artificial(RNA):** Este modelo es extraído desde la literatura, específicamente en el documento de Lekouch, publicado en el año 2012. En este documento se consideran las siguientes características: la estructura de la RNA, división de datos de entrenamientos y pruebas, ambos de forma aleatoria, por último utilizo el algoritmo de entrenamiento Levenberg Marquardt [11].

El modelo tiene como objetivo encontrar una relación adecuada entre los datos meteorológicos diarios (entradas independientes) y el flujo de rocío diario, tal cual como se puede divisar en la siguiente figura 3.5.

La RNA contiene una capa de entrada con 4 neuronas, una única capa oculta con 5 neuronas ocultas con la función de activación sigmoidea tangente y por último 1 capa de salida, con función de activación log-sigmoidea, de acuerdo a esta estructura de capas podemos estimar la cantidad de agua de rocío.

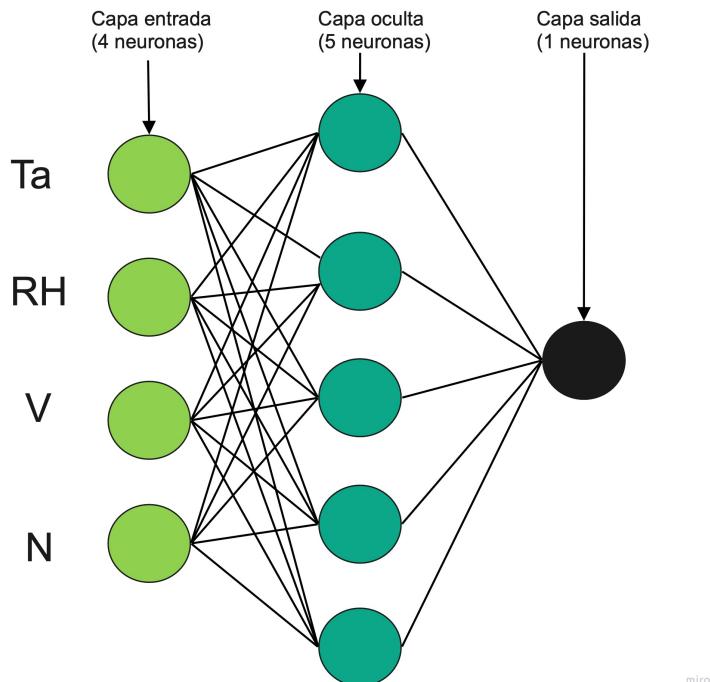


Figura 3.5: Diagrama de arquitectura RNA “MLP”.

El modelo de la red neuronal contienen 4 parámetros de entrada, los cuales son: La temperatura del aire, humedad relativa, velocidad de viento y por último la nubosidad. Estas variables de entradas son calculadas inicialmente por nuestro modelo, asociados a épocas que son el número de entrenamientos provocado a nuestro modelo, ya que nos entrega las estimaciones como resultado de salida, en simples palabras es la recolección de flujo del rocío, tal y como se puede apreciar en la siguiente figura 3.6.

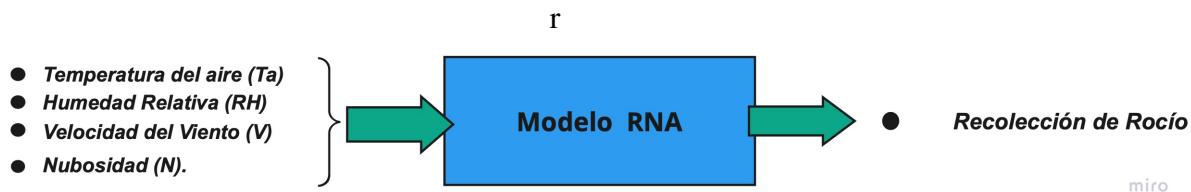


Figura 3.6: Diagramas de modelo de la RNA.

b) **Modelo matemático de Beysens:** Este modelo es propuesto y obtenido con base en la literatura empleada por Daniel Beysens en el año 2016, desarrollado para predecir la recolección de agua de rocío en función de parámetros ambientales como entrada [10].

La altura es de 0.923 kilómetros, por lo que será despreciable; Sin embargo, la fórmula contienen 4 parámetros importantes de entrada, estos son: Temperatura aire, Temperatura de rocío, Viento y Nubosidad, donde su variable de salida es el flujo del rocío tal como se visualiza en la siguiente figura 3.7, en el caso de obtener predicciones negativas serán igual a 0.



Figura 3.7: Diagrama de modelo matemático realizado por D. Beysens

Una vez implementados ambos modelos matemáticos estimativos presentados anteriormente en el punto a y b. El fenómeno del rocío, se considera tanto en el documento de I. Lekouch [11], como en el de D. Beysens [10], donde sugieren una acumulación luego de los resultados correlaciones calculados por el modelo de datos experimentales, como se presentan en las siguientes sumatorias.

$$\sum_{\text{días}=i}^{\text{día final}} \left( \left( \frac{dh}{dt} \right)_{\text{exp}} \right) \quad \sum_{\text{días}=i}^{\text{día final}} \left( \left( \frac{dh}{dt} \right)_{\text{calc}} \right) \quad (3.3)$$

5. La quinta etapa corresponde a las correlaciones obtenidas de los resultados de los modelos matemáticos, la correlación indica la fuerza y dirección de la relación lineal, además del proporcional entre dos variables estadísticas, en el caso de la teoría se efectúan en los datos de entrenamiento así como también en las pruebas.

En función a la experimentación, luego de obtener los resultados, establecimos 4 parámetros de acuerdo a la correlación lineal y ecuación de la recta que se puede divisar en la siguiente fórmula.

$$y = mx + b \quad (3.4)$$

A continuación se describirá cada parámetro obtenido basándonos en la correlación y acumulaciones, estos son;

- a)  $R^2$  : Es una medida estadística que representa la proporción de la varianza de una variable dependiente que se explica por una o varias variables independientes en un modelo de regresión o correlación.
  - b)  $m$  : Pendiente, representada por la variable ‘m’ en la fórmula, indica el grado de inclinación respecto a la horizontal.
  - c)  $b$  : Este coeficiente de posición se representa con la letra ‘b’ indicando donde se interceptará al eje y.
  - d)  $SE$  : Error Estándar (*Standard Error*) La regresión es el valor que muestra la diferencia entre los valores reales y los estimados de una regresión. Es utilizado para valorar si existe una correlación entre la regresión y los valores medidos obteniendo el intervalo de confianza de la medida.
6. La última etapa seis corresponde al plan de trabajo efectuado en la metodología, donde se realizarán las tablas de visualizaciones de los modelos de RNA y D. Beysens para establecer diferencias y discusión de los resultados.

Las primeras tablas 3.1 y 3.2 se desarrollan para comparar los modelos matemáticos en función de los datos diarios y Acumulados, el objetivo principal es visualizar y comparar los comportamientos del modelo con datos de entrenamiento y de validación.

Día	Entrenamiento	Validación
RNA	PD	PD
	6 AM	6 AM
Beysens	PD	PD
	6 AM	6 AM

Tabla 3.1: Resultados Día a día

Acumulados	Entrenamientos	Validación
RNA	PD	PD
	6 AM	6 AM
Beysens	PD	PD
	6 AM	6 AM

Tabla 3.2: Resultados Acumulados

En las posteriores tablas enmarcamos los modelos desarrollados de Validación o testing que solo es el 1/3 de los datos, se realiza esta implementación para visualizar el desempeño de correlación en horarios 6 AM y PD de los modelos de D. Beysens además de las RNA calculadas en datos diarios y acumulado, estos gráficos serán presentados en las siguientes tablas 3.3 para explicar las siguientes correlaciones.

PD / 6 AM	Beysens	RNA
Día	Validación	Validación
Acumulados	Validación	Validación

Tabla 3.3: Resultados PD y 6 AM

Por último, la discusión de resultados se realizó a partir de una tabla de resumen de los parámetros vistos en el punto 5 “metodología”, que principalmente son las siguientes variables:  $R^2$ ,  $m$ ,  $b$  y  $SE$ . A través de estos parámetros, se podrá deducir cuál modelo y en qué horario es más adecuado: PD, 6 AM, Día a día o Acumulado. Posteriormente, mostraré cómo se implementará la tabla 3.4 de “resumen de resultados” en modelos de RNA y el modelo de D. Beysens. Los resultados son referentes al ingreso de las variables mencionadas, donde los resultados se aprovechan para calcular la media aritmética y la desviación estándar. Tras estos cálculos, se representarán en la tabla de resumen de resultados, lo que nos permitirá diferenciar y analizar el modelo predominante.

PD	Día	Beysens
		RNA
Acumulados	Beysens	
		RNA
6 AM	Día	Beysens
		RNA
Acumulados	Beysens	
		RNA

Tabla 3.4: Resumen de resultados de los casos de estudio.

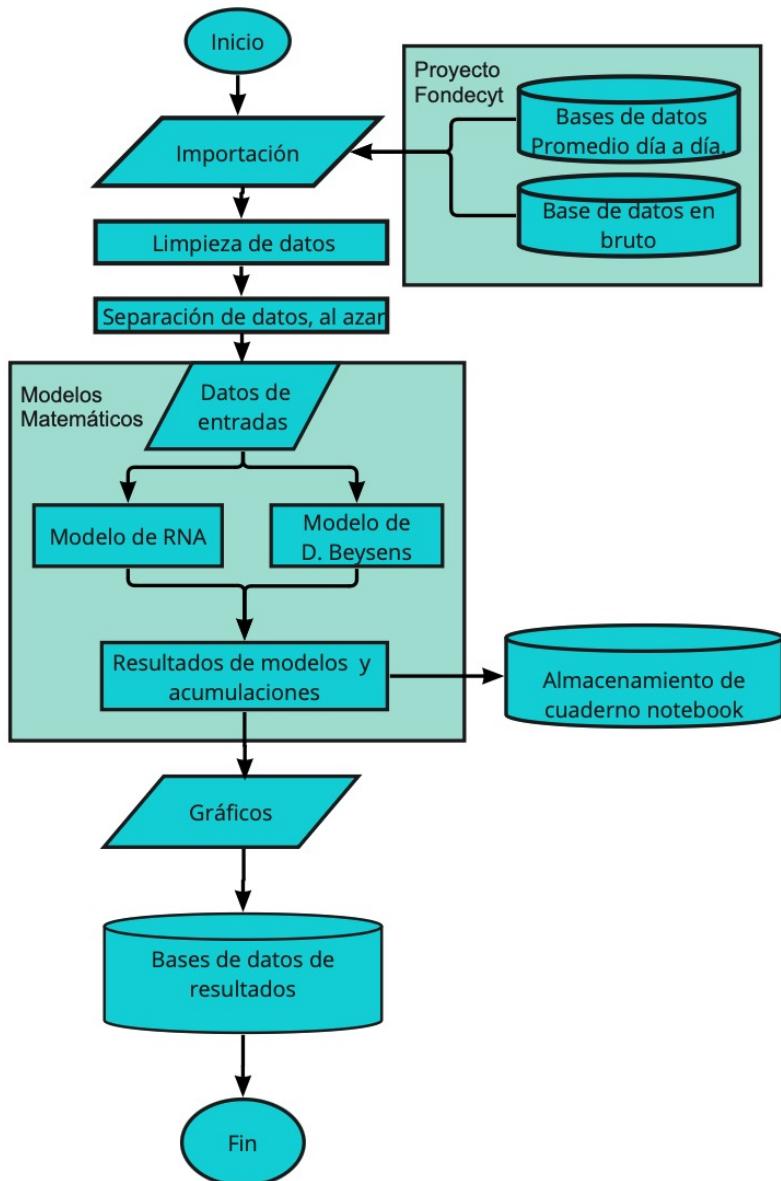
# **Capítulo 4**

## **Diseño de la solución**

### **4.1. Solución del problema**

La solución del problema es realizada a partir de la formación de la problemática, los objetivos específicos, además del objetivo general y la metodología, esto nos permite diseñar una solución propuesta de acuerdo a la hipótesis que se especifica en la construcción de ambos modelos matemáticos estimativos. El desarrollo de esta investigación será mediante el lenguaje de programación Python, con el software de código abierto Jupyter Notebook y por último Docker que nos ayudará a emplear el modelo arquitectónico.

El diagrama de solución es equivalente a una prueba como en la siguiente Figura 4.1, esta inicia con la importación de las bases de datos del proyecto FONDECYT 11140863 ejecutado por la Universidad de La Serena y bajo la dirección del profesor Danilo Carvaljal, luego se genera una limpieza de datos NaN y vacíos, antes de ingresar a los modelos matemáticos los datos se dividen en entrenamiento (2/3) y en entrenamiento (1/3), posteriormente los datos de entrada o input ingresan a cada modelo, los cuales tras procesos matemáticos generan una salida única que será la estimación del flujo recolectado de rocío. Basándose en estos datos podremos visualizar las correlaciones así como también las acumulaciones, tras este paso almacenamos en GitHub toda la prueba ejecutada. Al finalizar se efectuaron diez iteraciones de pruebas de los modelos para obtener la media aritmética y desviación estándar del conjunto de pruebas, permitiendo generar una bases de datos de los resultados a partir los parámetros como;  $R^2$ ,  $m$ ,  $b$ ,  $SE$ . Estos parámetros los utilizaremos para analizar, discutir y comprobar el experimento realizado por una tabla de resumen de resultados.



miro

Figura 4.1: Diagrama de la solución del problema

## 4.2. Técnicas de análisis

De acuerdo a los resultados de las correlaciones de los modelos y la ecuación de la recta generada, aplicamos las técnicas de análisis que corresponden a la media aritmética y desviación estándar que a continuación describiremos.

- **Media aritmética:** La media aritmética se utiliza para determinar la tendencia general de un conjunto de datos. Consiste en sumar todos los valores en una lista y luego dividir la suma por el número de elementos en esa lista. La media aritmética proporciona una instantánea rápida de la información y ayuda a obtener una idea general de los datos.
- **Desviación estándar:** La desviación estándar se utiliza para determinar la dispersión de los puntos de datos. Una desviación estándar alta indica que los datos se distribuyen más ampliamente desde la media, lo que implica una mayor variabilidad. Por otro lado, una desviación estándar baja indica que los datos se alinean más cerca de la media, lo que implica una menor variabilidad. La desviación estándar ayuda a comprender la dispersión de los datos y a identificar si hay valores atípicos que puedan afectar la interpretación de los resultados.
- **Análisis de regresión:**

Esta técnica se utiliza para evaluar la relación entre una variable dependiente y una o más variables independientes. En el contexto del trabajo de título, puedes aplicar el análisis de regresión para examinar cómo las variables de entrada (como las condiciones ambientales) afectan la recolección de agua de rocío.

# Capítulo 5

## Experimentación

### 5.1. Diseños de experimentos

→ Etapas de experimentación:

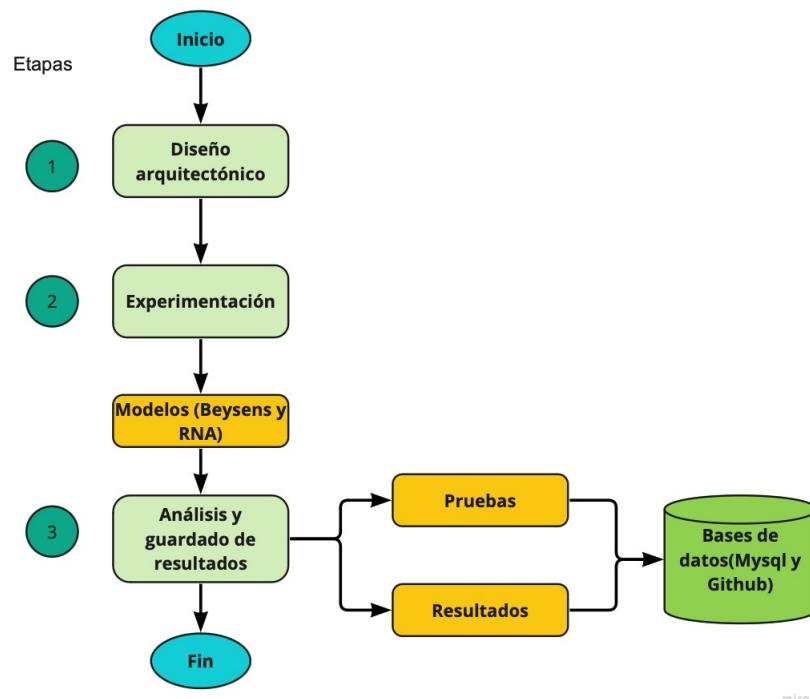


Figura 5.1: Etapas de la experimentación.

El diseño del experimento contiene tres etapas principalmente, como visualizamos en la figura anterior 5.1, en primer lugar es el diseño arquitectónico que representa la parte de infraestructura de nuestro trabajo de título, en segundo lugar la experimentación es de acuerdo a los modelos matemáticos experimentales, los cuales en la tercera etapa tiene como objetivos guardar las pruebas y resultados del experimento que posteriormente serán analizados.

- 1. Etapa del diseño arquitectónico** El diseño del diagrama 5.2, representa la comunicación cliente-servidor proporcionada por software gratuitos; sin embargo, esta parte fundamental nombrada infraestructura, obtenemos nuestro experimento de forma aislada en un sistema operativo que obtenemos implementación y ejecución de este, por último nos permite guardar los resultados en GitHub y en MySQL.

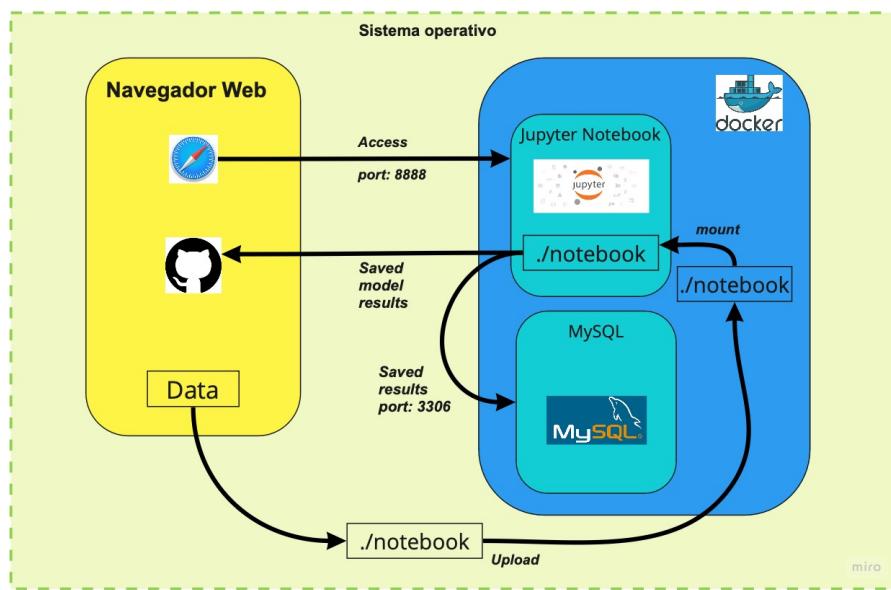


Figura 5.2: Diagrama cliente-servidor.

El funcionamiento del sistema arquitectónico ocurre cuando ejecutamos nuestro contenedor de Docker, el cual es un archivo Dockerfile, el cual genera una imagen aislada de Docker por medio el comando “Docker-Compose”, que nos facilitará la infraestructura de nuestro experimento, tal y como visualizamos en la siguiente figura 5.3.

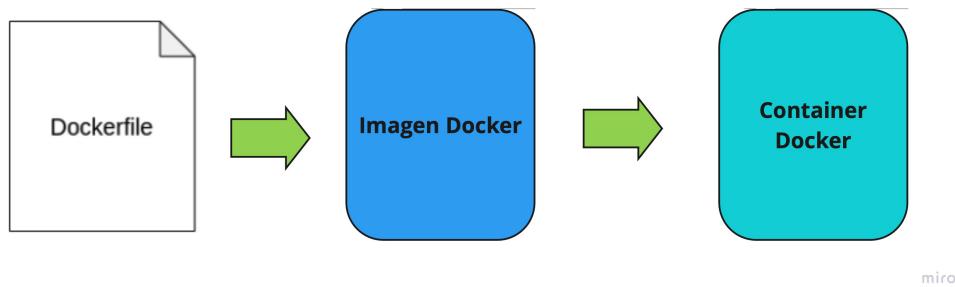


Figura 5.3: Funcionamiento del contenedor de Docker.

## 2. Etapa de Experimentación

¿Cómo realizamos los experimentos?

La etapa de experimentación se desarrolla a partir del diseño del experimento ejecutado en diez iteraciones denominado como pruebas. El diseño del experimento está basado por dos modelos de la literatura, los cuales coinciden en cuatro variables independientes o parámetros experimentales de entrada similares, para finalizar con la estimación del parámetro de salida de la predicción de recolección del flujo de rocío, tal como se puede divisar en la siguiente imagen del diseño del experimento 5.4.

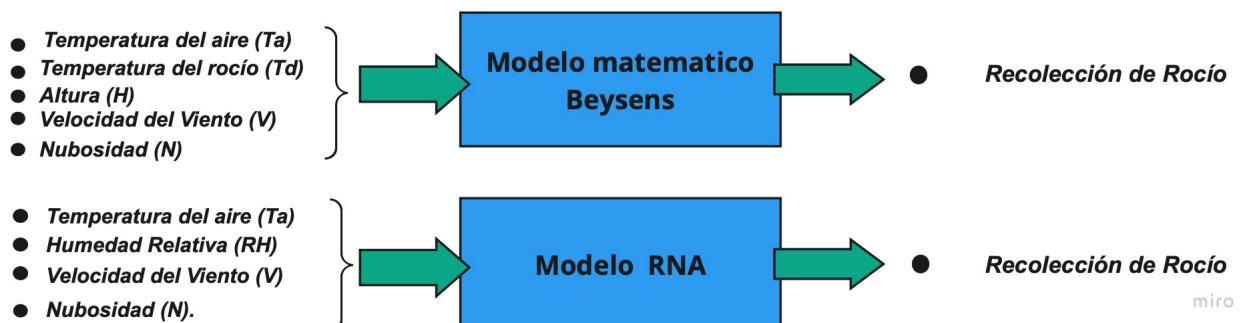


Figura 5.4: Diseño de la experimentación.

### 3. Etapa de análisis y guardado de resultados

Esta última etapa tiene el objetivo de registrar evidencias respecto a los resultados generados en los modelos, al igual que el entorno de trabajo ejecutado por medio de Jupyter Notebook designado como pruebas, con el propósito de guardar los experimentos y de dejar evidencia, además de analizar los resultados tras las diez experimentaciones consecutivas y con datos aleatorios entre experimentos.

- **Definición de variables:** Las variables a medir del modelo, son el ingreso y la salida del desarrollo de la experimentación, esto lo podemos ver registrado en la primera parte de las etapas de la metodología en la figura 3.2, puesto a los datos recolectados son propios de la zona Combarbalá proveniente de datos experimentales y geo espaciales, los cuales son recolectados a lo largo de dos años, proveniente del diseño de experimento el cual contiene parámetros experimentales de entrada y salida los cuales están representados en el diseño del experimento figura 5.4.
- **Variables de entrada:** Se trata de las variables experimentales independientes de ingreso del modelo, obtenidas desde Combarbalá, las cuales se detallarán a continuación:
    - Modelo RNA
      - 1. Temperatura del aire ( $T_a$ )
      - 2. Humedad relativa ( $HR$ )
      - 3. Nubosidad ( $N$ )
      - 4. Velocidad del viento ( $V$ )
    - Modelo D. Beysens
      - 1. Temperatura del aire ( $T_a$ )
      - 2. Temperatura del rocío ( $T_d$ )
      - 3. Nubosidad ( $N$ )
      - 4. Velocidad del viento ( $V$ )
      - 5. Altura ( $H$ )
  - **Variable de salida:** La variable experimental de salida de los modelos matemáticos, es el flujo recolectado de agua de rocío por toda la noche:
    - Modelo RNA y D. Beysens
      - 1. Volumen de agua recolectada de Rocío ( $dh/dt$ ).

## 5.2. Datasets y Casos de estudio

- **Descripción de datasets:** Los datos o parámetros experimentales meteorológicos son de acuerdo al clima, una zona semiárida, escasa de precipitación y vegetación, estos parámetros son la temperatura, la nubosidad, la humedad relativa, la velocidad del viento. Por último, el flujo de rocío colectado como se evidencia en la figura 5.5. Este flujo de agua de rocío es recolectado utilizado en el proyecto FONDECYT de N.º 11140863, el cual estaba a cargo del Profesor Danilo Carvajal, quien generó esta base de datos a lo largo del tiempo establecido del proyecto.



Figura 5.5: Colector de rocío, proyecto FONDECYT de N.º 11140863, situado en Combarbalá.

Es un fenómeno natural además de físico, se observa principalmente en la noche, el cual se logra apreciar en plantas, paredes, techos de casas y superficies heladas. Se considera la recolección de los datos en los meses de mayo del año 2016 hasta octubre del año 2017, por medio de sensores y pluviómetros, que nos permiten la captación de estos datos experimentales, precisamente en el lugar el Chañar, comuna Combarbalá, Región de Coquimbo, como se visualiza en la siguiente imagen 5.6 a  $31^{\circ} 7'57.54''S$   $71^{\circ} 7'53.26''O$  a 923 metros de altura a 50 km del Océano Pacífico.

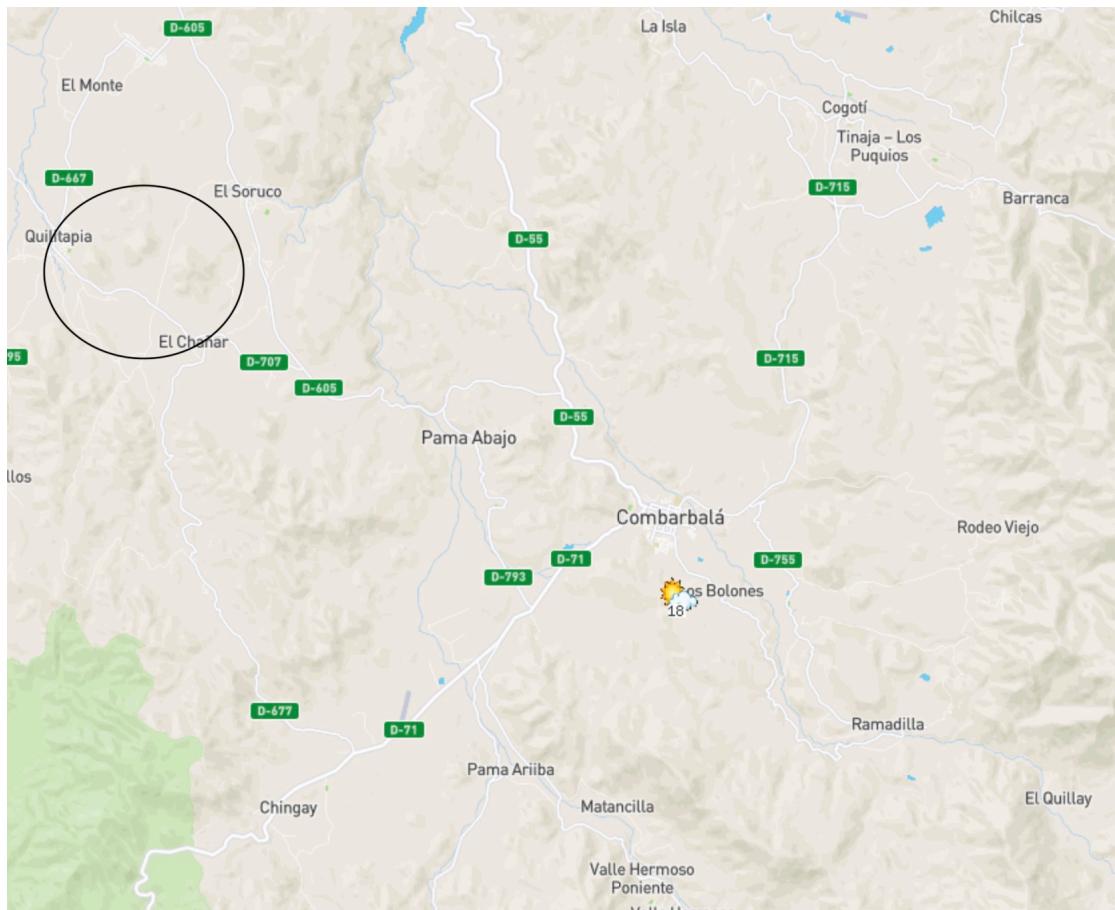


Figura 5.6: Imagen de la ubicación de Combarbalá, fuente de referencia [5].

Se logra obtener dos bases de datos. La primera corresponde a los datos promedio por día (PD), mientras que la segunda es una base de datos en bruto con intervalos de 15, 16 y 30 minutos. El almacenamiento se generó en Excel en formato CSV a través del instrumento denominado PLC, que se utiliza principalmente para recopilar datos del entorno. Está equipado con módulos de entrada y salida que se conectan a sensores ambientales. Se puede visualizar este dispositivo en la siguiente imagen 5.7.



PLC

Figura 5.7: PLC utilizado en el proyecto FONDECYT de N.<sup>o</sup> 11140863.

A continuación, mostraremos unos gráficos que representan la evolución de las variables medioambientales recopiladas en el proyecto a lo largo del tiempo, con un enfoque en los datos registrados a las 6 AM, como se evidencia en la imagen 5.8, y en horario PD. La diferencia se visualiza en la siguiente imagen 5.9.

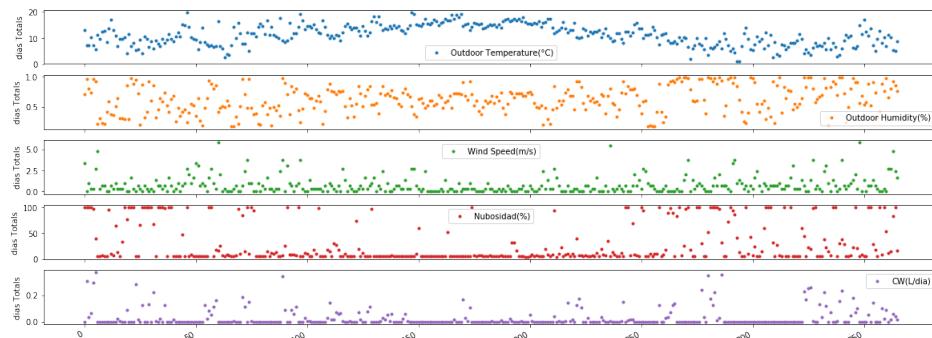


Figura 5.8: Parámetros promedio diario (PD), proyecto FONDECYT de N.<sup>o</sup> 11140863.

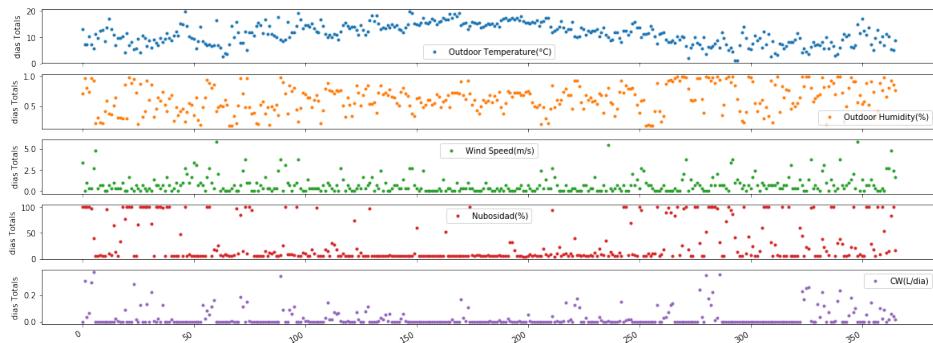


Figura 5.9: Parámetros en horarios (6 AM), proyecto FONDECYT de N.<sup>o</sup> 11140863.

#### → Descripción detallada de casos de estudio

La descripción de casos de estudio, fue mencionada en la tabla 3.4, en donde se consideran los datos que pueden ser en horario PD o 6 AM, los métodos hacen función a los modelos matemáticos utilizados en la literatura denominados modelos de D. Beysens y la RNA.

Datos	Métodos	Correlación	#
PD	Beysens	Día a día	1
		Acumulado	2
	RNA	Día a día	3
		Acumulado	4
6 AM	Beysens	Día a día	5
		Acumulado	6
	RNA	Día a día	7
		Acumulado	8

Tabla 5.1: Conclusión de modelos.

A continuación se presentan ocho casos de estudio relacionados con la estimación y correlación de datos horarios utilizando diferentes modelos matemáticos. A continuación, se detallará cada caso y los resultados obtenidos, así como las etapas de desarrollo y las comparaciones realizadas.

- **Caso 1:** Se consideran las bases de datos de horario PD como entrada para el modelo matemático propuesto por D. Beysens. El objetivo principal es obtener resultados de estimación y calcular la correlación entre los datos estimados y medidos.

- **Caso 2:** Se realiza un análisis acumulado de los resultados diarios obtenidos en el Caso 1. Se suman de forma acumulada los días estimados y medidos para obtener las correlaciones acumuladas del modelo de D. Beysens.
- **Caso 3:** Utilizamos el mismo conjunto de datos de horarios PD que en el Caso 1, pero se emplea el modelo de RNA conocido como "MLP" (Perceptrón Multicapa) para estimar y calcular la correlación diaria de los resultados obtenidos por la RNA.
- **Caso 4:** Realizamos un análisis acumulado similar al Caso 2, pero utilizando los resultados estimados y medidos obtenidos en el Caso 3. El objetivo es obtener las correlaciones acumuladas del modelo de RNA.
- **Caso 5:** Trabajamos con datos de horarios 6 AM, utilizando la misma fecha que en los Casos 1 y 3. Los datos de horario 6 AM ingresan al modelo matemático propuesto por D. Beysens, y los resultados estimados son a partir del proceso matemático de la RNA, Luego se calcula la correlación día a día.
- **Caso 6:** Utilizamos los resultados estimados obtenidos mediante la fórmula de Beysens en el Caso 5. Estos resultados se suman de forma acumulada a los resultados medidos, con el fin de obtener las correlaciones acumuladas del modelo de Beysens con horario 6 AM.
- **Caso 7:** Desarrollamos un análisis similar al Caso 4, pero utilizando los datos de horarios 6 AM que se utilizaron en el Caso 4. Se emplea el modelo de RNA y se obtiene la estimación como resultado, también como la correlación día a día.
- **Caso 8:** Continuamos tras finalizar el caso 7 utilizando los resultados estimados y medidos obtenidos. Se suman de forma acumulada y se calculan las correlaciones acumuladas de la RNA para el horario 6 AM.

En resumen, los casos de estudio presentados en esta investigación abarcan diferentes escenarios y enfoques utilizando datos de horarios PD y 6 AM, así como modelos matemáticos propuestos por D. Beysens y la RNA. Se realizó una comparación de resultados de estimación y correlación entre los datos estimados y medidos en cada caso.

Los resultados obtenidos en estos casos contribuyen a validar y evaluar los modelos matemáticos utilizados, brindando una mayor comprensión de la relación entre los datos estimados y medidos. Además, se pudo determinar la eficacia de los modelos en la estimación de datos horarios y calcular la correlación correspondiente.

En general, estos hallazgos proporcionan una base sólida para futuras investigaciones y aplicaciones prácticas en el campo de la estimación y correlación de datos horarios. Los resultados obtenidos pueden ser útiles para tomar decisiones informadas en diversas áreas que requieran análisis horarios y predicción.

Esta investigación se centra en el análisis comparativo de modelos matemáticos propuestos por D. Beysens y la Red Neuronal Artificial (RNA) para la estimación y correlación de datos horarios. A través de la generación de gráficos y la realización de cuatro etapas, cada una compuesta por dos casos de estudio, se llevó a cabo una exhaustiva comparación de la correlación de datos 6 AM o PD para cada método matemático utilizado y vistos en los casos de estudios anticipados de la tabla 5.2, que visualizaremos en la siguiente imagen 5.10.

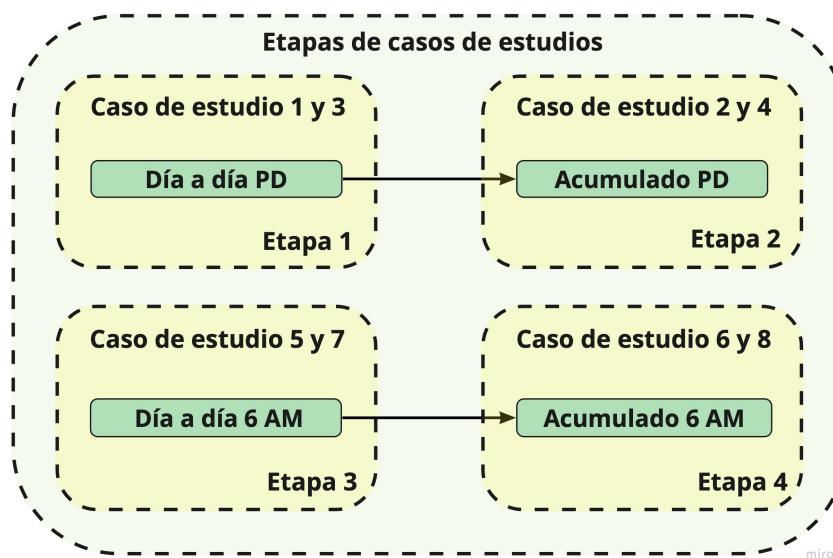


Figura 5.10: Etapas de casos de estudios.

La primera etapa se centró en los casos 1 y 2, donde se consideraron las bases de datos de horario PD como entrada para el modelo matemático de D. Beysens. Obtuvimos los resultados de estimación y calculamos las correlaciones entre los datos estimados y medidos. Posteriormente, en la segunda etapa, se analizaron los casos 2 y 4, empleando el modelo de RNA “MLP” para estimar y calcular la correlación diaria de los resultados utilizando datos de horarios PD.

El tercer paso fue para los casos 5 y 6, donde se trabajó con datos de horarios 6 AM. Aprovechando la misma fecha que en los casos anteriores, se ingresaron los datos de horario 6 AM al modelo matemático de Beysens, obteniendo resultados estimados mediante el proceso matemático de la RNA. Se calculó la correlación día a día y se realizaron análisis comparativos.

En la cuarta etapa, se abordaron los casos 7 y 8, continuando con el análisis de datos de horarios 6 AM. Utilizando el modelo de RNA y los datos de horarios 6 AM del

caso 4, se obtuvo la estimación y se calculó la correlación día a día, comparamos los resultados con los casos anteriores, para evaluar la eficacia de la RNA en la estimación y correlación de datos en horarios 6 AM y PD.

La generación de un gráfico descriptivo permitió visualizar de manera clara la organización de las cuatro etapas y los casos estudiados en cada una.

Este gráfico facilitó la comprensión de los diferentes modelos matemáticos utilizados y la comparación de la correlación de datos 6 AM y PD en cada caso.

Para finalizar se realizará una tabla de resumen de conclusión de los resultados de los casos de estudios.

A través de la realización de un análisis comparativo de modelos matemáticos en cuatro fases, se ha conseguido una evaluación minuciosa de la estimación y correlación de datos horarios. Los resultados obtenidos aportan conocimientos valiosos para futuras investigaciones y aplicaciones prácticas en el campo de análisis horarios y predicción de datos.

### 5.3. Interpretación de resultados

La visualización de datos desempeña un papel fundamental en la comunicación efectiva de la cantidad, datos e información obtenidos a partir de la experimentación. En este sentido, se utilizaron diversas técnicas de visualización para representar los resultados obtenidos.

- **Gráfico de correlación:** Es una herramienta que nos permite demostrar la relación entre dos clases de datos, además de cuantificar la intensidad de dicha relación.
- **Gráfico de puntos:** Consta de puntos de datos representados en un gráfico con ejes x e y, se utilizan para describir gráficamente ciertas tendencias o grupos de datos.
- **Gráfico de línea:** Consiste en un instrumento que nos admite visualizar los cambios a lo largo de un rango continuo, en este caso la cantidad de prueba del experimento.

El primer tipo de visualización utilizado fue el gráfico de correlación, el cual nos permitió demostrar la relación entre dos clases de datos y cuantificar la intensidad de dicha relación. Este tipo de gráfico resultó útil para identificar patrones y tendencias en los datos, así como para comprender la influencia de una variable sobre otra.

Además, se empleó el gráfico de puntos, el cual consiste en representar los puntos de datos en un gráfico con ejes x e y. Este tipo de visualización fue utilizado para describir gráficamente ciertas tendencias o grupos de datos, lo que permitió una mejor comprensión de la distribución de los datos de medición y los estimados generados por los modelos matemáticos PD y 6 AM.

Por último, se utilizó el gráfico de línea, el cual resultó especialmente útil para visualizar los cambios a lo largo de un rango continuo, en este caso, la cantidad de prueba del experimento. A través de este tipo de visualización, se pudieron observar claramente los resultados de variables como el coeficiente de determinación ( $R^2$ ), la pendiente ( $m$ ), el coeficiente al origen ( $c$ ) y el error estándar ( $SE$ ). Esto proporcionó una comprensión más profunda de los resultados obtenidos y permitió realizar comparaciones entre diferentes pruebas y casos de estudio.

# Capítulo 6

## Implementación

### 6.1. Software utilizado

El software utilizado es Jupyter Notebook, el contenedor es Docker y para nuestro repositorio Github, los cuales se visualiza en el diagrama de componentes 6.1, que nos permite nuestra implementación de los experimentos, los cuales se enumeran además de describir a continuación.

1. **Jupyter Notebook:** El software Jupyter Notebook es una aplicación web que sirve de modo de puente constante entre el código y los textos explicativos de forma cliente-servidor, adicional como se menciona en la documentación “Que es un documento compatible que combina el código informático, descripciones en lenguaje sencillo, datos, visualizaciones enriquecidas en modelos 3D [20]. Los principales usos son la depuración de datos, creación y entrenamiento de modelos de aprendizaje automático, modelos matemáticos y la visualización de datos además de los resultados de D. Beysens y de la RNA. Como logramos visualizar en el diagrama 6.1 de componente del sistema Jupyter Notebook, se sitúa en la parte de presentación de nuestro sistema, cuyo objetivo es encargarse de la interfaz web para la visualización plasmada del código donde se realizará los algoritmos matemáticos.
2. **Docker:** Es una plataforma que nos permite crear, probar e implementar aplicaciones rápidamente, Docker es encargado de empaquetar software en contenedores que incluyen lo necesario para que el software se ejecute correctamente de forma aislada, permitiendo el despliegue de la aplicación. Adicional Docker Compose es una herramienta para definir y ejecutar aplicaciones Docker de varios contenedores. Con Compose, utilizas un archivo YAML para configurar los servicios de tu aplicación. Luego, con un solo comando, creas e inicias todos los servicios desde tu configuración [21]. En el diagrama 6.1 se puede visualizar en la parte superior en el componente

nombrado como Infraestructura que tiene como objetivo permitir el despliegue de la aplicación de Jupyter Notebook.

3. **GitHub:** La plataforma GitHub es una herramienta de hospedaje de código destinada al control de versiones y la colaboración [22], siendo la herramienta más utilizada por los programadores. En el diagrama 6.1 de componentes se visualiza en la parte de bases de datos que tiene como finalidad almacenar los modelos y los resultados.
4. **Nginx:** Nginx es uno de los servidores web más populares del mundo. Puede manejar correctamente altas cargas con muchas conexiones de clientes concurrentes y puede funcionar fácilmente como servidor web, servidor de correo o servidor de proxy inverso [23]. En el contexto de tu trabajo de titulación, Nginx puede estar desempeñando un papel importante en la configuración de la infraestructura para la implementación de tu aplicación. Por ejemplo, Nginx podría ser utilizado para dirigir las solicitudes de los usuarios a tu aplicación Jupyter Notebook que se ejecuta en el contenedor Docker, asegurando una conexión segura a través de HTTPS y equilibrando la carga entre múltiples instancias de tu aplicación si es necesario.
5. **MySQL:** MySQL es un sistema de administración de bases de datos de código abierto, que comúnmente se instala como parte de la popular pila LAMP (Linux, Apache, MySQL, PHP/Python/Perl). Implementa el modelo relacional y utiliza el Structured Query Language (más conocido como SQL) para administrar los datos [24]. De acuerdo al trabajo de título utilizaremos MySQL para guardar los resultados de los modelos matemáticos.

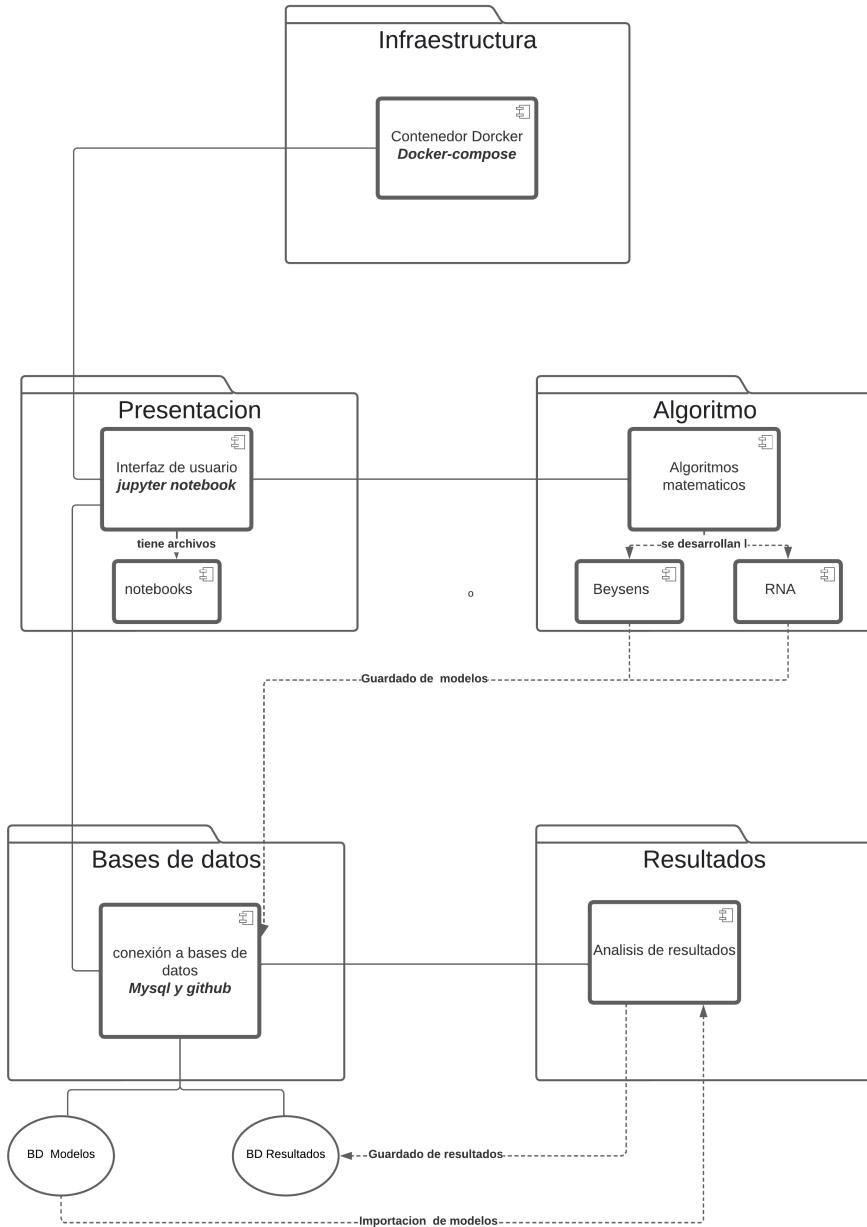


Figura 6.1: Diagrama de Componente del sistema.

## 6.2. Hardware utilizado

El trabajo experimental se realizó con el hardware MacBook Pro 13 m1 2020 ejemplificado en el siguiente esquema 6.2 el cual contiene: Memoria principal, CPU, Dispositivos de entrada, almacenamiento y salida.

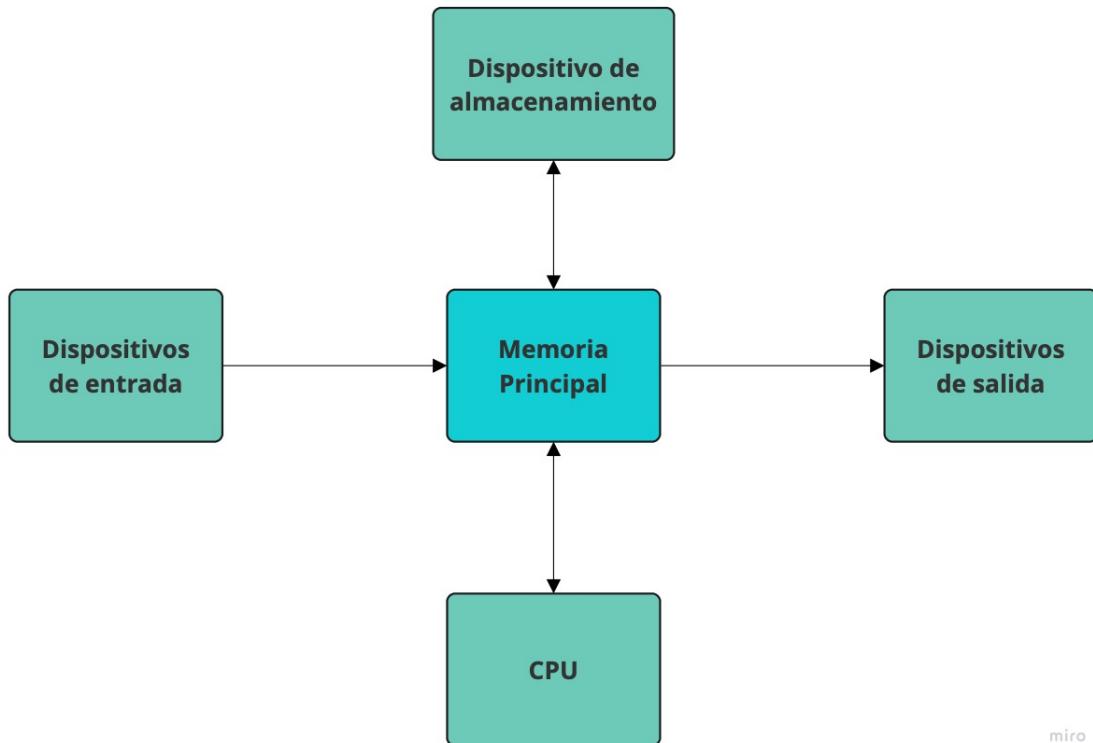


Figura 6.2: Esquema Hardware.

Componentes del Hardware utilizados localmente:

1. **Dispositivo de almacenamiento:** Contiene una SSD de 256 GB.
2. **Dispositivo de entrada:** Mouse y el teclado.
3. **Memoria Principal:** Es la Memoria RAM 8 GB.
4. **CPU:** Contiene una CPU Apple M1 silicon de 8 núcleos con 4 núcleos de rendimiento y 4 de eficiencia.
5. **Dispositivo de salida:** Pantalla de retina.

Servidor(Digital-Ocean):

1. **Dispositivo de almacenamiento:** Contiene una SSD de 50 GB.
2. **Dispositivo de entrada:** Mouse y el teclado.
3. **Memoria Principal:** la Memoria RAM 4 GB.
4. **CPU:** Contiene una CPU 4 núcleos de rendimiento.
5. **Dispositivo de salida:** Pantalla de retina.

### 6.3. Lenguajes de programación

El primer lenguaje de programación es Python, visualizado en la siguiente imagen 6.3, se define, como un lenguaje de programación versátil de múltiple plataforma, su principal objetivo es la automatización de procesos, nos permite utilizar múltiples librerías gratuitas, actualmente es distribuido como código abierto, ya que contiene una comunidad grande de desarrolladores, es de fácil uso por la similitud con el inglés, Python puede ejecutarse en sistemas Mac, Windows y Unix, también ha sido portado a Java y máquinas virtuales .NET, se considera un lenguaje de script, como Rubí o Perl. A menudo se usa para crear Web aplicaciones, la agilidad que aporta este lenguaje, así como la potencia de procesamiento de datos de las RNA hace que ambos elementos en conjunto sean la opción más eficaz y eficiente de altos rendimientos, además de las RNA y sus campos. Tanto como Machine Learning y deep learning han conseguido resultados impresionantes en áreas como la robótica, los pronósticos comerciales y estimaciones temporales, también es usado principalmente en Big Data, Data Science, Frameworks de pruebas y desarrollo web. Esto lo convierte en un lenguaje de propósito general de gran nivel, debido a su extensa biblioteca.

En efecto, adquirir conocimientos de programación es sencillo y hasta puede ser una labor entretenida para personas de todas las edades.

Esta herramienta en el año 2023 se encuentra a la vanguardia de las tecnologías, mediante su entorno amigable con el desarrollador y su enorme comunidad. Las personas que más utilizan Python son los desarrolladores, analistas y científicos de datos.



Figura 6.3: Imagen del logo de Python

El segundo lenguaje es el denominado SQL tal cual como se evidencia en la imagen ??, que significa Structured Query Language. Es un lenguaje especializado que se utiliza para trabajar con bases de datos relacionales. Con SQL, puedo hacer muchas cosas, como buscar información en la base de datos, agregar nuevos datos, actualizar o borrar datos existentes, y también crear o cambiar la estructura de las tablas y otros elementos de la base de datos.

SQL es muy importante porque es el lenguaje estándar que se utiliza en muchos sistemas de gestión de bases de datos, como MySQL, Oracle, Microsoft SQL Server y PostgreSQL. También es fácil de usar junto con otros lenguajes de programación.

En resumen, en mi investigación descubrí que SQL sigue siendo una herramienta muy importante en el año 2023. Me permite gestionar y manipular bases de datos de una manera eficiente, y su relevancia ha aumentado aún más con las nuevas tecnologías de análisis de datos y aprendizaje automático disponibles en el campo de la informática.



Figura 6.4: Imagen del logo de SQL

## 6.4. Estrategia de implementación

Las estrategias de implementación están constituidas por las siguientes fases, las cuales están de forma secuencial, como se muestra en la siguiente imagen del esquema 6.5. A continuación se describirá cada parte de la estrategia a utilizar:

1. **Instalación de software:** En esta primera fase, se procederá a instalar las aplicaciones de software necesarias, que son Jupyter Notebook y Docker. Estas herramientas serán fundamentales para el desarrollo del proyecto.
2. **Implementación del software:** En la segunda fase, se procede a implementar el software necesario. Esto implica obtener las aplicaciones activas y en funcionamiento. Con Docker, que sigue una arquitectura cliente-servidor, se logra el acceso mediante un puerto en el navegador web. Esto permite utilizar Jupyter Notebook para visualizar y trabajar con los datos de manera gráfica.
3. **Pruebas de los modelos:** En la tercera fase, se realizan pruebas exhaustivas en las que se someten los modelos matemáticos estimativos y acumulados a diez escenarios diferentes. Estas pruebas nos permiten evaluar el comportamiento de los modelos y su rendimiento en situaciones variadas.
4. **Resultados:** En la cuarta fase, se obtienen los resultados de cada modelo. Estos resultados se analizan detenidamente para extraer conclusiones relevantes y tomar decisiones basadas en ellos.
5. **Guardar resultados:** La quinta fase, se guardan los resultados de  $R^2$ , pendiente, coeficiente y error estándar, además los archivos notebook ejecutados como experimentos.
6. **Conclusión:** La última fase consiste en generar conclusiones basadas en los casos de estudio realizados. Se evalúa cada modelo en función de su coeficiente de determinación ( $R^2$ ) y su pendiente. Asimismo, se calculan las medias aritméticas y la desviación estándar de los veinte resultados de los modelos y los acumulados.

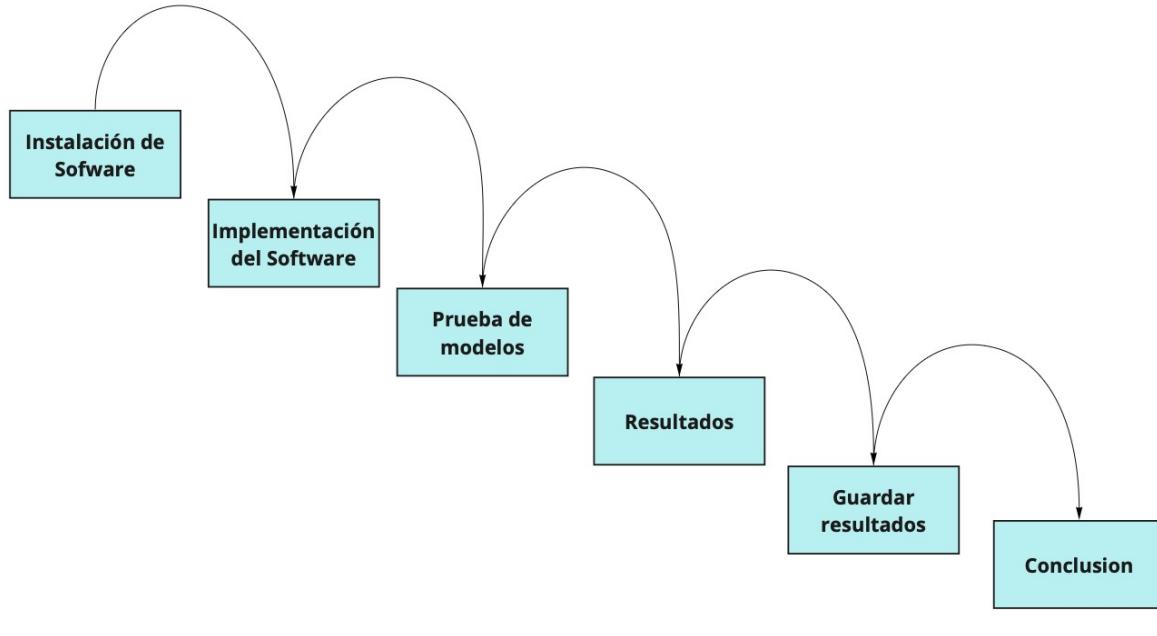


Figura 6.5: Esquema de estrategia de implementación.

A través de estas estrategias de implementación, se espera lograr un análisis riguroso de los modelos matemáticos además de obtener conclusiones sólidas en este trabajo de título. Estas conclusiones serán respaldadas en bases de datos por los resultados obtenidos en los experimentos o pruebas, que y nos permitirán obtener un mejor entendimiento de los modelos y su rendimiento.

## **6.5. Visualización**

En una primera instancia Adjuntaremos los gráficos de experimento número 4 que corresponden a cuadros comparativos de gráficos con los resultados en  $R^2$  de los modelos matemáticos con los datos de Entrenamiento y validación, en horarios PD además de 6 AM.

En una segunda instancia visualizaremos los gráficos de los modelos de validación de Beysens y RNA en horarios 6 AM y PD.

En una tercera oportunidad se empleará los cuadros de los casos de estudio y un resumen de este.

## 1. Cuadro de prueba número 1, RNA Entrenamiento y validación.

a) Modelo Beysens:

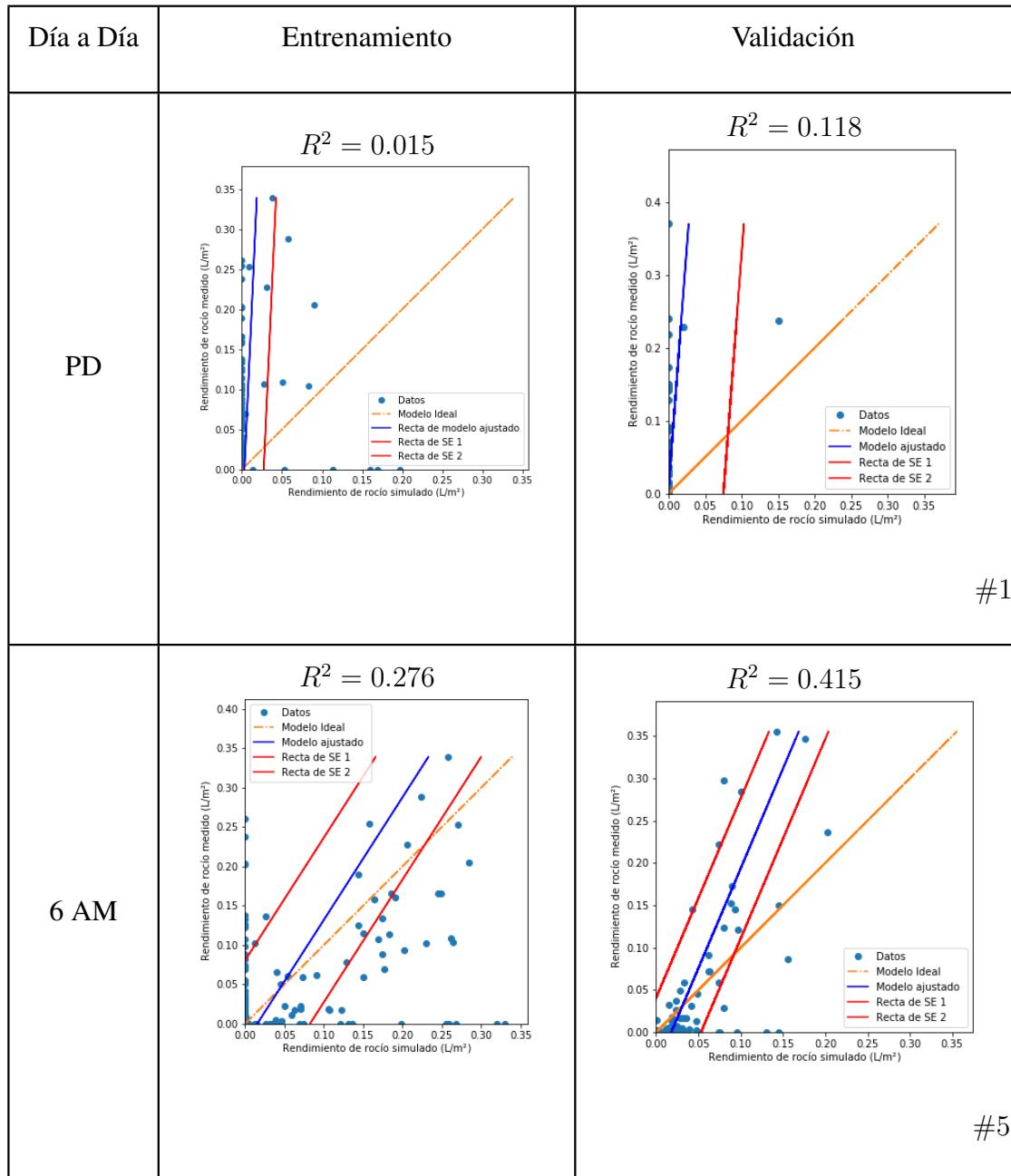


Tabla 6.1: Tabla de modelos D.Beysens

b) Modelo RNA:

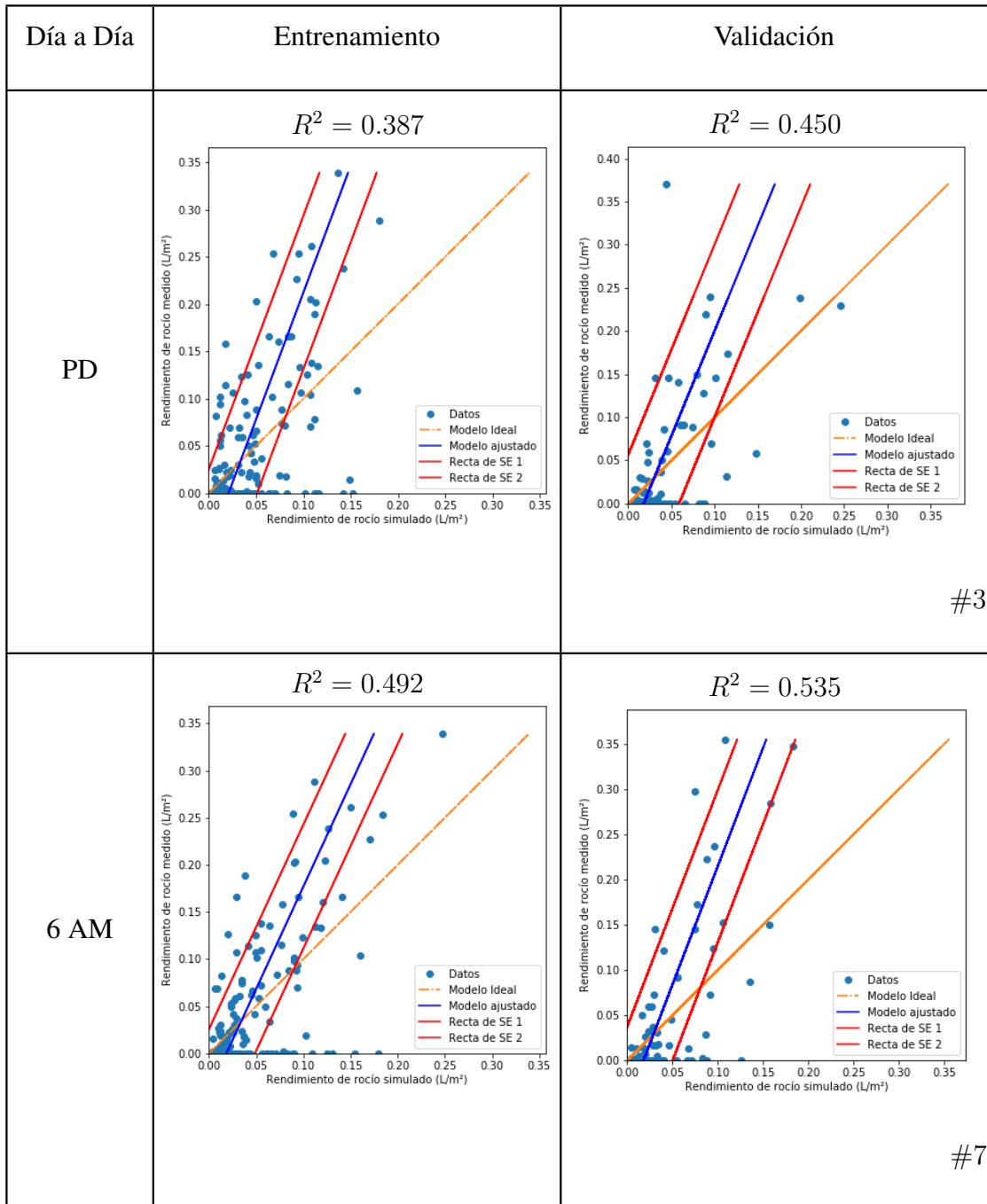


Tabla 6.2: Tabla de modelos RNA

c) Acumulado Beysens;

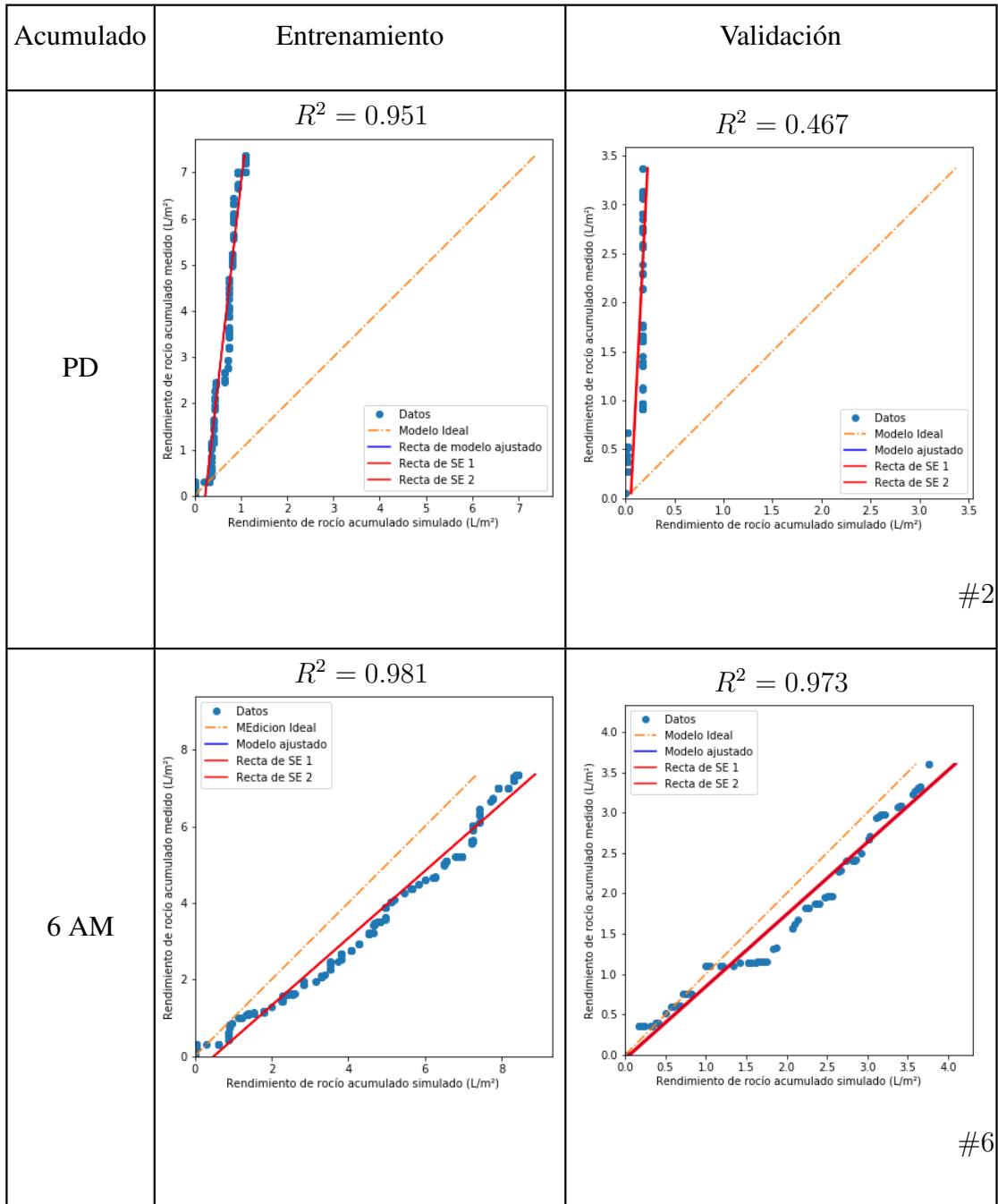


Tabla 6.3: Tabla de modelos acumulados D. Beysens

d) Acumulado RNA:

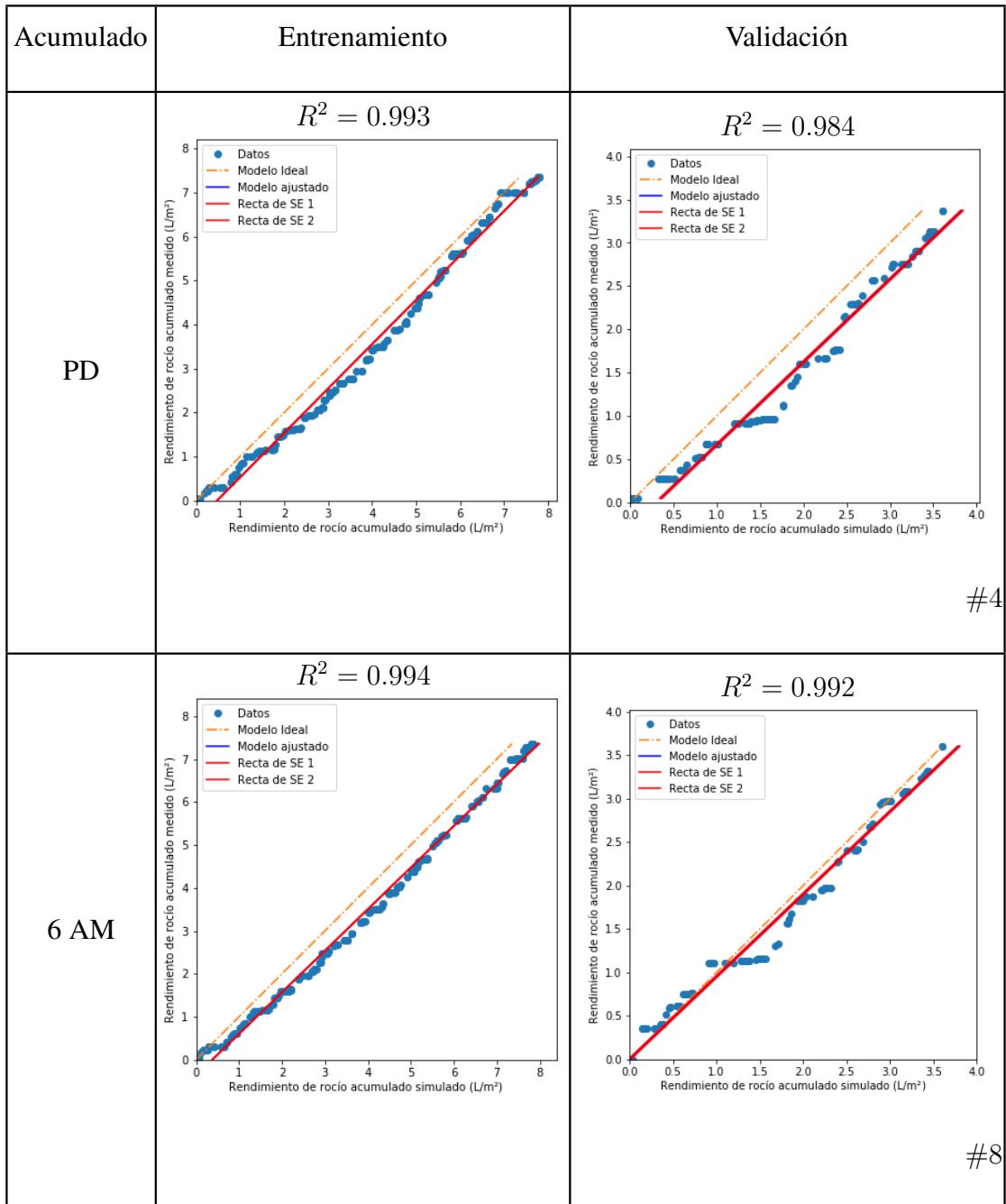


Tabla 6.4: Tabla de modelos acumulados RNA

Como discusión, logramos identificar que en el primer cuadro de "modelo de Beysens" en horario "PD" día a día y el 3 cuadro de modelo acumulado de D. Beysens, obtenemos una baja  $R^2$  y todos los datos fluctúan en el eje "Y", cabe señalar que el modelo matemático de Beysens no funciona o no es adaptable con datos PD.

**2. Cuadro de modelos de datos de Validación comparativo de Beysens y RNA en horarios PD y 6 AM.**

- Modelo horario PD:

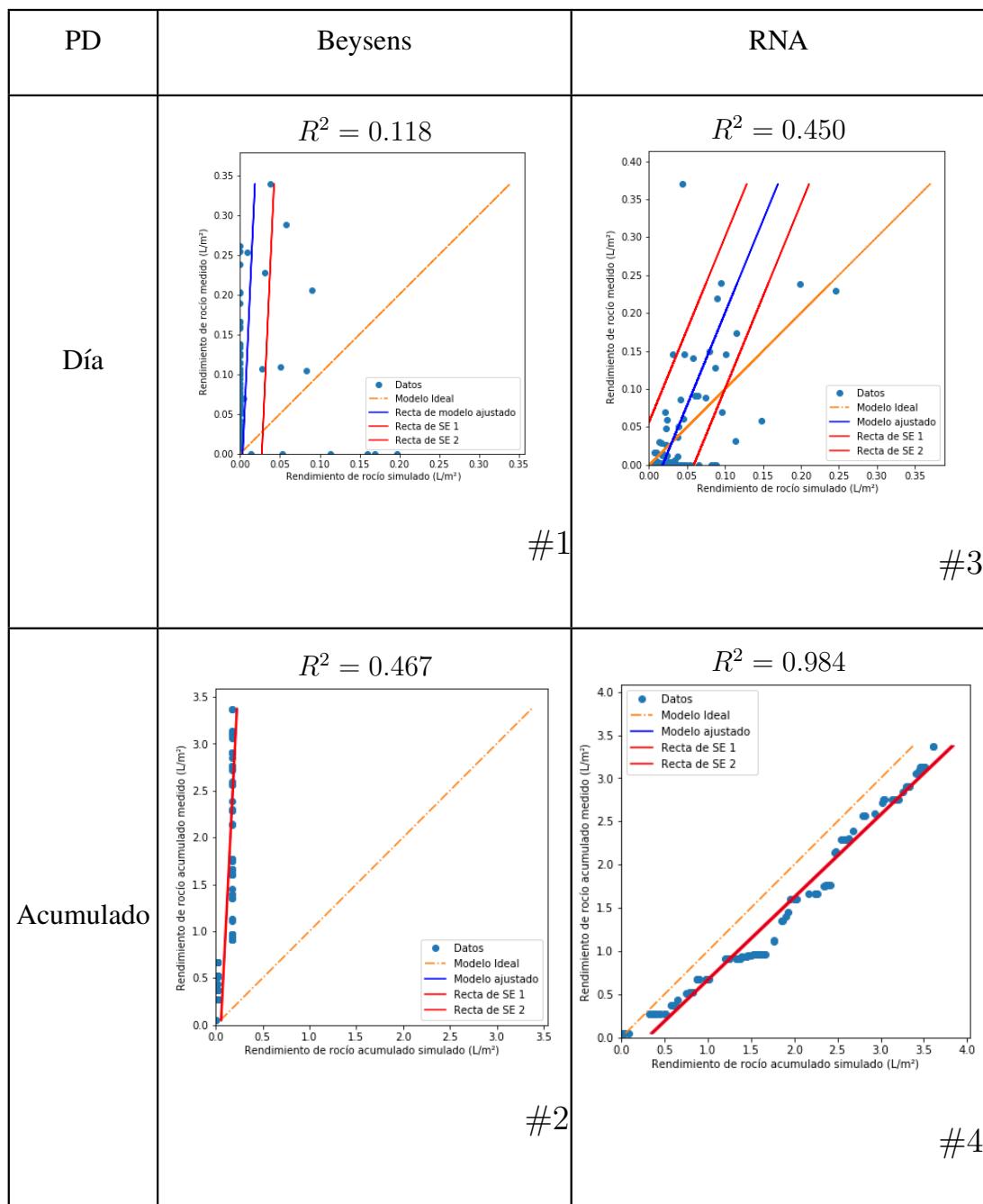


Tabla 6.5: Tabla de modelos horario PD

■ Modelo horario 6 AM:

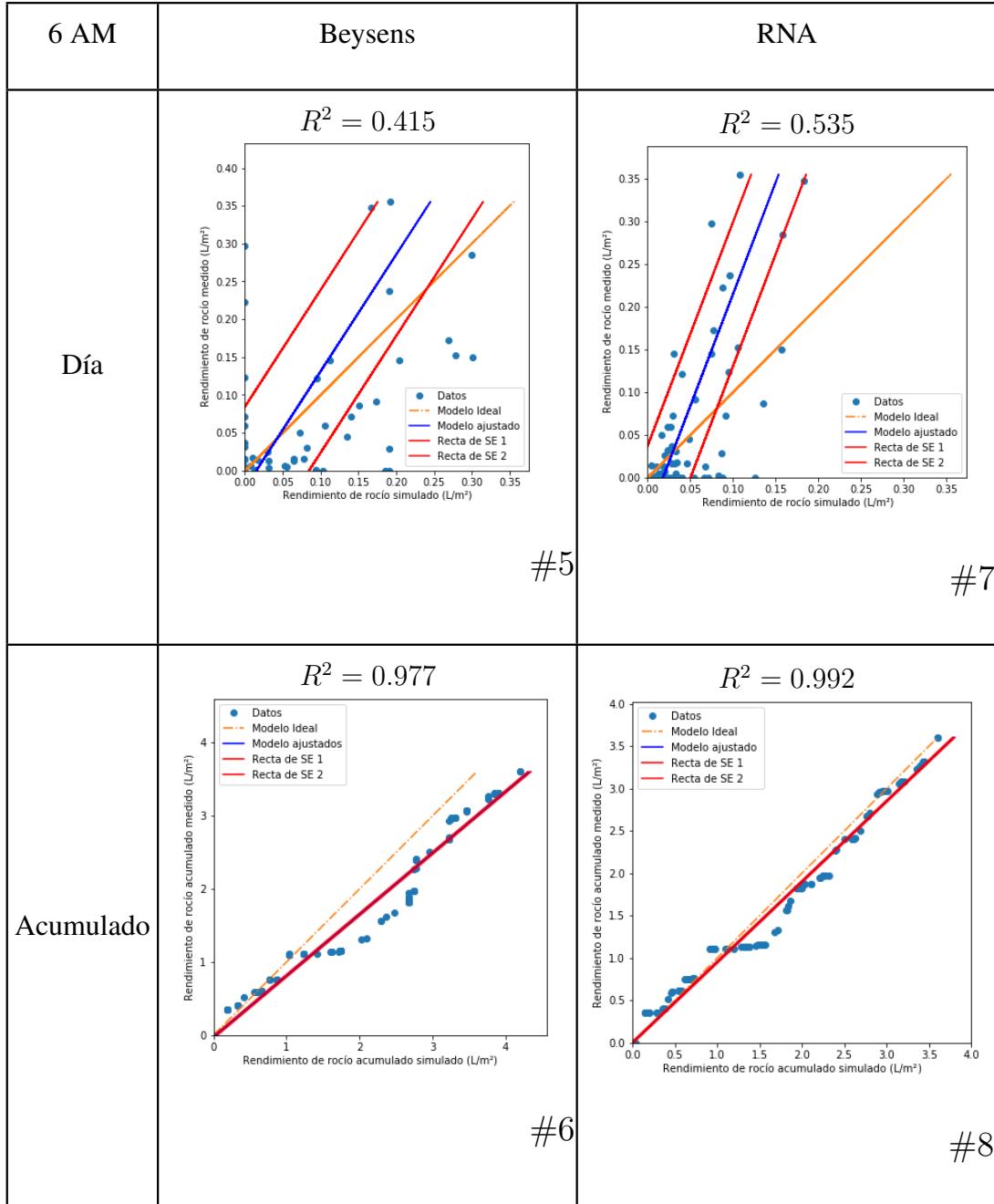


Tabla 6.6: Tabla de modelos horario 6 AM

En resumen luego de evidenciar el resultados de la prueba número 4, el modelo de D. Beysens en horario 6 AM funciona en horario diario y acumulado, en el caso que solo disponemos de datos 6 AM ya se podría predecir. En comparación con la RNA podemos garantizar que funciona en horarios PD y 6 AM, aunque el horario a las 6 AM predice y obtiene una mayor correlación de  $R^2$  por último visualizamos que el modelo de RNA supera al propuesto por D. Beysens, al obtener una correlación de  $R^2$  más alta.

### **3. Cuadro de resumen de los modelos de casos de estudios.**

Se llevó a cabo una tercera etapa importante, en la cual se realizaron cuadros comparativos relacionados con los 4 casos de estudio. Estos cuadros permitieron interpretar los modelos matemáticos y las acumulaciones correspondientes a las 10 pruebas totales del experimento. Este análisis nos permitió visualizar, diferenciar y analizar cada una de las pruebas realizadas, así como los modelos utilizados.

Como resultado de este análisis comparativo, se pudo obtener una mejor comprensión de los resultados y establecer conclusiones relevantes. Además, se generó una tabla de resumen que recopila los resultados obtenidos en el gráfico previamente analizado. Esta tabla proporciona una visión general y concisa de los resultados, permitiendo una fácil referencia y análisis de los mismos.

Con estas acciones, se logró un análisis exhaustivo y riguroso de los datos recopilados, lo que contribuye a la validez y confiabilidad de los resultados obtenidos en mi trabajo de título.

- a) Cuadro comparativo de los casos de estudios.

<b>Caso 1 y 3</b>	<b>Beysens</b> $R^2 = 0.067 \pm 0.064$ <b>RNA</b> $R^2 = 0.275 \pm 0.116$	<b>Beysens</b> $m = 0.052 \pm 0.034$ <b>RNA</b> $m = 0.356 \pm 0.099$																																																																		
RNA y Beysens diario en horario PD	<table border="1"> <caption>Data for R-squared vs Test Count</caption> <thead> <tr> <th>Cantidad de Pruebas</th> <th>R<sup>2</sup> del Modelo Beysens PD</th> <th>R<sup>2</sup> del Modelo MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.11</td><td>0.42</td></tr> <tr><td>2</td><td>0.02</td><td>0.20</td></tr> <tr><td>3</td><td>0.01</td><td>0.28</td></tr> <tr><td>4</td><td>0.17</td><td>0.35</td></tr> <tr><td>5</td><td>0.01</td><td>0.08</td></tr> <tr><td>6</td><td>0.01</td><td>0.17</td></tr> <tr><td>7</td><td>0.09</td><td>0.37</td></tr> <tr><td>8</td><td>0.14</td><td>0.32</td></tr> <tr><td>9</td><td>0.11</td><td>0.18</td></tr> <tr><td>10</td><td>0.03</td><td>0.34</td></tr> </tbody> </table>	Cantidad de Pruebas	R <sup>2</sup> del Modelo Beysens PD	R <sup>2</sup> del Modelo MLP LM PD	1	0.11	0.42	2	0.02	0.20	3	0.01	0.28	4	0.17	0.35	5	0.01	0.08	6	0.01	0.17	7	0.09	0.37	8	0.14	0.32	9	0.11	0.18	10	0.03	0.34	<table border="1"> <caption>Data for Slopes vs Test Count</caption> <thead> <tr> <th>Cantidad de Pruebas</th> <th>Pendiente del Modelo Beysens PD</th> <th>Pendiente del Modelo MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.07</td><td>0.41</td></tr> <tr><td>2</td><td>0.07</td><td>0.50</td></tr> <tr><td>3</td><td>0.01</td><td>0.45</td></tr> <tr><td>4</td><td>0.02</td><td>0.37</td></tr> <tr><td>5</td><td>0.01</td><td>0.22</td></tr> <tr><td>6</td><td>0.01</td><td>0.27</td></tr> <tr><td>7</td><td>0.01</td><td>0.29</td></tr> <tr><td>8</td><td>0.09</td><td>0.36</td></tr> <tr><td>9</td><td>0.07</td><td>0.21</td></tr> <tr><td>10</td><td>0.07</td><td>0.42</td></tr> </tbody> </table>	Cantidad de Pruebas	Pendiente del Modelo Beysens PD	Pendiente del Modelo MLP LM PD	1	0.07	0.41	2	0.07	0.50	3	0.01	0.45	4	0.02	0.37	5	0.01	0.22	6	0.01	0.27	7	0.01	0.29	8	0.09	0.36	9	0.07	0.21	10	0.07	0.42
Cantidad de Pruebas	R <sup>2</sup> del Modelo Beysens PD	R <sup>2</sup> del Modelo MLP LM PD																																																																		
1	0.11	0.42																																																																		
2	0.02	0.20																																																																		
3	0.01	0.28																																																																		
4	0.17	0.35																																																																		
5	0.01	0.08																																																																		
6	0.01	0.17																																																																		
7	0.09	0.37																																																																		
8	0.14	0.32																																																																		
9	0.11	0.18																																																																		
10	0.03	0.34																																																																		
Cantidad de Pruebas	Pendiente del Modelo Beysens PD	Pendiente del Modelo MLP LM PD																																																																		
1	0.07	0.41																																																																		
2	0.07	0.50																																																																		
3	0.01	0.45																																																																		
4	0.02	0.37																																																																		
5	0.01	0.22																																																																		
6	0.01	0.27																																																																		
7	0.01	0.29																																																																		
8	0.09	0.36																																																																		
9	0.07	0.21																																																																		
10	0.07	0.42																																																																		
	<b>Beysens</b> $b = 0.001 \pm 0.002$ <b>RNA</b> $b = 0.021 \pm 0.003$	<b>Beysens</b> $SE = 0.021 \pm 0.013$ <b>RNA</b> $SE = 0.055 \pm 0.017$																																																																		
	<table border="1"> <caption>Data for Coefficients vs Test Count</caption> <thead> <tr> <th>Resultados de coeficientes</th> <th>Resultados coef<sup>b</sup> del Modelo Beysens PD</th> <th>Resultados coef<sup>b</sup> del Modelo MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.001</td><td>0.018</td></tr> <tr><td>2</td><td>0.005</td><td>0.023</td></tr> <tr><td>3</td><td>0.002</td><td>0.023</td></tr> <tr><td>4</td><td>0.001</td><td>0.019</td></tr> <tr><td>5</td><td>0.002</td><td>0.025</td></tr> <tr><td>6</td><td>0.002</td><td>0.026</td></tr> <tr><td>7</td><td>0.001</td><td>0.018</td></tr> <tr><td>8</td><td>0.001</td><td>0.018</td></tr> <tr><td>9</td><td>0.001</td><td>0.019</td></tr> <tr><td>10</td><td>0.004</td><td>0.023</td></tr> </tbody> </table>	Resultados de coeficientes	Resultados coef <sup>b</sup> del Modelo Beysens PD	Resultados coef <sup>b</sup> del Modelo MLP LM PD	1	0.001	0.018	2	0.005	0.023	3	0.002	0.023	4	0.001	0.019	5	0.002	0.025	6	0.002	0.026	7	0.001	0.018	8	0.001	0.018	9	0.001	0.019	10	0.004	0.023	<table border="1"> <caption>Data for Standard Error vs Test Count</caption> <thead> <tr> <th>Cantidad de Pruebas</th> <th>SE del Modelo Beysens PD</th> <th>SE del Modelo MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.02</td><td>0.04</td></tr> <tr><td>2</td><td>0.05</td><td>0.08</td></tr> <tr><td>3</td><td>0.03</td><td>0.06</td></tr> <tr><td>4</td><td>0.005</td><td>0.011</td></tr> <tr><td>5</td><td>0.03</td><td>0.07</td></tr> <tr><td>6</td><td>0.025</td><td>0.05</td></tr> <tr><td>7</td><td>0.03</td><td>0.035</td></tr> <tr><td>8</td><td>0.015</td><td>0.04</td></tr> <tr><td>9</td><td>0.02</td><td>0.04</td></tr> <tr><td>10</td><td>0.04</td><td>0.05</td></tr> </tbody> </table>	Cantidad de Pruebas	SE del Modelo Beysens PD	SE del Modelo MLP LM PD	1	0.02	0.04	2	0.05	0.08	3	0.03	0.06	4	0.005	0.011	5	0.03	0.07	6	0.025	0.05	7	0.03	0.035	8	0.015	0.04	9	0.02	0.04	10	0.04	0.05
Resultados de coeficientes	Resultados coef <sup>b</sup> del Modelo Beysens PD	Resultados coef <sup>b</sup> del Modelo MLP LM PD																																																																		
1	0.001	0.018																																																																		
2	0.005	0.023																																																																		
3	0.002	0.023																																																																		
4	0.001	0.019																																																																		
5	0.002	0.025																																																																		
6	0.002	0.026																																																																		
7	0.001	0.018																																																																		
8	0.001	0.018																																																																		
9	0.001	0.019																																																																		
10	0.004	0.023																																																																		
Cantidad de Pruebas	SE del Modelo Beysens PD	SE del Modelo MLP LM PD																																																																		
1	0.02	0.04																																																																		
2	0.05	0.08																																																																		
3	0.03	0.06																																																																		
4	0.005	0.011																																																																		
5	0.03	0.07																																																																		
6	0.025	0.05																																																																		
7	0.03	0.035																																																																		
8	0.015	0.04																																																																		
9	0.02	0.04																																																																		
10	0.04	0.05																																																																		

Tabla 6.7: Tabla de descripción Casos de estudio 1 y 3 “Día a día PD”.

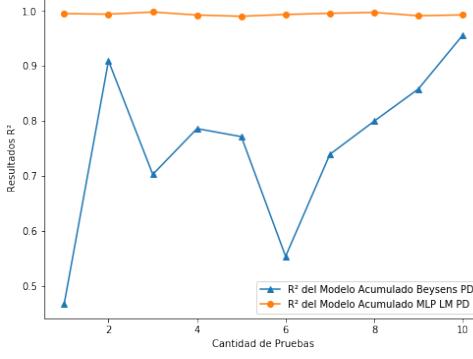
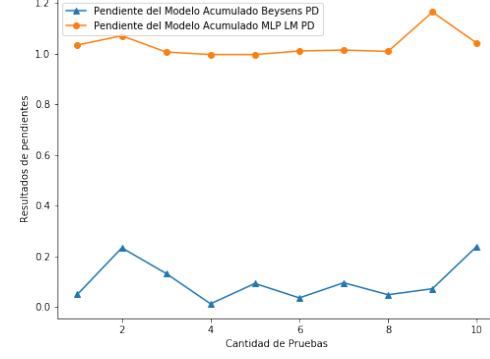
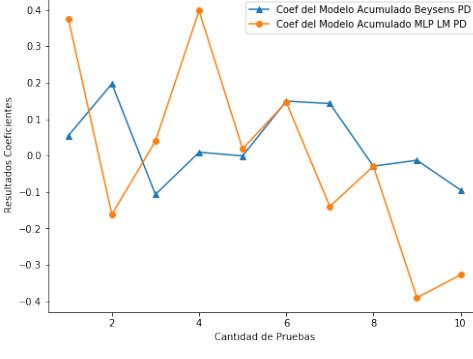
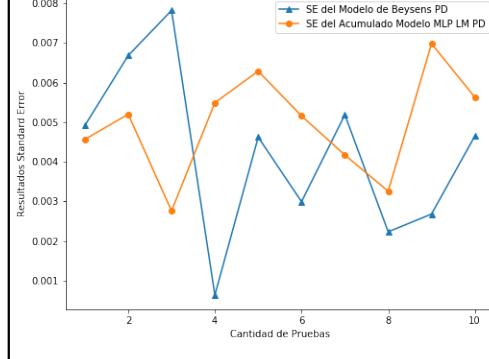
<b>Caso 2 y 4</b>	<b>Beysens</b> $R^2 = 0.754 \pm 0.150$ <b>RNA</b> $R^2 = 0.994 \pm 0.002$	<b>Beysens</b> $m = 0.102 \pm 0.051$ <b>RNA</b> $m = 1.033 \pm 0.051$																																																																		
RNA y Beysens Acumulado en horario PD	 <table border="1"> <caption>Data for R-squared vs Trials (RNA y Beysens Acumulado en horario PD)</caption> <thead> <tr> <th>Pruebas</th> <th>R<sup>2</sup> del Modelo Acumulado Beysens PD</th> <th>R<sup>2</sup> del Modelo Acumulado MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.48</td><td>1.00</td></tr> <tr><td>2</td><td>0.92</td><td>1.00</td></tr> <tr><td>3</td><td>0.71</td><td>1.00</td></tr> <tr><td>4</td><td>0.79</td><td>1.00</td></tr> <tr><td>5</td><td>0.77</td><td>1.00</td></tr> <tr><td>6</td><td>0.55</td><td>1.00</td></tr> <tr><td>7</td><td>0.75</td><td>1.00</td></tr> <tr><td>8</td><td>0.80</td><td>1.00</td></tr> <tr><td>9</td><td>0.86</td><td>1.00</td></tr> <tr><td>10</td><td>0.95</td><td>1.00</td></tr> </tbody> </table>	Pruebas	R <sup>2</sup> del Modelo Acumulado Beysens PD	R <sup>2</sup> del Modelo Acumulado MLP LM PD	1	0.48	1.00	2	0.92	1.00	3	0.71	1.00	4	0.79	1.00	5	0.77	1.00	6	0.55	1.00	7	0.75	1.00	8	0.80	1.00	9	0.86	1.00	10	0.95	1.00	 <table border="1"> <caption>Data for Slopes vs Trials (RNA y Beysens Acumulado en horario PD)</caption> <thead> <tr> <th>Pruebas</th> <th>Pendiente del Modelo Acumulado Beysens PD</th> <th>Pendiente del Modelo Acumulado MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.05</td><td>1.02</td></tr> <tr><td>2</td><td>0.25</td><td>1.05</td></tr> <tr><td>3</td><td>0.15</td><td>1.00</td></tr> <tr><td>4</td><td>0.02</td><td>1.00</td></tr> <tr><td>5</td><td>0.08</td><td>1.00</td></tr> <tr><td>6</td><td>0.05</td><td>1.00</td></tr> <tr><td>7</td><td>0.08</td><td>1.00</td></tr> <tr><td>8</td><td>0.05</td><td>1.00</td></tr> <tr><td>9</td><td>0.08</td><td>1.15</td></tr> <tr><td>10</td><td>0.22</td><td>1.02</td></tr> </tbody> </table>	Pruebas	Pendiente del Modelo Acumulado Beysens PD	Pendiente del Modelo Acumulado MLP LM PD	1	0.05	1.02	2	0.25	1.05	3	0.15	1.00	4	0.02	1.00	5	0.08	1.00	6	0.05	1.00	7	0.08	1.00	8	0.05	1.00	9	0.08	1.15	10	0.22	1.02
Pruebas	R <sup>2</sup> del Modelo Acumulado Beysens PD	R <sup>2</sup> del Modelo Acumulado MLP LM PD																																																																		
1	0.48	1.00																																																																		
2	0.92	1.00																																																																		
3	0.71	1.00																																																																		
4	0.79	1.00																																																																		
5	0.77	1.00																																																																		
6	0.55	1.00																																																																		
7	0.75	1.00																																																																		
8	0.80	1.00																																																																		
9	0.86	1.00																																																																		
10	0.95	1.00																																																																		
Pruebas	Pendiente del Modelo Acumulado Beysens PD	Pendiente del Modelo Acumulado MLP LM PD																																																																		
1	0.05	1.02																																																																		
2	0.25	1.05																																																																		
3	0.15	1.00																																																																		
4	0.02	1.00																																																																		
5	0.08	1.00																																																																		
6	0.05	1.00																																																																		
7	0.08	1.00																																																																		
8	0.05	1.00																																																																		
9	0.08	1.15																																																																		
10	0.22	1.02																																																																		
	<b>Beysens</b> $b = 0.031 \pm 0.103$ <b>RNA</b> $b = -0.006 \pm 0.264$	<b>Beysens</b> $SE = 0.004 \pm 0.002$ <b>RNA</b> $SE = 0.002 \pm 0.001$																																																																		
	 <table border="1"> <caption>Data for Coefficients vs Trials (RNA y Beysens Acumulado en horario PD)</caption> <thead> <tr> <th>Pruebas</th> <th>Coef del Modelo Acumulado Beysens PD</th> <th>Coef del Modelo Acumulado MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.08</td><td>0.38</td></tr> <tr><td>2</td><td>0.20</td><td>-0.18</td></tr> <tr><td>3</td><td>-0.05</td><td>0.08</td></tr> <tr><td>4</td><td>0.02</td><td>0.40</td></tr> <tr><td>5</td><td>-0.02</td><td>0.05</td></tr> <tr><td>6</td><td>0.12</td><td>0.12</td></tr> <tr><td>7</td><td>0.15</td><td>-0.12</td></tr> <tr><td>8</td><td>-0.05</td><td>-0.05</td></tr> <tr><td>9</td><td>-0.05</td><td>-0.38</td></tr> <tr><td>10</td><td>-0.08</td><td>-0.32</td></tr> </tbody> </table>	Pruebas	Coef del Modelo Acumulado Beysens PD	Coef del Modelo Acumulado MLP LM PD	1	0.08	0.38	2	0.20	-0.18	3	-0.05	0.08	4	0.02	0.40	5	-0.02	0.05	6	0.12	0.12	7	0.15	-0.12	8	-0.05	-0.05	9	-0.05	-0.38	10	-0.08	-0.32	 <table border="1"> <caption>Data for Standard Error vs Trials (RNA y Beysens Acumulado en horario PD)</caption> <thead> <tr> <th>Pruebas</th> <th>SE del Modelo de Beysens PD</th> <th>SE del Acumulado Modelo MLP LM PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.0045</td><td>0.0045</td></tr> <tr><td>2</td><td>0.0060</td><td>0.0050</td></tr> <tr><td>3</td><td>0.0025</td><td>0.0025</td></tr> <tr><td>4</td><td>0.0008</td><td>0.0055</td></tr> <tr><td>5</td><td>0.0045</td><td>0.0060</td></tr> <tr><td>6</td><td>0.0030</td><td>0.0050</td></tr> <tr><td>7</td><td>0.0045</td><td>0.0035</td></tr> <tr><td>8</td><td>0.0020</td><td>0.0030</td></tr> <tr><td>9</td><td>0.0030</td><td>0.0070</td></tr> <tr><td>10</td><td>0.0045</td><td>0.0055</td></tr> </tbody> </table>	Pruebas	SE del Modelo de Beysens PD	SE del Acumulado Modelo MLP LM PD	1	0.0045	0.0045	2	0.0060	0.0050	3	0.0025	0.0025	4	0.0008	0.0055	5	0.0045	0.0060	6	0.0030	0.0050	7	0.0045	0.0035	8	0.0020	0.0030	9	0.0030	0.0070	10	0.0045	0.0055
Pruebas	Coef del Modelo Acumulado Beysens PD	Coef del Modelo Acumulado MLP LM PD																																																																		
1	0.08	0.38																																																																		
2	0.20	-0.18																																																																		
3	-0.05	0.08																																																																		
4	0.02	0.40																																																																		
5	-0.02	0.05																																																																		
6	0.12	0.12																																																																		
7	0.15	-0.12																																																																		
8	-0.05	-0.05																																																																		
9	-0.05	-0.38																																																																		
10	-0.08	-0.32																																																																		
Pruebas	SE del Modelo de Beysens PD	SE del Acumulado Modelo MLP LM PD																																																																		
1	0.0045	0.0045																																																																		
2	0.0060	0.0050																																																																		
3	0.0025	0.0025																																																																		
4	0.0008	0.0055																																																																		
5	0.0045	0.0060																																																																		
6	0.0030	0.0050																																																																		
7	0.0045	0.0035																																																																		
8	0.0020	0.0030																																																																		
9	0.0030	0.0070																																																																		
10	0.0045	0.0055																																																																		

Tabla 6.8: Tabla de descripción Casos de estudio 2 y 4 “Acumulado PD”.

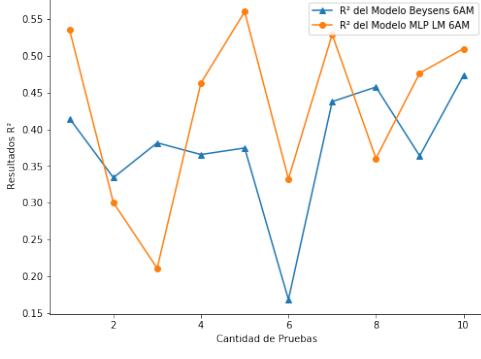
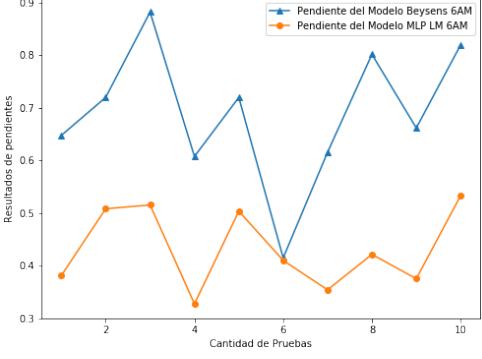
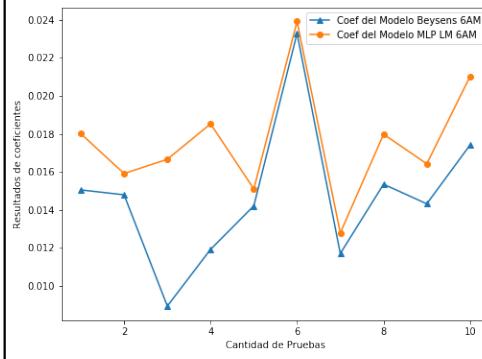
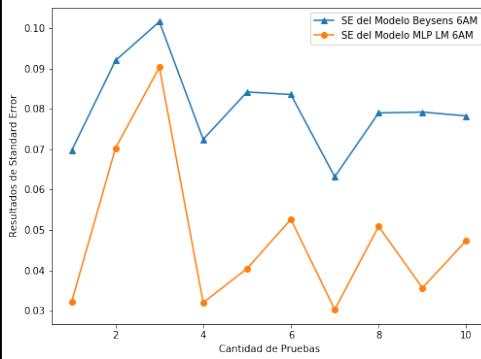
<b>Caso 5 y 7</b>	<b>Beysens</b> $R^2 = 0.377 \pm 0.086$ <b>RNA</b> $R^2 = 0.427 \pm 0.119$	<b>Beysens</b> $m = 0.688 \pm 0.133$ <b>RNA</b> $m = 0.433 \pm 0.075$
RNA y Beysens diario en horario 6 AM		
	<b>Beysens</b> $b = 0.014 \pm 0.004$ <b>RNA</b> $b = 0.018 \pm 0.003$	<b>Beysens</b> $SE = 0.080 \pm 0.021$ <b>RNA</b> $SE = 0.048 \pm 0.019$
		

Tabla 6.9: Tabla de descripción Casos de estudio 5 y 7, Día a día en horario 6 AM.

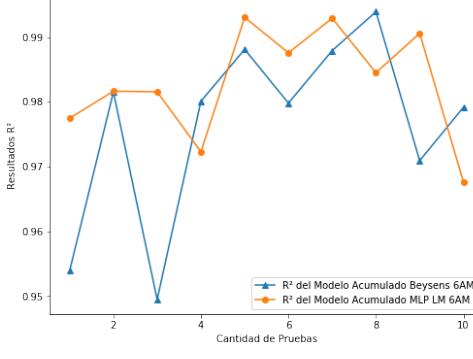
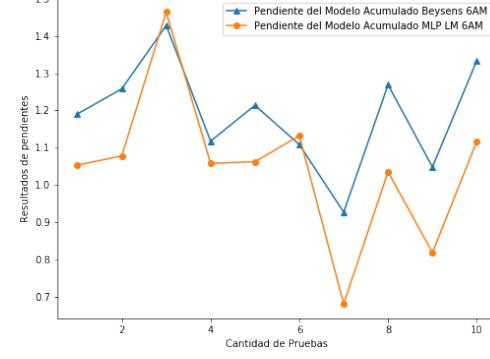
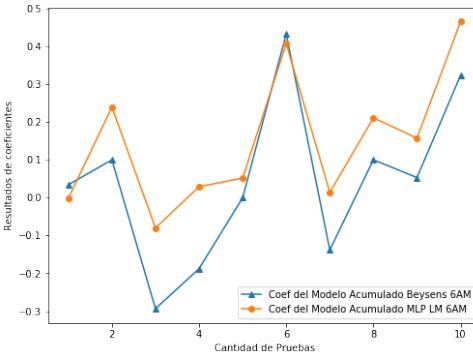
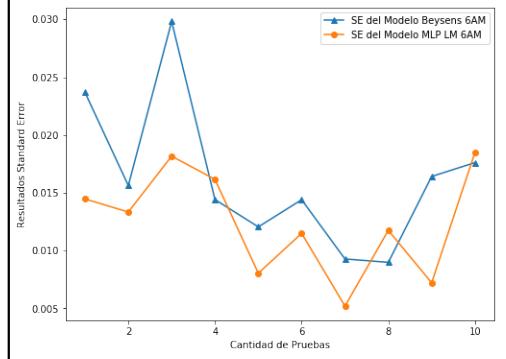
<b>Caso 6 y 8</b>	<b>Beysens</b> $R^2 = 0.976 \pm 0.014$ <b>RNA</b> $R^2 = 0.983 \pm 0.008$	<b>Beysens</b> $m = 1.189 \pm 0.145$ <b>RNA</b> $m = 1.050 \pm 0.203$																																																																		
RNA y Beysens acumulado en horario 6 AM	 <table border="1"> <caption>Data for R-squared vs Trials (6 AM)</caption> <thead> <tr> <th>Pruebas</th> <th>R<sup>2</sup> del Modelo Acumulado Beysens 6AM</th> <th>R<sup>2</sup> del Modelo Acumulado MLP LM 6AM</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.955</td><td>0.978</td></tr> <tr><td>2</td><td>0.982</td><td>0.982</td></tr> <tr><td>3</td><td>0.952</td><td>0.982</td></tr> <tr><td>4</td><td>0.978</td><td>0.970</td></tr> <tr><td>5</td><td>0.990</td><td>0.992</td></tr> <tr><td>6</td><td>0.978</td><td>0.985</td></tr> <tr><td>7</td><td>0.985</td><td>0.990</td></tr> <tr><td>8</td><td>0.988</td><td>0.982</td></tr> <tr><td>9</td><td>0.972</td><td>0.990</td></tr> <tr><td>10</td><td>0.978</td><td>0.968</td></tr> </tbody> </table>	Pruebas	R <sup>2</sup> del Modelo Acumulado Beysens 6AM	R <sup>2</sup> del Modelo Acumulado MLP LM 6AM	1	0.955	0.978	2	0.982	0.982	3	0.952	0.982	4	0.978	0.970	5	0.990	0.992	6	0.978	0.985	7	0.985	0.990	8	0.988	0.982	9	0.972	0.990	10	0.978	0.968	 <table border="1"> <caption>Data for Slopes vs Trials (6 AM)</caption> <thead> <tr> <th>Pruebas</th> <th>Pendiente del Modelo Acumulado Beysens 6AM</th> <th>Pendiente del Modelo Acumulado MLP LM 6AM</th> </tr> </thead> <tbody> <tr><td>1</td><td>1.20</td><td>1.05</td></tr> <tr><td>2</td><td>1.25</td><td>1.08</td></tr> <tr><td>3</td><td>1.45</td><td>1.48</td></tr> <tr><td>4</td><td>1.15</td><td>1.05</td></tr> <tr><td>5</td><td>1.20</td><td>1.08</td></tr> <tr><td>6</td><td>1.12</td><td>1.12</td></tr> <tr><td>7</td><td>0.92</td><td>0.68</td></tr> <tr><td>8</td><td>1.28</td><td>1.05</td></tr> <tr><td>9</td><td>1.05</td><td>0.82</td></tr> <tr><td>10</td><td>1.30</td><td>1.12</td></tr> </tbody> </table>	Pruebas	Pendiente del Modelo Acumulado Beysens 6AM	Pendiente del Modelo Acumulado MLP LM 6AM	1	1.20	1.05	2	1.25	1.08	3	1.45	1.48	4	1.15	1.05	5	1.20	1.08	6	1.12	1.12	7	0.92	0.68	8	1.28	1.05	9	1.05	0.82	10	1.30	1.12
Pruebas	R <sup>2</sup> del Modelo Acumulado Beysens 6AM	R <sup>2</sup> del Modelo Acumulado MLP LM 6AM																																																																		
1	0.955	0.978																																																																		
2	0.982	0.982																																																																		
3	0.952	0.982																																																																		
4	0.978	0.970																																																																		
5	0.990	0.992																																																																		
6	0.978	0.985																																																																		
7	0.985	0.990																																																																		
8	0.988	0.982																																																																		
9	0.972	0.990																																																																		
10	0.978	0.968																																																																		
Pruebas	Pendiente del Modelo Acumulado Beysens 6AM	Pendiente del Modelo Acumulado MLP LM 6AM																																																																		
1	1.20	1.05																																																																		
2	1.25	1.08																																																																		
3	1.45	1.48																																																																		
4	1.15	1.05																																																																		
5	1.20	1.08																																																																		
6	1.12	1.12																																																																		
7	0.92	0.68																																																																		
8	1.28	1.05																																																																		
9	1.05	0.82																																																																		
10	1.30	1.12																																																																		
	<b>Beysens</b> $b = 0.042 \pm 0.220$ <b>RNA</b> $b = 0.149 \pm 0.181$	<b>Beysens</b> $SE = 0.016 \pm 0.006$ <b>RNA</b> $SE = 0.012 \pm 0.005$																																																																		
	 <table border="1"> <caption>Data for Coefficients vs Trials (6 AM)</caption> <thead> <tr> <th>Pruebas</th> <th>Coef del Modelo Acumulado Beysens 6AM</th> <th>Coef del Modelo Acumulado MLP LM 6AM</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.05</td><td>0.02</td></tr> <tr><td>2</td><td>0.12</td><td>0.22</td></tr> <tr><td>3</td><td>-0.28</td><td>-0.08</td></tr> <tr><td>4</td><td>-0.18</td><td>0.05</td></tr> <tr><td>5</td><td>0.02</td><td>0.08</td></tr> <tr><td>6</td><td>0.42</td><td>0.40</td></tr> <tr><td>7</td><td>-0.12</td><td>0.02</td></tr> <tr><td>8</td><td>0.10</td><td>0.20</td></tr> <tr><td>9</td><td>0.08</td><td>0.15</td></tr> <tr><td>10</td><td>0.32</td><td>0.45</td></tr> </tbody> </table>	Pruebas	Coef del Modelo Acumulado Beysens 6AM	Coef del Modelo Acumulado MLP LM 6AM	1	0.05	0.02	2	0.12	0.22	3	-0.28	-0.08	4	-0.18	0.05	5	0.02	0.08	6	0.42	0.40	7	-0.12	0.02	8	0.10	0.20	9	0.08	0.15	10	0.32	0.45	 <table border="1"> <caption>Data for Standard Error vs Trials (6 AM)</caption> <thead> <tr> <th>Pruebas</th> <th>SE del Modelo Beysens 6AM</th> <th>SE del Modelo MLP LM 6AM</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.025</td><td>0.014</td></tr> <tr><td>2</td><td>0.015</td><td>0.013</td></tr> <tr><td>3</td><td>0.030</td><td>0.017</td></tr> <tr><td>4</td><td>0.008</td><td>0.015</td></tr> <tr><td>5</td><td>0.012</td><td>0.007</td></tr> <tr><td>6</td><td>0.012</td><td>0.011</td></tr> <tr><td>7</td><td>0.007</td><td>0.005</td></tr> <tr><td>8</td><td>0.008</td><td>0.012</td></tr> <tr><td>9</td><td>0.015</td><td>0.006</td></tr> <tr><td>10</td><td>0.017</td><td>0.018</td></tr> </tbody> </table>	Pruebas	SE del Modelo Beysens 6AM	SE del Modelo MLP LM 6AM	1	0.025	0.014	2	0.015	0.013	3	0.030	0.017	4	0.008	0.015	5	0.012	0.007	6	0.012	0.011	7	0.007	0.005	8	0.008	0.012	9	0.015	0.006	10	0.017	0.018
Pruebas	Coef del Modelo Acumulado Beysens 6AM	Coef del Modelo Acumulado MLP LM 6AM																																																																		
1	0.05	0.02																																																																		
2	0.12	0.22																																																																		
3	-0.28	-0.08																																																																		
4	-0.18	0.05																																																																		
5	0.02	0.08																																																																		
6	0.42	0.40																																																																		
7	-0.12	0.02																																																																		
8	0.10	0.20																																																																		
9	0.08	0.15																																																																		
10	0.32	0.45																																																																		
Pruebas	SE del Modelo Beysens 6AM	SE del Modelo MLP LM 6AM																																																																		
1	0.025	0.014																																																																		
2	0.015	0.013																																																																		
3	0.030	0.017																																																																		
4	0.008	0.015																																																																		
5	0.012	0.007																																																																		
6	0.012	0.011																																																																		
7	0.007	0.005																																																																		
8	0.008	0.012																																																																		
9	0.015	0.006																																																																		
10	0.017	0.018																																																																		

Tabla 6.10: Tabla de descripción Casos de estudio 1 y 3 “Día a día PD”

b) Resumen de cuadros comparativos de los casos de estudios.

Después de revisar y analizar las tablas anteriores correspondientes a los casos de estudio, se ha generado una tabla de resumen que simplifica la interpretación de los resultados. Esta tabla resume de manera concisa los datos obtenidos y permite analizarlos utilizando técnicas de análisis como la media aritmética y la desviación estándar.

La tabla se divide en dos partes. La primera parte representa los resultados del primer y segundo caso de estudio, mientras que la segunda parte representa los resultados del tercer y cuarto caso de estudio. Esta división facilita la comparación entre los diferentes casos y ayuda a identificar patrones o tendencias comunes.

Al utilizar la media aritmética y la desviación estándar como técnicas de análisis, se busca obtener una medida promedio de los resultados y evaluar la variabilidad de los mismos. Esto proporciona una visión más clara de la consistencia de los resultados y permite realizar conclusiones más sólidas.

En resumen, la tabla de resumen es una herramienta clave en mi tesis de informática, ya que condensa y simplifica la información obtenida de los casos de estudio, permitiendo un análisis más fácil e interpretación de los resultados a través de técnicas de análisis estadístico.

PD	Beysens	RNA
Día a día	$R^2 = 0.067 \pm 0.064$ $m = 0.052 \pm 0.034$ $b = 0.021 \pm 0.003$ $SE = 0.002 \pm 0.001$ <span style="text-align: right;">#caso1</span>	$R^2 = \mathbf{0.275} \pm 0.116$ $m = 0.356 \pm 0.099$ $b = 0.021 \pm 0.003$ $SE = 0.055 \pm 0.017$ <span style="text-align: right;">#caso3</span>
Acumulado	$R^2 = 0.754 \pm 0.150$ $m = 0.102 \pm 0.078$ $b = 0.031 \pm 0.103$ $SE = 0.004 \pm 0.002$ <span style="text-align: right;">#caso2</span>	$R^2 = 0.994 \pm 0.002$ $m = \mathbf{1.033} \pm 0.051$ $b = -0.006 \pm 0.264$ $SE = 0.002 \pm 0.001$ <span style="text-align: right;">#caso4</span>

Tabla 6.11: Resumen de los parámetros de correlación (promedio y desviación estándar de las 10 pruebas), resultados PD.

<b>6 AM</b>	Beysens	RNA
Día a día	$R^2 = \mathbf{0.377} \pm 0.086$ $m = 0.688 \pm 0.133$ $b = 0.014 \pm 0.004$ $SE = 0.080 \pm 0.011$ <span style="text-align: right;">#caso5</span>	$R^2 = \mathbf{0.427} \pm 0.119$ $m = 0.433 \pm 0.075$ $b = 0.014 \pm 0.004$ $SE = 0.048 \pm 0.019$ <span style="text-align: right;">#caso7</span>
Acumulado	$R^2 = 0.976 \pm 0.014$ $m = \mathbf{1.189} \pm 0.257$ $b = 0.042 \pm 0.220$ $SE = 0.016 \pm 0.006$ <span style="text-align: right;">#caso6</span>	$R^2 = 0.983 \pm 0.008$ $m = \mathbf{1.050} \pm 0.203$ $b = 0.149 \pm 0.181$ $SE = 0.012 \pm 0.005$ <span style="text-align: right;">#caso8</span>

Tabla 6.12: Resumen de los parámetros de correlación (promedio y desviación estándar de las 10 pruebas), resultados 6 AM.

En la etapa final de mi tesis de informática, al concluir la revisión de las tablas de resumen, se ha llegado a la deducción de que el modelo RNA y el modelo acumulado-RNA son superiores en rendimiento en comparación con el modelo D. Beysens y el modelo acumulados-Beysens. Esta conclusión se basa en el análisis de los datos recopilados y procesados en las tablas correspondientes.

En particular, al examinar la segunda tabla denominada “6.6 - 6 AM”, se observa que los resultados obtenidos son superiores en comparación con la primera tabla “6.5 - PD”. Esto implica que los modelos RNA y acumulado-RNA presentan un mejor desempeño en términos de los criterios y métricas utilizadas en mi investigación.

Estos hallazgos son significativos, ya que demuestran que los modelos basados en RNA ofrecen mejores resultados en relación con los modelos basados en D. Beysens. Estos resultados respaldan y refuerzan la validez y efectividad de los modelos RNA y acumulado-RNA en el contexto de la investigación.

En resumen, los resultados obtenidos de las tablas de resumen concluyen que los modelos RNA y acumulado-RNA son superiores a los modelos D. Beysens y acumulados-Beysens. Además, se destaca que la segunda tabla “6.6 - 6 AM” presenta resultados superiores en comparación con la primera tabla “6.5 - PD”.

De acuerdo a los resultados según los gráficos y la tabla de resumen, logramos identificar que el modelo de Beysens en horarios PD no funciona bien, ya que el modelo de D. Beysens no se adapta con datos en horario PD, predice valores muy pequeños de rocío, esto se evidenció en la tabla de modelo y acumulado de Beysens. Cabe señalar que estos hallazgos son relevantes y respaldan la importancia de implementar estos los modelos basados en RNA.

# **Capítulo 7**

## **Conclusiones**

En conclusión, este trabajo de título ha abordado de manera exitosa la tarea de predecir la cantidad de rocío, comparando el modelo de Beysens y una red neuronal artificial (RNA). El objetivo principal fue contribuir parcialmente a resolver la problemática de escasez hídrica que afecta tanto a nivel local como global.

Los datos analizados, obtenidos a través del proyecto FONDECYT Iniciación “11140863” y correspondientes a la localidad de Combarbalá, respaldaron y fundamentaron los resultados obtenidos. Siguiendo las referencias bibliográficas [6] y [1], se utilizó el modelo matemático de Beysens y se implementó una red neuronal en este estudio.

Inicialmente, se concluyó que el modelo requería un filtro de datos antes de aplicar cualquier enfoque adicional. Además, se encontró que la RNA de entrenamiento Adamax mostró un desempeño superior en comparación con Levenberg-Marquardt, especialmente en relación con los datos del norte de Chile.

Se realizaron predicciones utilizando datos de mediciones de promedios diarios y mediciones tomadas a las 6 AM. Además, se consideraron las correlaciones tanto en un enfoque “día a día” como en datos “acumulados”. Los resultados obtenidos se resumen en la tabla presentada, donde se detallan los métodos utilizados, la correlación obtenida, el desempeño y los valores de  $R^2$  y  $m$  correspondientes.

Datos	Métodos	Correlación	#	Desempeño	$R^2 / m$
PD	Beysens	Día a día	1	Malo	$R^2 << 1$
		Acumulado	2	Malo	$m = 0.102 \pm 0.078$
	RNA	Día a día	3	Marginal	$R^2 = 0.275 \pm 0.116$
		Acumulado	4	Bueno	$m = 1.033 \pm 0.051$
6 AM	Beysens	Día a día	5	Regular	$R^2 = 0.377 \pm 0.086$
		Acumulado	6	Bueno	$m = 1.189 \pm 0.257$
	RNA	Día a día	7	Regular	$R^2 = 0.427 \pm 0.119$
		Acumulado	8	Bueno	$m = 1.050 \pm 0.203$

Tabla 7.1: Conclusión de modelos

En resumen, este trabajo ha logrado desarrollar y comparar modelos matemáticos para predecir la cantidad de rocío. La RNA implementada ha demostrado ser eficiente y cumplir con los objetivos establecidos. Los resultados obtenidos son prometedores y contribuyen al campo de estudio, especialmente en la etapa de visualización. Estos hallazgos pueden ser de utilidad para abordar la problemática de escasez hídrica y establecer medidas de mitigación en diversas zonas afectadas. Este trabajo sienta las bases para futuras investigaciones y mejoras en los modelos propuestos, con el objetivo de brindar soluciones más precisas y efectivas en la predicción del rocío.

De acuerdo a la tabla anterior, en los experimentos realizados se observó que el modelo de D. Beysens no presenta un buen desempeño para el caso de promedio diario, día a día, ni promedio diario acumulado (número 1 y 2 de la Tabla 7.1). Por ejemplo, el caso promedio diario, día a día, generó un  $R^2$  despreciable, con respecto a 1, son de desempeño malo. Por otro lado, la predicción usando el modelo de RNA de datos PD, en el caso día a día obtiene un resultado marginal de  $R^2$  “0.102” y en el acumulado se obtiene un buen desempeño al  $m$  igual a “1.033”, entonces es cercano a 1 (número 3 y 4 de la Tabla 7.1) presentó un mejor modelo estimativo a comparación del desempeño anterior con el modelo de D. Beysens.

Considerando ahora los casos de estudio 6 AM, el modelo matemático de D. Beysens se comportó con un desempeño regular al utilizar el día a día ( $R^2 = 0.377$ ), en el caso acumulado obtenemos un buen desempeño ( $m = 1.189$ ), como se pueden divisar en la anterior Tabla en los números 5 y 6. La predicción realizada por el modelo de la RNA 6 AM, se comportó con un desempeño regular al utilizar promedio diario en donde se obtiene un  $R^2$  igual a 0.427 mejor que el día a día 6 AM de D. Beysens 0.377. En el caso de los acumulados obtenemos buenos desempeños, ya que en el de RNA 6 AM obtenemos un  $m$  igual a 1.050, más cerca de 1, en Beysens acumulado obtenemos un ( $m = 1.189$ ), Estos datos se visualizan en la Tabla en las líneas número 7 y 8. Sin embargo, destacamos número 4 de la tabla denominada como PD, de la columna de “RNA” y de fila “acumulados”, ennegrecido donde se obtuvo él ( $R^2 = 0.994$ ) y pendiente ( $m = 1.033$ ), demuestra que en el caso de solo considerar datos Promedio diario, podíamos ocupar este modelo para predecir el flujo de agua.

Los problemas más desafiantes es el enfrentar nuevas tecnologías fueron: las variables climatológicas, filtros de datos brutos, la predicción de modelos matemáticos no lineales y la poca información en la literatura referente a este fenómeno en particular. La limitación del proyecto es la cantidad de datos en función del tiempo y poca información, el difícil modelamiento de los modelos estimativos.

la primera instancia es implementar mediante un script en Python desde la página web de la NASA Worldview, para obtener el parámetro de nubosidad de forma automatizada, es decir el generar un software que descargue el video y luego separe en imágenes, que son guardados como contenido estático, permitiéndonos un generador de imágenes desde satélites de la NASA.

En una segunda instancia añadiría una API REST con tecnología de Flask para emplear nuestro modelo y realizar estimaciones día a día referente al ingreso de parámetros de entrada mediante una página web HTML, con el fin de poder utilizar estos modelos guardados y en un futuro implementar el condensador de rocío tal y como lo evidencia la documentación de Lekouch a partir de la estimación del modelo generado por la RNA, similar al tipo de pronóstico anticipado que ocurre en las lluvias.

# **Capítulo 8**

## **Anexo 1**

### **8.1. Implantación**

#### **8.1.1. Requerimientos**

##### **1. Requerimientos mínimos:**

Los requerimientos mínimos para efecto del experimento son los que se mencionan en el siguiente recuadro:

Requerimientos mínimos	Hardware o Software
Base de datos	El programa MySQL
Aplicación	Jupyter Notebook (Python 3)
Memoria RAM	4 GB de memoria RAM
Sistema Operativo	OS Windows 8, 10 o Mac de 64 bit
CPU	Se recomienda con intel i5/i7/Xeon
Cuenta del repositorio	Cuenta en GitHub
Disco duro	50 GB

Tabla 8.1: Requerimientos mínimos

##### **2. Requerimientos de Recomendados**

Los requerimientos recomendados facilita el efecto del experimento, cada vez que un laptop tenga mayor característica de sus componentes implica en el tiempo de ejecución de los algoritmos es menor, estos serán presentados en la siguiente tabla:

Requerimientos mínimos	Hardware o Software
Base de datos	El programa MySQL
Aplicación	Jupyter Notebook (Python 3)
Memoria RAM	Más de 8 GB de RAM
Sistema Operativo	OS Windows 10 o Mac de 64 bit
CPU	Se recomienda quad-core o hexa-core Intel i7/ Intel i9/M1
Cuenta de repositorio	Cuenta en GitHub
Disco duro	50 GB

Tabla 8.2: Requerimiento recomendado

### 3. Requerimientos Utilizados

Requerimientos	Hardware o Software
Base de datos	El programa MySQL
Aplicación	Jupyter Notebook (Python 3)
Memoria RAM	4 GB
Sistema Operativo	Ubuntu 20.04
CPU	CPU (quad-core)
Cuenta de repositorio	Cuenta en GitHub
Disco duro	50 GB

Tabla 8.3: Requerimiento Utilizados.

### **8.1.2. Preparación de Ambiente**

La preparación del ambiente se realizará de acuerdo a la implementación por medio del software y hardware utilizado. En un principio debemos tener una cuenta creada en GitHub, para guardar nuestro códigos y resultados, luego instalar Docker en conjunto con Docker-compose, con el fin de desplegar 3 software en un contenedor de forma aislada, que utilizaremos para desarrollar nuestro experimento y preparación de nuestro ambiente.

1. El primer paso es instalar Docker y Docker-Compose.

<https://docs.docker.com/desktop/install/ubuntu/>

2. El segundo paso es descargar el código mediante el siguiente comando:

`git clone https://github.com/Rodrigomontenegrofarias/Result-Model.git`

3. En el tercer paso, es fundamental modificar las variables de entorno disponibles para que se direccionen correctamente hacia nuestros puertos, servidores y servicios. El archivo que requiere edición se nombra ”./env”, además de las variables de MySQL, puede encontrarse en los siguientes dominios.

Sitio GitHub de archivo variable ENV.

Sitio GitHub de archivo variable base de datos.

4. En el cuarto paso, procedemos a ejecutar la sentencia para Nginx, la cual implica la creación de una red externa. Puede encontrar más detalles sobre este proceso en el enlace siguiente.
5. En quinto paso, para ejecutar el archivo docker-compose.yml, como se visualiza en la siguiente figura 8.1, se desarrolla a partir del comando para generar nuestro contenedor que contiene un sistema operativo Ubuntu 20.04 para implementar nuestro Jupyter Notebook, MySQL y Nginx:

```
docker-compose up -d
```

Figura 8.1: Ejecución Docker network y Docker compose.

## 8.2. Manual de usuario

En el manual de usuario explicamos el paso a paso, en el siguiente listado.

1. El primer paso se realiza luego de haber preparado el ambiente, garantizamos que cada servicio esté corriendo en el puerto indicado, con el siguiente comando, el cual podemos visualizar en la imagen 8.2:

docker-compose ps

```
root@Ubuntu-s-2vcpu-4gb-nyc3-01:/proyectos/Modelo-tesis# docker-compose ps
WARN[0000] network default: network.external.name is deprecated. Please set network.name with external: true
WARN[0000] network red_proxy: network.external.name is deprecated. Please set network.name with external: true
Name                  Command               State            Status          Ports
modelo-tesis-myapp-1 "jupyter notebook"   up 15 hours    Up 15 hours   0.0.0.0:8888->8888/tcp, :::8888->8888/tcp
mysql                "docker-entrypoint.s..." up 15 hours    Up 15 hours   0.0.0.0:3306->3306/tcp, :::3306->3306/tcp, 33060/tcp
nginx-proxy          "jwilder/nginx-proxy"  up 15 hours    Up 15 hours   0.0.0.0:80->80/tcp, :::80->80/tcp, 0.0.0.0:443->443/tcp, :::443->443/tcp
```

Figura 8.2: Listado de aplicaciones con el comando Docker compose ps.

2. Segundo paso es visualizar que esté corriendo correctamente en la página web, en mi caso por la IP de mi servidor es en la ruta, como se distingue en la imagen 8.3:

<http://159.203.136.164:8888/>

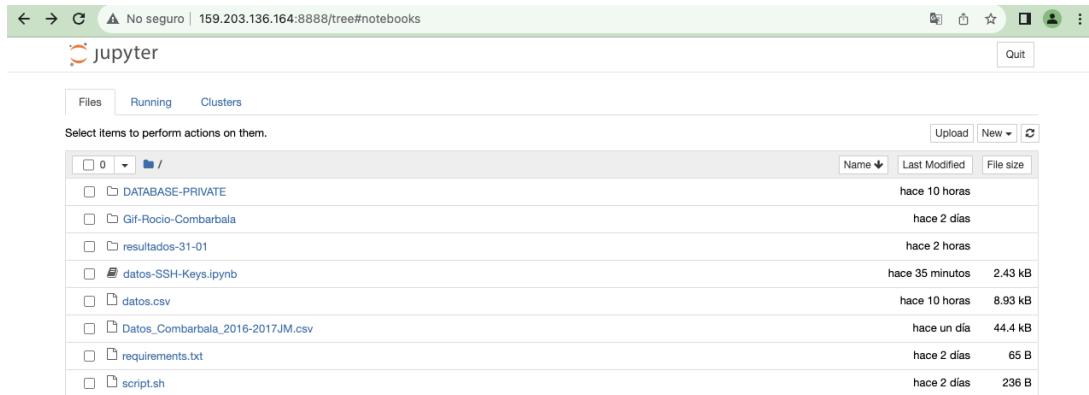


Figura 8.3: Jupyter Notebook, IP y puerto.

3. El tercer paso es ejecutar el archivo datos-SSH-Keys.ipynb, el cual nos permite lograr visualizar las contraseñas de keys para realizar la conexión con GitHub y ejecutar el archivo script para la descarga de los experimentos a realizar, tal cual como se muestra en la siguiente representación 8.4.

```
In [2]: cat ~/ssh/id_rsa
-----BEGIN RSA PRIVATE KEY-----
MIIEJjwIBAAKCAgEAtewHluEEUlxnVckB8Bu0yY31C0Sv4Cd5Yot6h0mK8eU1cXHHk
vPk0uZwgIEInTNEUdugptqmUu21qEmbjuogh0Pbjkl4Y1KUCD1/0IFUG7ccnf4
Eg84U0t0KB2SpRej/KKbjjCg45mjk/UsznkohP50FPR-PIYN0uQkrls/15/C1X31w0
2W7ph1h3r9P01P+kFXakxA0rDT4PoPuFyK+UozLeFZTAXt51dLMSrnYKFh35
vw6ZGBWPmgrrr0V8hnmAM05eG0YjNmYxH7eXTN/V23287+uMgsP000D2RzP/
mGsh2tBh2iirUo2BzsSMqR0e0N3hH4007PEoMcBsji4U45Os/eDzLj8Xn61T
1os+9n68LYp4ZPu0bHLj2bs8+a2B0M1ETxraLzM63oawgwun3mrzmz75A097k
zfwf32kzLwE41c0obpr0x04V1F+6R+ZH08L5bwMqag5Cg2XdbR+m35Z1kj
YaxWHimdmKASV+gsg+0Aq8ckFF5MVPC650vVg9p//Mu00p0p0vv551Vrygfavvh
v3yKLmgYErB0nqTfY0NLcLb7dVwq0lPHJ3qrm51pVz/Fgwkp0J1YDC0yQck
xUc7g/dh06x8dkVTu19G10mgT1Z228xhuvC66nNzUg40ubwMnp-Ppi+cAwEA
AOQkCAG4DRGIEpwUwFUoGc/lnD7G80Y4F82MKh9w8SNyvCK/YfJmbQB085Jj015A
RsB2umEB8BeTB04pHuq+A1ZwN/7aaeUHaPj89K6aa70Mwv0bbj/z/y0hA1D
2aYvfjoVHULsgVhdekg16uvhQ07EVxg0g9xAcu1mTe9v0kdBQUSyUf0f5IRs80
bdmSa0pGKxs9017n2620tzmKS9dMpeCb9loqK3jj/wG0sZcf2X971y5AXIea
bjj3Ksm0dKRSzbe475+Kxmn6Mpzkjtua-Zpf80JF51gFoHcN1Ew+ZBd9Yc/Vde
EY7Yo/b6HBSIMreIM5sYYzE/f6jkh3a4KK7+1h7r7j1r3mHfN22XueN5pwEE
9pe2fMGy2KhbmwxLysVkhZomhSkvKcT19rls+InkkMet2Dvhdb8d0jV92wZ2tb9
wBzHusb5b5vEnw40LX8beFca9g1xbUQoX17nHxRgmjeFAEmAkvjug8vzNgHl
UU01xh10m0xSxE6L68J0+BAfqlrcw1kgvSA@0bbyy4400s.ZsJXV2MdfrNgwXEpvn
fa335U0fkMDMUenUrjhjz28HR8j+h-6J3E9nARK65a7e8v+nbohc
2fbNbLu+01mtCfmTPxdJ35030rANdUzu0B6/0bYU1MMPJx0oKCAQE/RTyg/fY
W92nNQB35s0TgnRtnisuRucnchFsY0e/j/cUx24K1739GnPcvhUyqL3oZGTLc18
33GjneptFs/WWIPcAMPzCg0WLEbaAS0gUvFPkjtwRX8M/307677Gggs487+RgngZKS
mcNG5LDha70ajL2FQdwhpGwqEwPmbj1Kf/ahNCntSuHBcC4Cvg
YMDP0E1+2WcS90DR/T22T44t+nxjij16K-/T63a60rRAxvC10h0bcobXbw4xcerx
G1AV1LQb917ngIaaLFcuUc31Ew2p48d22fMceyv2CfzHDNqsPn/uJm0qcrXH109n
Wf1bh0uBf97qKCAQEAE8Pbry-musTTIBZaNm/qM1g/L0/n1910/0PA4ExIR65/wAoPzbqXPM6L
QICRYung#40ZmQ1wHeaf1nV8SSazGg8t/E54bcAcrhxzqMj+A+lwpRCTc440J
SEK11yh2s32WYtHv35FR0UK4drnN-Hj/q8pxx/AqbbAxr09yc/WSarj/v50715
Nn2e0NxZ0//aja0eeew06a1z7VqgWlUaqCs0e5Gh6J7QwH2j5M8qxa0j1HfAcY
ihhwgSPekqpRtv0xApKcy1+6+nByYSk6wmn20Ld0dwKCAQaQnt41CEbl03Im@mcu
efYtmztU77tzoyitKfw/te40RAjieewmbd7tRyUgj1-HYVY0uE196GYwmbQj9w
c48PTh6vc/I8cRhQ17VUlwTKbhlKgpIn77h40A2fAyipj+IVNetgPa@0ItzV
jk2HxKwJ4L8gTuVndA8n0mu+C1ejnfe4w5iXfd0kn60TQjn55LsqFDUUDng7
Zddh6SucwL7b55UWZhu0hdoh2/nx1+6<+bw6ezn1b5389PpJRUvFaonsaYn/zb
zAoSzCubZMT0B38Zw0Sa-WFvEwL409e5xR5Y6rF5y6rFcbgHVYrtuAKRp+u51Rv+L
HTrRa0IBAX21YL0aiNf7365tAFLByp-X0Tl12wSMYg/Pri8mU0M1191ZdW4pd2
PtNx/00gqj2Ymkmz7S.lapdh0vU0jibks547Cizh2dwjGTqjxrcQ2cdyHc/Z1k
m39XY0Y62yfjhzKackKyinVaC169frwPaz150ab0Panin/aj8u0My0Rm2Nn
E0u21yrUSdho/m/RLrjhjeafh1LGa3+1KYwkrK30wysaYjXsfZGVXR+sLTbAVdm+
xV0TImxdrk2zx65cqjDhkdpqcbjpR065jwQDC4BE70110/U74nDtYdLn2CjgbQ
oaWbz2n1Z3/14Vctv8wam9uXXz/EPUNELggAEJasnbWNHnJ35j/yE9h/FktH00cF
5xR015r0XPCvY1bmfWMNPcY6XhQp4YKVJ8Pveea02nf7qikRlcYf0jMhpLjsSLX
E0g410/20qht-Ef0fZ0K7k1z7vTe63DC25j/mz19IuviFa/dzEoGAXavGh84H/dV
g3WYh21lhxZ1C2wGvarzDoblnNOxbE6s3UKxxRR95v0xXY80nsQ+zmzjkRz21
o1h0x90/0U+MK5xFsgJhoqfMYYg+Ld87s or40AIHBMWNKj1eODU3WHD0IJ5
yh3se4xXcioXb8Xeg6Zwbd/gchHarUV6C1vsbjjQ8v+eRx1d120M/0PO/g===
-----END RSA PRIVATE KEY-----
```

```
In [3]: cat ~/ssh/id_rsa.pub
-----BEGIN RSA PUBLIC KEY-----
ssh-rsa AAAAB3NzaC1yc2EAAQABAAAQd0t5aE4QRQzKUQEG47jfUJBk/gJ3lii3qHSYrxTvxce08+T55nCAQidM0sRR26mC2qbW5Tyi
o5Zu061ChQ8G60vNjUpQ1O0/xgV0btxy/Tg5d2h5M60fLK1f6P8rE60MKdjma0t9730cq1E/k49H48hg35SCSwuL8VfchbTZYdumKlev0860/6
QVdorEBCs01Pg+Cg+5+5Taj0UQV1MBe2LmMszmudg0wHf/DpkYfwmacu7RUGGYBszT14zhCM2zjEfVrDm39UnYnzv7EsawW9dQ4PZfnn+Ya
yha0GfLwtTagH0IyPe5g3eHofj07s85goxwGyN0LhTjk0x78J1nUydrfVpWiz72frwlinhk+7PrcsnNt6zz5YGV02U51e1qXWlbehrCDC6fez
Gb0bvKA73u7N9Yxfat0vB57jVzShumtDEPhWix7pFh5kTTwlvDcyaqjDkKDpZ1f6bflmKsNgDfYeKZ2yoBJx6CyD44Crxy0UXkxU8r0LnRWD2n8
x1475k6m+9LkhWvKp+Fq+G/fiouuapgRutCe1Nh/0twtvt2/3LCp2w8cneplGbnU+9n+d8bCSm84mJgMLTJByTFRzuD92FDrFfx2Rv0630aLsaeB0
VllTzGEq681Bqc1m6rjRvAyen4+k5w== rodrigo.montenegro@alumnos.uv.cl
-----END RSA PUBLIC KEY-----
```

Figura 8.4: Archivo de datos SSH Keys.

Con el fin de relacionar git con nuestro repositorio Jupyter, colocaremos el token en GitHub en como se visualiza en la siguiente imagen 8.7.

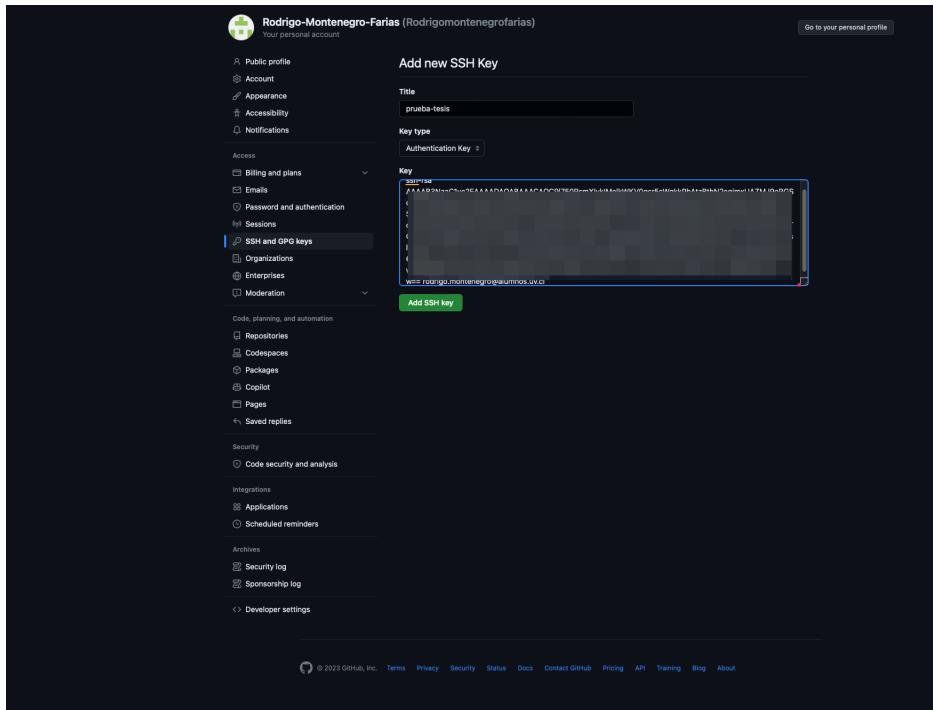


Figura 8.5: Conexión GitHub y Jupyter Notebook.

4. El cuarto paso implica la realización de pruebas para cada experimento. En este caso, se llevaron a cabo un total de 10 pruebas utilizando datos aleatorios variados, los archivos de las pruebas se divisa en la siguiente imagen 8.8. Una vez completado el último experimento número 10 que se visualiza en la figura 8.9, los resultados de estas pruebas se almacenan en una base de datos MySQL, tal como se evidencia en la siguiente imagen.

Figura 8.6. Colaboración GitHub y Jupyter Notebook.

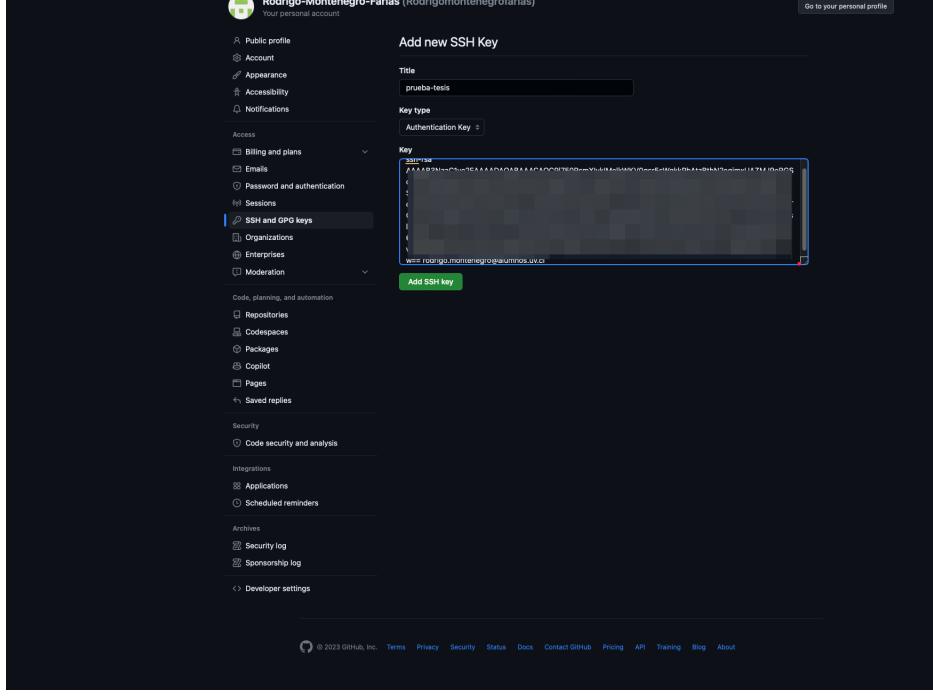


Figura 8.7: Conexión GitHub y Jupyter Notebook.

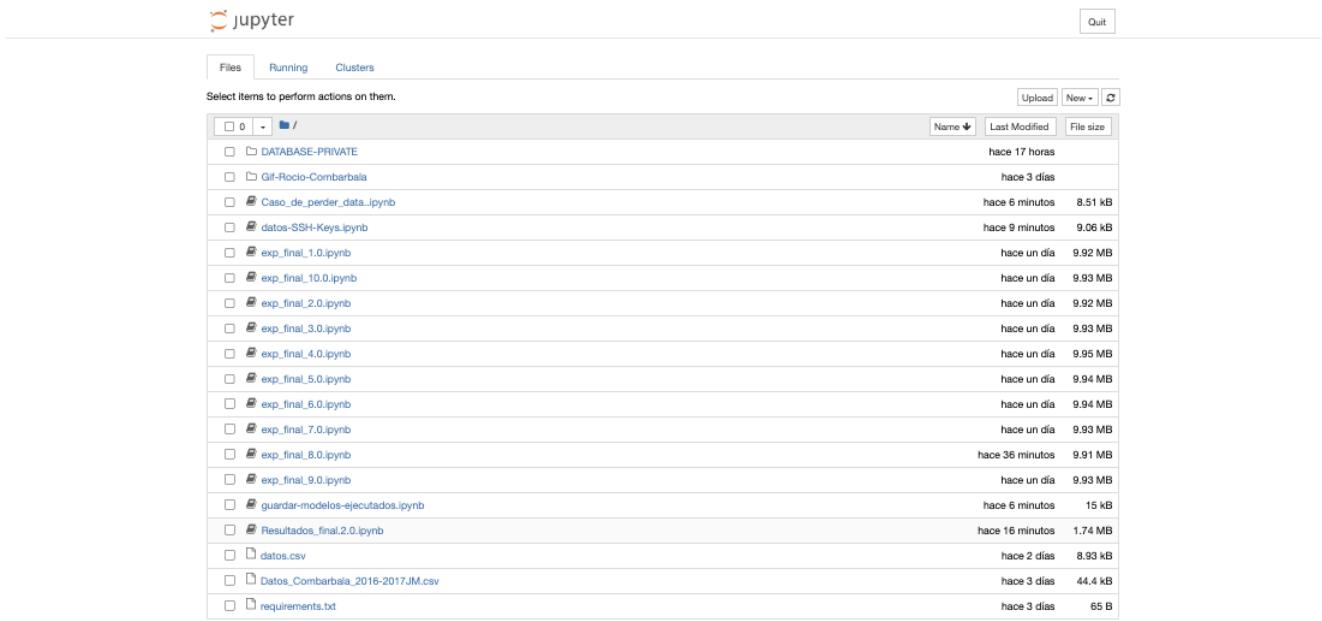


Figura 8.8: Archivos del repositorio para efectuar los experimentos.

```

jupyter exp_final_10.0 Last Checkpoint: hace 2 minutos (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Trusted Python 3
In [1]: ls
DATABASE-PRIVATE/
Datos_Combarbala_2016-2017JM.csv
datos.csv
exp_final_1.0.ipynb
exp_final_1.0.ipynb
exp_final_1.0.ipynb
exp_final_1.0.ipynb
exp_final_1.0.ipynb
exp_final_2.0-copy1.ipynb
exp_final_2.0.ipynb
exp_final_3.0.ipynb
exp_final_4.0.ipynb
resultados-31-01/

In [2]: cd ..
/root

In [3]: cd ..
/

In [4]: rm -rf DATABASE-PRIVATE/ Rodrigo(n)/ Rodrigo(n)/ Rodrigo(result/ resultados-modelos-tesis/ Guardar-notebook-jupyter)

In [5]: ls
bin/ lib/ opencv_contrib-3.4.5/ resultados/ srv/
boot/ lib64/ opencv_contrib.tgz resultados-31-01/ sys/
dev/ media/ opencv.tgz root/ tmp/
etc/ mnt/ opt/ run/ usr/
home/ opencv-3.4.5/ proc/ sbin/ var/

```

**Git clone database private.**

```

In [6]: import subprocess
repository_url = "git@github.com:Rodrigomontenegrofarias/DATABASE-PRIVATE.git"
# Obtener la huella digital de clave pública y clonar el repositorio
cmd_get_fingerprint = ['ssh-keyscan', '-H', 'github.com']
cmd_clone_repo = ['git', 'clone', repository_url]

try:
    fingerprint_result = subprocess.Popen(cmd_get_fingerprint, stdout=subprocess.PIPE, stderr=subprocess.PIPE)
    fingerprint_output, fingerprint_error = fingerprint_result.communicate()

    if fingerprint_result.returncode == 0:
        fingerprint = fingerprint_output.decode("utf-8")

        with open("/root/.ssh/known_hosts", "a") as known_hosts:
            known_hosts.write(fingerprint)

        subprocess.run(cmd_clone_repo, check=True)
        print("Repositorio clonado exitosamente.")
    else:
        print("Error obteniendo la huella digital:", fingerprint_error.decode("utf-8"))
except subprocess.CalledProcessError as e:
    print("Error:", e)

Repositorio clonado exitosamente.

```

```

In [7]: n=float(10)

In [8]: ls

```

Figura 8.9: Experimento número 10.

5. Quinto paso, es realizado una vez que ejecutamos nuestros 10 experimentos, para concluir ejecutaremos el Notebook de “resultados”, como se visualiza en la siguiente imagen 8.10, con el fin de obtener las métricas totales del experimento y ambos modelos matemáticos de predicción. Con esto originamos los experimentos de forma aleatoria y podemos comprobar cuál modelo es el mayor en términos de correlación, desviación estándar, error y pendiente.

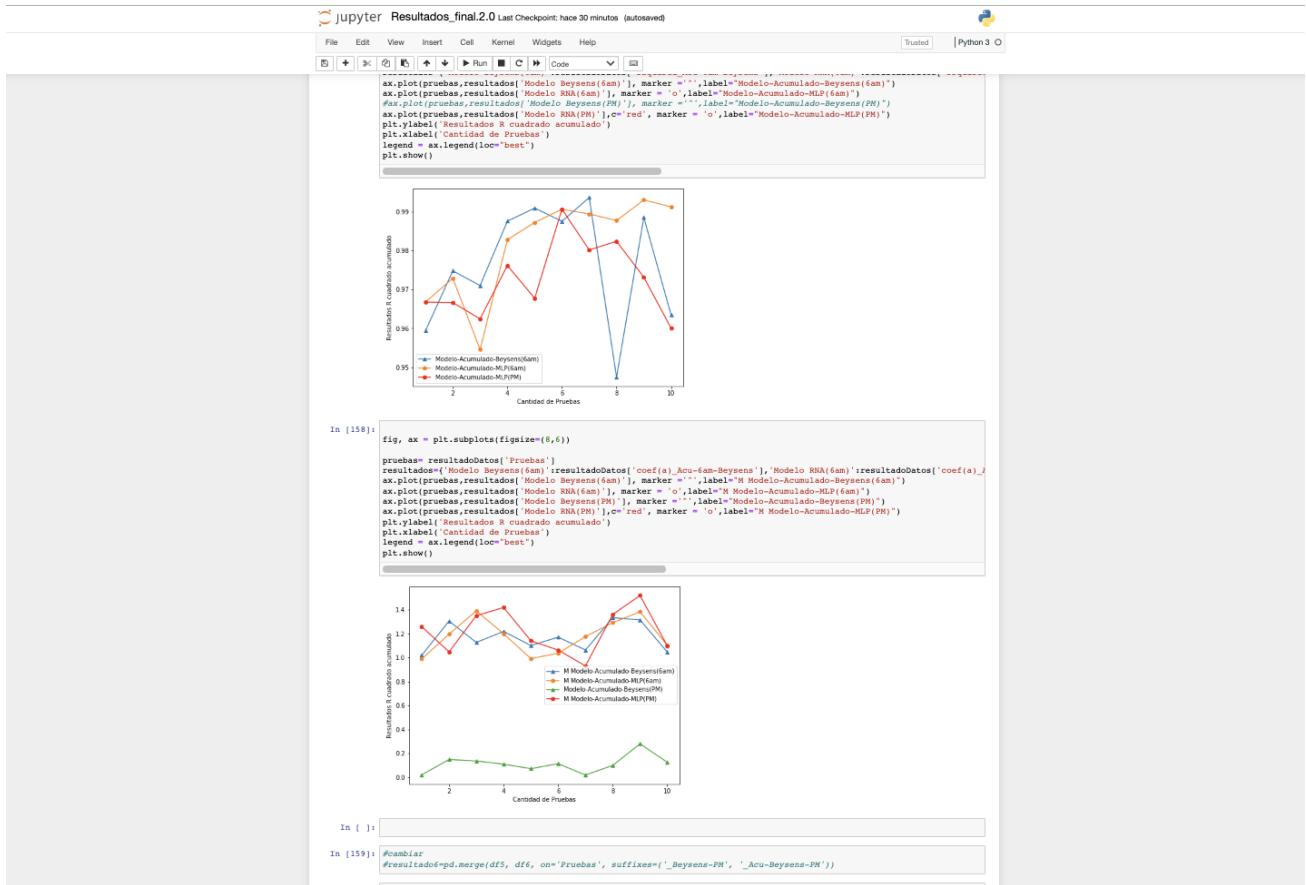


Figura 8.10: Imagen de conexión GitHub y Jupyter notebook.

## **Anexo 2**

### **8.3. Procedimiento de la captura y correlación de la variable de nubosidad utilizando la página web Worldview, proporcionado por la NASA**

El presente segmento describe el procedimiento desarrollado para capturar el valor numérico de nubosidad por medio de imágenes, estas son obtenidas a partir de los videos descargados en la página web de la nasa. Este procedimiento se realiza con la finalidad de crear e implementar un programa automatizado denominado script, para comprobar la correlación o  $R^2$  de los datos calculados por el script y los datos experimentales.

a continuación adjuntamos el diagrama que nos indica el procedimiento desde que ingresamos a la página web “<https://worldview.earthdata.nasa.gov>” provista por la NASA. En donde Se utilizó la opción de añadir capas (layers) y se seleccionó el parámetro de nubosidad, específicamente “cloud fraction”, indicando la fecha deseada, permitiendo descargar Los videos que contienen información sobre la nubosidad capturada por el sensor MODIS Cloud Fraction del satélite Terra (MOD06), como se puede visualizar en la siguiente figura 8.11.

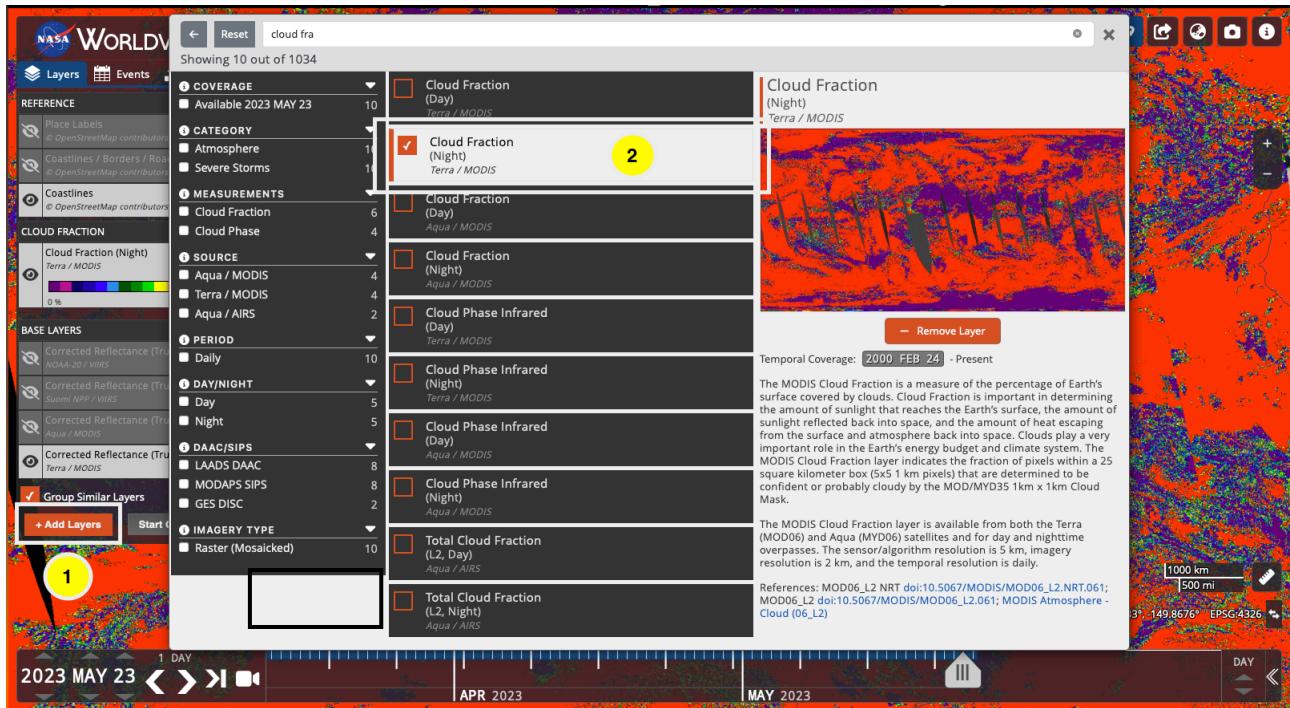


Figura 8.11: Imagen de la página Web la NASA.

Estos videos presentan una resolución espacial de 1 km y una resolución temporal diaria. Una vez descargado el video es sometido al script, el cual en un principio divide el video en frame, una imagen por cada día. Luego estas imágenes son sometidas a una función que nos permite obtener el promedio de 32 píxeles de los colores de cada imagen, para garantizar un buen resultado numérico de nubosidad. Para finalizar el script genera un gráfico de correlación de los datos experimentales calculados de forma manual con el equipo de Danilo Carbajal y los datos numéricos obtenidos a partir del script, se evidencia en el siguiente diagrama 8.12.

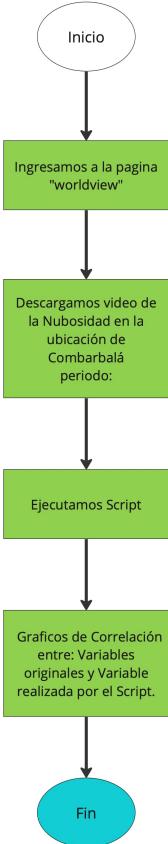


Figura 8.12: Diagrama para obtener la nubosidad desde la página web Woldview.

El script divide o separa los vídeo en imágenes individuales, generando una imagen de 36 píxeles por cada día, donde eliminaremos los 4 píxeles de cada esquina, el cual nos permite mantener la forma circular de interés, como se puede divisar en la siguiente Figura 8.13.

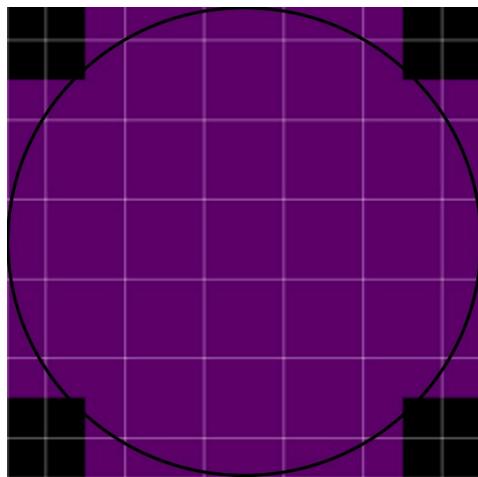


Figura 8.13: Imagen o pixel del color asociado a la nubosidad, ubicado en Combarbala.

Luego el script une todas estas imágenes segmentadas del párrafo anterior y procedemos al cálculo del color promedio de los píxeles de cada imagen, tomando en cuenta 32 de los 36 píxeles obtenidos, el color es representado en una escala del 0 al 100 % que indica la cantidad numérica de nubosidad por día, al finalizar el script genera gráficos de correlación entre el script calculado de forma manual y los automatizados por el nuevo código implementado, como podemos visualizar en el siguiente diagrama 8.14.



Figura 8.14: Diagrama del script.

Adjunto la evidencia del gráfico de correlación de la Nubosidad calculada manual y la del script, originalmente de los primeros 4 meses que equivalen a 125 días, obteniendo un  $R^2$  igual a 0.933, con  $m$  igual a 0.973 y error estándar de 0.024, tal cual como se evidencia en la siguiente figura 8.15.

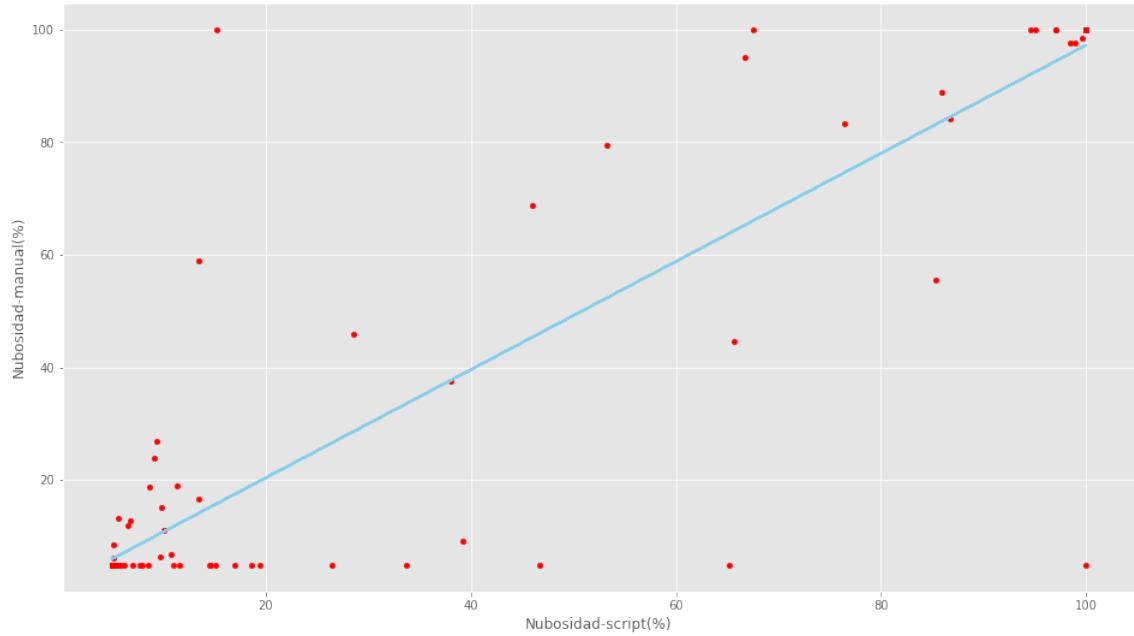


Figura 8.15: Imagen de correlación de nubosidad-script y nubosidad-manual, de un  $R^2$  igual a 0.933.

## **Anexo 3**

A continuación se presenta el modelo de la RNA que utiliza el entrenamiento del algoritmo ADAM, adicional también usaremos los casos de estudios en donde podemos visualizar.

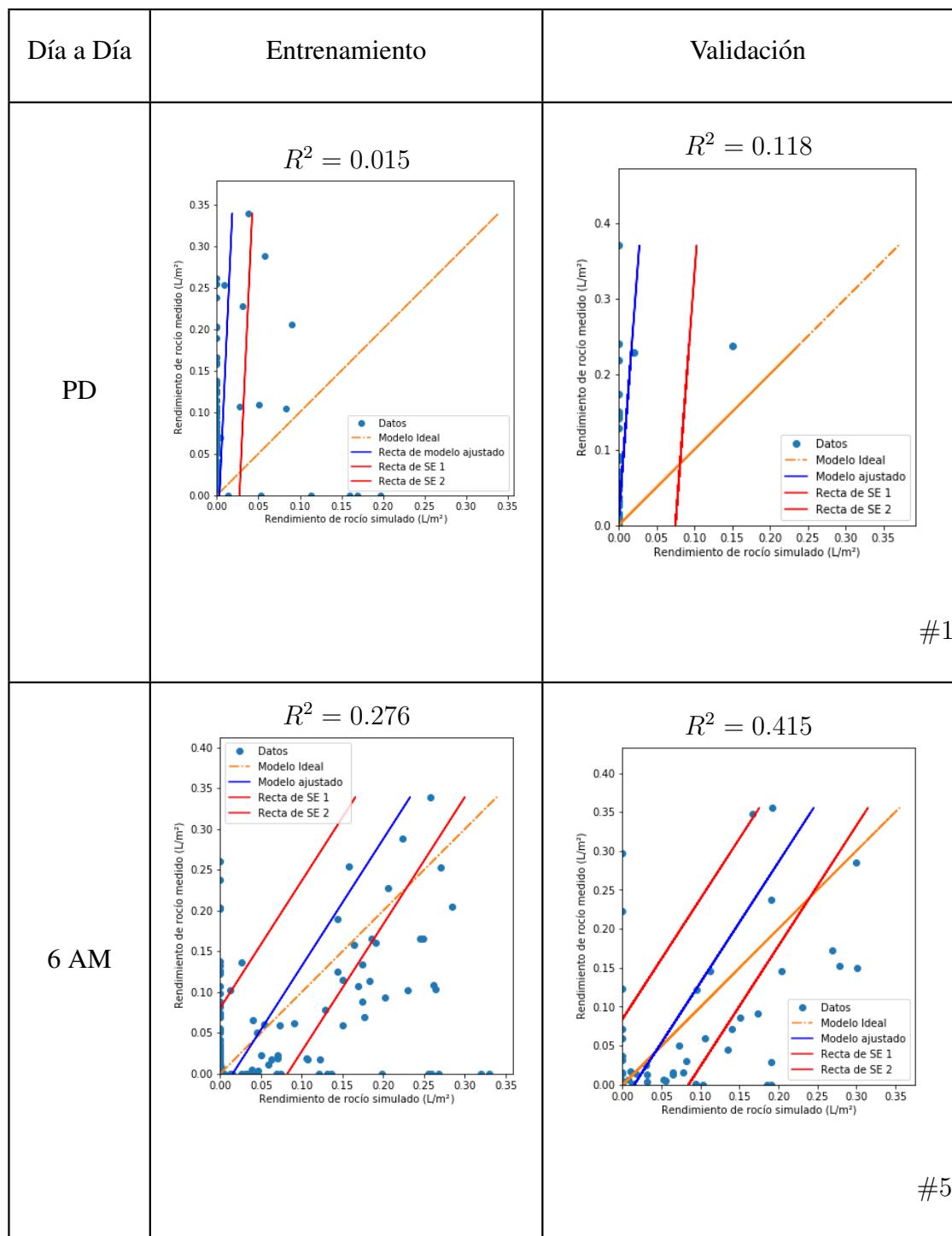


Tabla 8.4: Tabla de modelo D. Beysens, experimento número 1

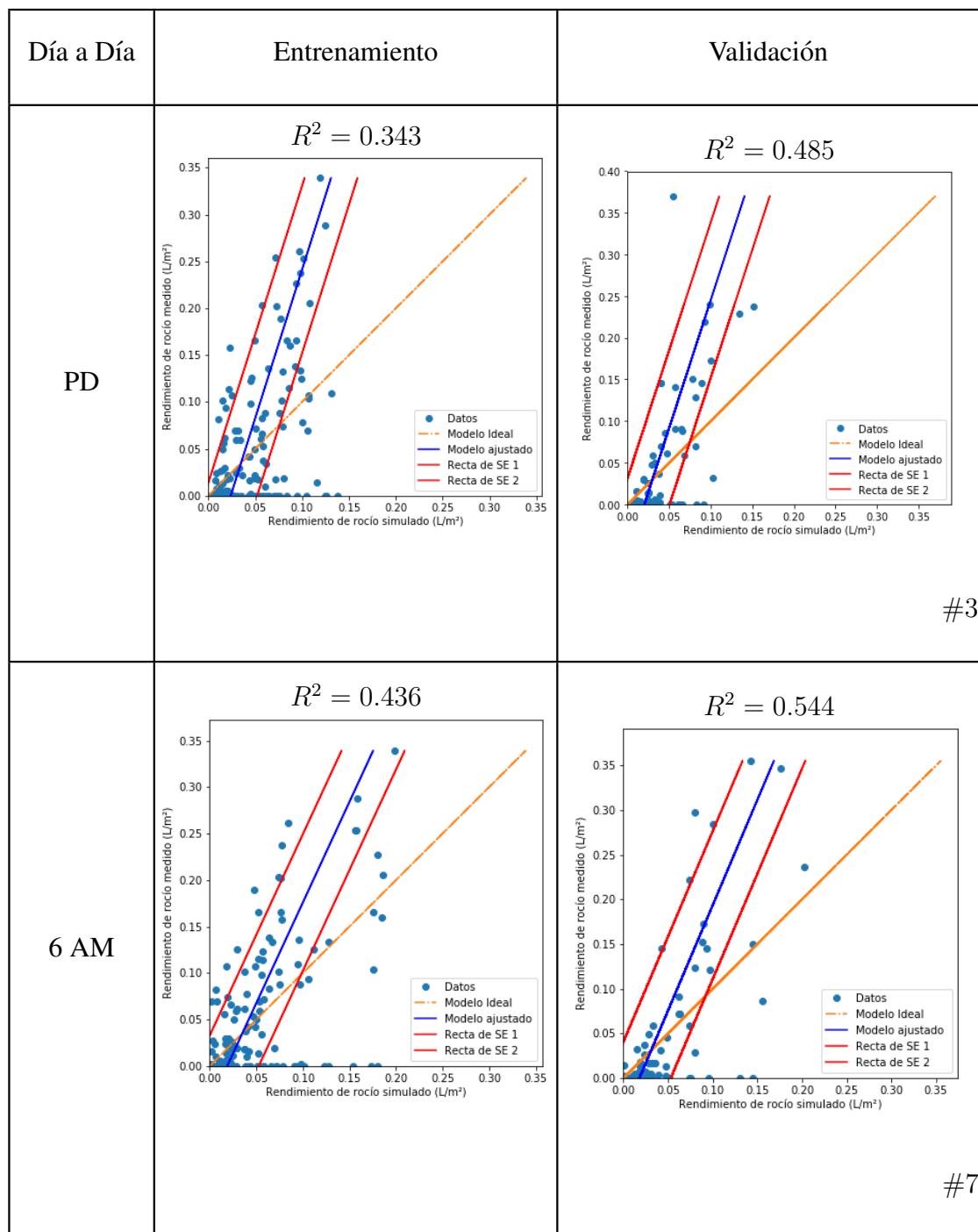


Tabla 8.5: Tabla de modelo RNA, experimento número 1

Acumulado	Entrenamiento	Validación
PD	<p><math>R^2 = 0.951</math></p> <p>Datos Modelo Ideal Recta de modelo ajustado Recta de SE 1 Recta de SE 2</p>	<p><math>R^2 = 0.467</math></p> <p>Datos Modelo Ideal Modelo ajustado Recta de SE 1 Recta de SE 2</p> <p>#2</p>
6 AM	<p><math>R^2 = 0.981</math></p> <p>Datos Medición Ideal Modelo ajustado Recta de SE 1 Recta de SE 2</p>	<p><math>R^2 = 0.954</math></p> <p>Datos Modelo Ideal Modelo ajustados Recta de SE 1 Recta de SE 2</p> <p>#6</p>

Tabla 8.6: Tabla de Acumulado RNA, experimento número 1

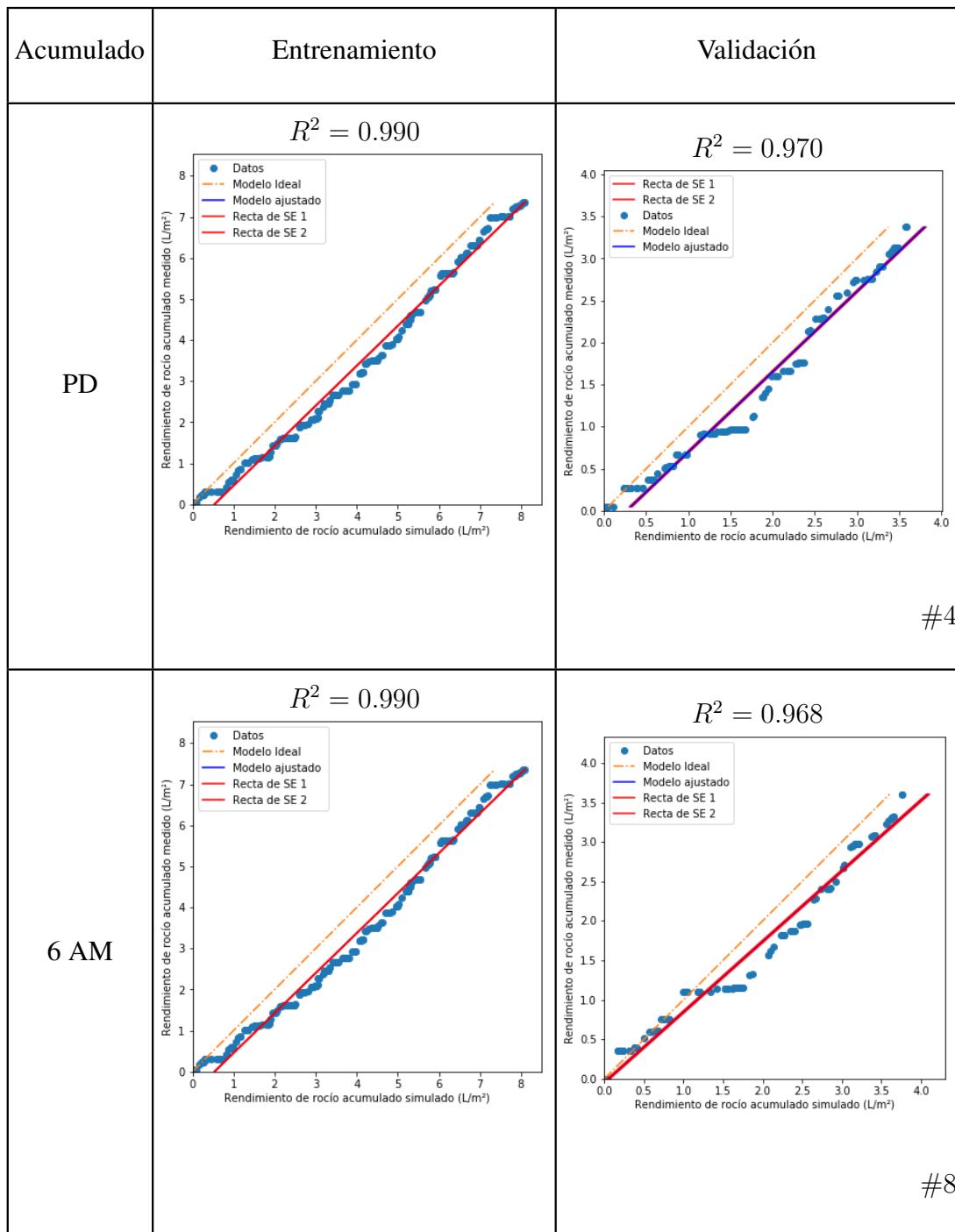


Tabla 8.7: Tabla de Acumulado D. Beysens, experimento número 1

<b>Caso 1 y 3</b>	<b>Beysens</b> $R^2 = 0.067 \pm 0.064$ <b>RNA</b> $R^2 = 0.295 \pm 0.092$	<b>Beysens</b> $m = 0.051 \pm 0.034$ <b>RNA</b> $m = 0.316 \pm 0.071$																																																																		
RNA y Beysens diario en horario PD	<table border="1"> <caption>Data for R-squared (R²)</caption> <thead> <tr> <th>Pruebas</th> <th>Beysens PD</th> <th>MLP PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.12</td><td>0.46</td></tr> <tr><td>2</td><td>0.02</td><td>0.18</td></tr> <tr><td>3</td><td>0.01</td><td>0.27</td></tr> <tr><td>4</td><td>0.17</td><td>0.29</td></tr> <tr><td>5</td><td>0.01</td><td>0.20</td></tr> <tr><td>6</td><td>0.01</td><td>0.21</td></tr> <tr><td>7</td><td>0.09</td><td>0.35</td></tr> <tr><td>8</td><td>0.15</td><td>0.28</td></tr> <tr><td>9</td><td>0.10</td><td>0.30</td></tr> <tr><td>10</td><td>0.03</td><td>0.35</td></tr> </tbody> </table>	Pruebas	Beysens PD	MLP PD	1	0.12	0.46	2	0.02	0.18	3	0.01	0.27	4	0.17	0.29	5	0.01	0.20	6	0.01	0.21	7	0.09	0.35	8	0.15	0.28	9	0.10	0.30	10	0.03	0.35	<table border="1"> <caption>Data for Slopes (Pendientes)</caption> <thead> <tr> <th>Pruebas</th> <th>Beysens PD</th> <th>MLP PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.08</td><td>0.33</td></tr> <tr><td>2</td><td>0.07</td><td>0.38</td></tr> <tr><td>3</td><td>0.02</td><td>0.44</td></tr> <tr><td>4</td><td>0.02</td><td>0.24</td></tr> <tr><td>5</td><td>0.01</td><td>0.28</td></tr> <tr><td>6</td><td>0.01</td><td>0.28</td></tr> <tr><td>7</td><td>0.01</td><td>0.25</td></tr> <tr><td>8</td><td>0.07</td><td>0.27</td></tr> <tr><td>9</td><td>0.07</td><td>0.25</td></tr> <tr><td>10</td><td>0.08</td><td>0.44</td></tr> </tbody> </table>	Pruebas	Beysens PD	MLP PD	1	0.08	0.33	2	0.07	0.38	3	0.02	0.44	4	0.02	0.24	5	0.01	0.28	6	0.01	0.28	7	0.01	0.25	8	0.07	0.27	9	0.07	0.25	10	0.08	0.44
Pruebas	Beysens PD	MLP PD																																																																		
1	0.12	0.46																																																																		
2	0.02	0.18																																																																		
3	0.01	0.27																																																																		
4	0.17	0.29																																																																		
5	0.01	0.20																																																																		
6	0.01	0.21																																																																		
7	0.09	0.35																																																																		
8	0.15	0.28																																																																		
9	0.10	0.30																																																																		
10	0.03	0.35																																																																		
Pruebas	Beysens PD	MLP PD																																																																		
1	0.08	0.33																																																																		
2	0.07	0.38																																																																		
3	0.02	0.44																																																																		
4	0.02	0.24																																																																		
5	0.01	0.28																																																																		
6	0.01	0.28																																																																		
7	0.01	0.25																																																																		
8	0.07	0.27																																																																		
9	0.07	0.25																																																																		
10	0.08	0.44																																																																		
RNA y Beysens diario en horario PD	<b>Beysens</b> $b = 0.026 \pm 0.013$ <b>RNA</b> $b = 0.046 \pm 0.015$	<b>Beysens</b> $SE = 0.026 \pm 0.013$ <b>RNA</b> $SE = 0.046 \pm 0.015$																																																																		
	<table border="1"> <caption>Data for Coefficients (Coef)</caption> <thead> <tr> <th>Pruebas</th> <th>Beysens PD</th> <th>MLP PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.001</td><td>0.020</td></tr> <tr><td>2</td><td>0.005</td><td>0.026</td></tr> <tr><td>3</td><td>0.002</td><td>0.024</td></tr> <tr><td>4</td><td>0.001</td><td>0.022</td></tr> <tr><td>5</td><td>0.003</td><td>0.021</td></tr> <tr><td>6</td><td>0.002</td><td>0.030</td></tr> <tr><td>7</td><td>0.001</td><td>0.021</td></tr> <tr><td>8</td><td>0.001</td><td>0.021</td></tr> <tr><td>9</td><td>0.001</td><td>0.023</td></tr> <tr><td>10</td><td>0.004</td><td>0.026</td></tr> </tbody> </table>	Pruebas	Beysens PD	MLP PD	1	0.001	0.020	2	0.005	0.026	3	0.002	0.024	4	0.001	0.022	5	0.003	0.021	6	0.002	0.030	7	0.001	0.021	8	0.001	0.021	9	0.001	0.023	10	0.004	0.026	<table border="1"> <caption>Data for Standard Error (SE)</caption> <thead> <tr> <th>Pruebas</th> <th>Beysens PD</th> <th>MLP PD</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.020</td><td>0.032</td></tr> <tr><td>2</td><td>0.050</td><td>0.072</td></tr> <tr><td>3</td><td>0.035</td><td>0.060</td></tr> <tr><td>4</td><td>0.005</td><td>0.032</td></tr> <tr><td>5</td><td>0.035</td><td>0.052</td></tr> <tr><td>6</td><td>0.025</td><td>0.048</td></tr> <tr><td>7</td><td>0.030</td><td>0.032</td></tr> <tr><td>8</td><td>0.015</td><td>0.040</td></tr> <tr><td>9</td><td>0.020</td><td>0.035</td></tr> <tr><td>10</td><td>0.040</td><td>0.052</td></tr> </tbody> </table>	Pruebas	Beysens PD	MLP PD	1	0.020	0.032	2	0.050	0.072	3	0.035	0.060	4	0.005	0.032	5	0.035	0.052	6	0.025	0.048	7	0.030	0.032	8	0.015	0.040	9	0.020	0.035	10	0.040	0.052
Pruebas	Beysens PD	MLP PD																																																																		
1	0.001	0.020																																																																		
2	0.005	0.026																																																																		
3	0.002	0.024																																																																		
4	0.001	0.022																																																																		
5	0.003	0.021																																																																		
6	0.002	0.030																																																																		
7	0.001	0.021																																																																		
8	0.001	0.021																																																																		
9	0.001	0.023																																																																		
10	0.004	0.026																																																																		
Pruebas	Beysens PD	MLP PD																																																																		
1	0.020	0.032																																																																		
2	0.050	0.072																																																																		
3	0.035	0.060																																																																		
4	0.005	0.032																																																																		
5	0.035	0.052																																																																		
6	0.025	0.048																																																																		
7	0.030	0.032																																																																		
8	0.015	0.040																																																																		
9	0.020	0.035																																																																		
10	0.040	0.052																																																																		

Tabla 8.8: Tabla de descripción Casos de estudio 1 y 3 “Día a día PD” con las 10 pruebas

<b>Caso 2 y 4</b>	<b>Beysens</b> $R^2 = 0.754 \pm 0.006$ <b>RNA</b> $R^2 = 0.985 \pm 0.006$	<b>Beysens</b> $m = 0.985 \pm 0.080$ <b>RNA</b> $m = 1.016 \pm 0.074$
RNA y Beysens Acumulado en horario PD		
	<b>Beysens</b> $b = 0.000 \pm 0.094$ <b>RNA</b> $b = 0.091 \pm 0.094$	<b>Beysens</b> $SE = 0.004 \pm 0.003$ <b>RNA</b> $SE = 0.012 \pm 0.004$

Tabla 8.9: Tabla de descripción Casos de estudio 2 y 4 “Acumulado PD”

<b>Caso 5 y 7</b>	<b>Beysens</b> $R^2 = 0.377 \pm 0.086$ <b>RNA</b> $R^2 = 0.425 \pm 0.108$	<b>Beysens</b> $m = 0.688 \pm 0.132$ <b>RNA</b> $m = 0.507 \pm 0.130$
RNA y Beysens diario en horario 6 AM		
	<b>Beysens</b> $b = 0.014 \pm 0.003$ <b>RNA</b> $b = 0.019 \pm 0.004$	<b>Beysens</b> $SE = 0.075 \pm 0.021$ <b>RNA</b> $SE = 0.047 \pm 0.010$

Tabla 8.10: Tabla de descripción Casos de estudio 5 y 7 Día a día en horario 6 AM

<b>Caso 6 y 8</b>	<b>Beysens</b> $R^2 = 0.976 \pm 0.014$ <b>RNA</b> $R^2 = 0.997 \pm 0.011$	<b>Beysens</b> $m = 1.189 \pm 0.145$ <b>RNA</b> $m = 1.170 \pm 0.282$																																																																		
RNA y Beysens acumulado en horario 6 AM	<p>Resultados <math>R^2</math> vs Cantidad de Pruebas</p> <table border="1"> <thead> <tr> <th>Cantidad de Pruebas</th> <th>Beysens 6AM (<math>R^2</math>)</th> <th>MLP 6AM (<math>R^2</math>)</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.955</td><td>0.970</td></tr> <tr><td>2</td><td>0.980</td><td>0.980</td></tr> <tr><td>3</td><td>0.950</td><td>0.970</td></tr> <tr><td>4</td><td>0.980</td><td>0.965</td></tr> <tr><td>5</td><td>0.990</td><td>0.995</td></tr> <tr><td>6</td><td>0.980</td><td>0.990</td></tr> <tr><td>7</td><td>0.990</td><td>0.995</td></tr> <tr><td>8</td><td>0.980</td><td>0.980</td></tr> <tr><td>9</td><td>0.970</td><td>0.985</td></tr> <tr><td>10</td><td>0.980</td><td>0.965</td></tr> </tbody> </table>	Cantidad de Pruebas	Beysens 6AM ( $R^2$ )	MLP 6AM ( $R^2$ )	1	0.955	0.970	2	0.980	0.980	3	0.950	0.970	4	0.980	0.965	5	0.990	0.995	6	0.980	0.990	7	0.990	0.995	8	0.980	0.980	9	0.970	0.985	10	0.980	0.965	<p>Resultados de pendientes vs Cantidad de Pruebas</p> <table border="1"> <thead> <tr> <th>Cantidad de Pruebas</th> <th>Beysens 6AM (Pendiente)</th> <th>MLP 6AM (Pendiente)</th> </tr> </thead> <tbody> <tr><td>1</td><td>1.15</td><td>1.12</td></tr> <tr><td>2</td><td>1.20</td><td>1.20</td></tr> <tr><td>3</td><td>1.45</td><td>1.75</td></tr> <tr><td>4</td><td>1.10</td><td>1.15</td></tr> <tr><td>5</td><td>1.20</td><td>1.15</td></tr> <tr><td>6</td><td>1.10</td><td>1.30</td></tr> <tr><td>7</td><td>0.90</td><td>0.75</td></tr> <tr><td>8</td><td>1.30</td><td>1.20</td></tr> <tr><td>9</td><td>1.05</td><td>0.90</td></tr> <tr><td>10</td><td>1.30</td><td>1.20</td></tr> </tbody> </table>	Cantidad de Pruebas	Beysens 6AM (Pendiente)	MLP 6AM (Pendiente)	1	1.15	1.12	2	1.20	1.20	3	1.45	1.75	4	1.10	1.15	5	1.20	1.15	6	1.10	1.30	7	0.90	0.75	8	1.30	1.20	9	1.05	0.90	10	1.30	1.20
Cantidad de Pruebas	Beysens 6AM ( $R^2$ )	MLP 6AM ( $R^2$ )																																																																		
1	0.955	0.970																																																																		
2	0.980	0.980																																																																		
3	0.950	0.970																																																																		
4	0.980	0.965																																																																		
5	0.990	0.995																																																																		
6	0.980	0.990																																																																		
7	0.990	0.995																																																																		
8	0.980	0.980																																																																		
9	0.970	0.985																																																																		
10	0.980	0.965																																																																		
Cantidad de Pruebas	Beysens 6AM (Pendiente)	MLP 6AM (Pendiente)																																																																		
1	1.15	1.12																																																																		
2	1.20	1.20																																																																		
3	1.45	1.75																																																																		
4	1.10	1.15																																																																		
5	1.20	1.15																																																																		
6	1.10	1.30																																																																		
7	0.90	0.75																																																																		
8	1.30	1.20																																																																		
9	1.05	0.90																																																																		
10	1.30	1.20																																																																		
	<b>Beysens</b> $b = 0.042 \pm 0.220$ <b>RNA</b> $b = 0.191 \pm 0.188$	<b>Beysens</b> $SE = 0.016 \pm 0.006$ <b>RNA</b> $SE = 0.015 \pm 0.064$																																																																		
	<p>Resultados de coeficientes vs Cantidad de Pruebas</p> <table border="1"> <thead> <tr> <th>Cantidad de Pruebas</th> <th>Beysens 6AM (Coef)</th> <th>MLP 6AM (Coef)</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.05</td><td>0.08</td></tr> <tr><td>2</td><td>0.10</td><td>0.22</td></tr> <tr><td>3</td><td>-0.28</td><td>0.05</td></tr> <tr><td>4</td><td>-0.15</td><td>0.12</td></tr> <tr><td>5</td><td>0.00</td><td>0.05</td></tr> <tr><td>6</td><td>0.45</td><td>0.48</td></tr> <tr><td>7</td><td>-0.12</td><td>0.35</td></tr> <tr><td>8</td><td>0.10</td><td>0.35</td></tr> <tr><td>9</td><td>0.05</td><td>0.18</td></tr> <tr><td>10</td><td>0.35</td><td>0.48</td></tr> </tbody> </table>	Cantidad de Pruebas	Beysens 6AM (Coef)	MLP 6AM (Coef)	1	0.05	0.08	2	0.10	0.22	3	-0.28	0.05	4	-0.15	0.12	5	0.00	0.05	6	0.45	0.48	7	-0.12	0.35	8	0.10	0.35	9	0.05	0.18	10	0.35	0.48	<p>Resultados Standard Error vs Cantidad de Pruebas</p> <table border="1"> <thead> <tr> <th>Cantidad de Pruebas</th> <th>Beysens 6AM (SE)</th> <th>MLP 6AM (SE)</th> </tr> </thead> <tbody> <tr><td>1</td><td>0.025</td><td>0.018</td></tr> <tr><td>2</td><td>0.015</td><td>0.015</td></tr> <tr><td>3</td><td>0.030</td><td>0.028</td></tr> <tr><td>4</td><td>0.012</td><td>0.018</td></tr> <tr><td>5</td><td>0.008</td><td>0.008</td></tr> <tr><td>6</td><td>0.014</td><td>0.012</td></tr> <tr><td>7</td><td>0.007</td><td>0.004</td></tr> <tr><td>8</td><td>0.008</td><td>0.012</td></tr> <tr><td>9</td><td>0.015</td><td>0.009</td></tr> <tr><td>10</td><td>0.016</td><td>0.020</td></tr> </tbody> </table>	Cantidad de Pruebas	Beysens 6AM (SE)	MLP 6AM (SE)	1	0.025	0.018	2	0.015	0.015	3	0.030	0.028	4	0.012	0.018	5	0.008	0.008	6	0.014	0.012	7	0.007	0.004	8	0.008	0.012	9	0.015	0.009	10	0.016	0.020
Cantidad de Pruebas	Beysens 6AM (Coef)	MLP 6AM (Coef)																																																																		
1	0.05	0.08																																																																		
2	0.10	0.22																																																																		
3	-0.28	0.05																																																																		
4	-0.15	0.12																																																																		
5	0.00	0.05																																																																		
6	0.45	0.48																																																																		
7	-0.12	0.35																																																																		
8	0.10	0.35																																																																		
9	0.05	0.18																																																																		
10	0.35	0.48																																																																		
Cantidad de Pruebas	Beysens 6AM (SE)	MLP 6AM (SE)																																																																		
1	0.025	0.018																																																																		
2	0.015	0.015																																																																		
3	0.030	0.028																																																																		
4	0.012	0.018																																																																		
5	0.008	0.008																																																																		
6	0.014	0.012																																																																		
7	0.007	0.004																																																																		
8	0.008	0.012																																																																		
9	0.015	0.009																																																																		
10	0.016	0.020																																																																		

Tabla 8.11: Tabla de descripción Casos de estudio 1 y 3 “Día a día PD” con las 10 pruebas

# Anexo 4

## 8.4. Códigos de implementación

El código a presentar es extenso, por ende adjuntaremos las URL del repositorio y algunas partes importante del código y estructura y script realizado para obtener el parámetro de nubosidad a partir de la página web Worldview.

<https://github.com/Rodrigomontenegrofarias/Modelo-tesis>  
<https://github.com/Rodrigomontenegrofarias/resultados-31-01>  
1) Código para Docker, para nuestro ambiente.

Listing 8.1: Ejemplo de código en Python

```
version: "3"
# Definir servicios
services:
  mysql:                      # Container para la base de datos .
    build:                      # Instrucciones para crear imagen
      context: ./sql           # Dóntexto actual
      dockerfile: Dockerfile   # Nombre Dockerfile
    restart: always
    mem_limit: 800m
    cpus: 0.5 # Limita la uso de CPU al 50%
    ports:
      - ${MYSQL_ROOT_PORT} # Nombre de la red asignad
    environment:              # Variables de entorno
      MYSQL_ROOT_PASSWORD: ${MYSQL_ROOT_PASSWORD}
      MYSQL_ROOT_HOST: ${MYSQL_ROOT_HOST}
      MYSQL_DATABASE: ${MYSQL_DATABASE}
      MYSQL_PASSWORD: ${MYSQL_PASSWORD}
    container_name: mysql # Nomnbre del container
  myapp:                      # Container para el myapp2
    build:                      # Instrucciones para crear imagen
      context: .               # Dóntexto actual
```

```

    dockerfile: Dockerfile # Nombre Dockerfile
    restart: always          # Indica que el container se reinicie siempre.
    ports:
      - ${JUPYTER_PORT}      # Varibale de la red asignada Port
    networks:
      - red_proxy            # Asignar una red al container
      - red_privada          # Nombre de la red asignada
    depends_on:
      - mysql
    environment:
      VIRTUAL_HOST: ${MYSQL_ROOT_HOST}# DNS asignado al container
    mem_limit: 3100m
    cpus: 0.95 # Limita la uso de CPU al 50%
  # servicio de proxy entre contenedores
  nginx-proxy:
    # comando para configurar Nginx
    command: "sh -c 'echo \"server_tokens off; proxy_connect_timeout 300; p
    restart: always          # reiniciar siempre
    image: jwilder/nginx-proxy # imagen desde docker HUB
    container_name: nginx-proxy # nombre del contenedor
    ports:
      - "80:80"
      - "443:443"
    volumes:
      - /var/run/docker.sock:/tmp/docker.sock:ro
  # Nombre de los volumenes son data , data1 , data2 , ... ,dataN segun correspon
  volumes:
    data:                   # Defenir volumenes
    data:                   # Nombre del volumen asignado al proyecto
  # El nombre de las redes no cambia (red_proxy y red_privada son estandar)
  networks:
    default:
      external:
        name: nginx-proxy    # Definir redes
    red_proxy:
      name: nginx-proxy     # Nombre de una red
    external:
      name: nginx-proxy     # Indica que es una red externa al docker-compose
    red_privada:
      name: nginx-proxy     # Nombre de la Red de uso publico entre Docker
      name: nginx-proxy     # Nombre de una red que depende del docker-compo

```

2) Cálculo formula matemática Beysens.

Listing 8.2: Ejemplo de código en Python

```

import math
import numpy as np

H = 0.9
H_2 = H**2
V0 = 4.4
a = 17.27
b = 237.7

c = np.log(train_dataset3ModeloMatematico["Outdoor_Humidity"
    ("\\%")])
alpha = a * ((train_dataset3ModeloMatematico["Outdoor_"
    "Temperature("\\degree_C)"]) / (b +
    train_dataset3ModeloMatematico["Outdoor_Temperature("\\"
        "degree_C)"])) + c
Td = (b * alpha) / (a - alpha)

exp20 = np.exp((-((train_dataset3ModeloMatematico["Wind_"
    "Speed(m/s)"]) / V0) ** 20))

A = (0.37 * (1 + (0.204323 * H) - (0.0238893 * H_2) - \
(18.0132 - (1.04963 * H) + 0.21891 * H_2) * \
(10 ** (-3) * Td)) * (((Td + 273.15) / 285) ** 4) * \
(1 - (train_dataset3ModeloMatematico["Nubosidad-octa"
    ("\\%")]) / 8)) + \
(0.06 * (Td - (train_dataset3ModeloMatematico["Outdoor_"
    "Temperature(C)"]))) * \
(1 + 100 * (1 - exp20))

train_dataset3ModeloMatematico["calculo"] = A
train_dataset3ModeloMatematico["sum(dh/dt)calc(mm)("
    "Prediccion-MMatematico)"] = A
train_dataset3ModeloMatematico["sum(dh/dt)exp(mm)(Medicion)"
    ] = df2["medicion"]
train_dataset3ModeloMatematico['calculo'] =
    train_dataset3ModeloMatematico['calculo'].where(
        train_dataset3ModeloMatematico['calculo'] >= 0, 0)

```

```

train_dataset3ModeloMatematico[ 'sum( dh / dt ) calc (mm) (
    Prediccion -MMatematico ) ] =
train_dataset3ModeloMatematico[ 'sum( dh / dt ) calc (mm) (
    Prediccion -MMatematico ) ].where(
train_dataset3ModeloMatematico[ 'sum( dh / dt ) calc (mm) (
    Prediccion -MMatematico ) ] >= 0 , 0)

```

3) Cálculo de RNA, entrenamiento LM.

Listing 8.3: Ejemplo de código en Python

```

model6 = keras.Sequential()
model6.add(Dense(units=4, input_dim=4))
model6.add(Dense(units=5, activation='tanh'))
model6.add(Dense(units=1, activation='sigmoid'))
f = lambda x: 1*x[0] + x[1]
def build_model():
    model1=model6
    optimizer = LevenbergMarquardt(learning_rate=0.0001, initial_weights=[1
        for _ in range(model6.count_params())])
    model1.compile(optimizer=optimizer, loss='mse',
                   metrics=['mae', 'mse'])

    return model1
# Display training progress by printing a single dot for each completed epoch
class PrintDot(keras.callbacks.Callback):
    def on_epoch_end(self, epoch, logs):
        if epoch % 100 == 0: print('')
        print('.', end=' ')
EPOCHS = 5000

history = model1.fit(
    train_dataset3, train_labels3, shuffle=True, validation_data=(test_dataset3,
    epochs=EPOCHS, verbose=0, callbacks=[PrintDot()]))

```

# **Agradecimientos a la Comisión de Revisión.**

Quiero expresar mi más sincero agradecimiento a la Comisión de Revisión, cuya dedicación y experiencia fueron fundamentales para enriquecer significativamente este trabajo.

Integrantes de la Comisión:

1. Ana Aguilera, Universidad de Valparaíso.
2. Danilo Carvajal, Universidad la Serena.
3. Jean-Gabriel Minonzio, Universidad de Valparaíso.

Cada miembro de la Comisión aportó valiosos comentarios y sugerencias que contribuyeron de manera significativa a la mejora continua de este documento. La perspicacia y el compromiso con la excelencia académica de cada uno de ustedes han dejado una marca indeleble en este trabajo.

Quisiera expresar mi más sincero y profundo agradecimiento al Profesor Jean-Gabriel Minonzio por su invaluable papel como Profesor Guía en este proyecto. Sus orientaciones, asesoramientos y dedicación fueron fundamentales para el desarrollo de la investigación. Su amplio conocimiento, especialmente al explicar el modelo de D. Beysens, y su ayuda en la interpretación de los resultados, han sido un pilar fundamental en este proceso. Su compromiso y apoyo han marcado una diferencia significativa en el éxito de este trabajo de título.

Adicional quiero agradecer al Profesor Danilo Carbalal, por cada explicación detallada de los modelos matemáticos y estimativos a partir de fórmulas y datos medioambientales, proporcionó una base sólida para este trabajo. Además, su valiosa contribución del proyecto Fondecyt de Iniciación N.º 11140863 mencionado en el trabajo de tesis y otros documentos, han enriquecido considerablemente la investigación.

La Profesora Ana ha sido una guía invaluable en el entendimiento y consideración de las Redes Neuronales Artificiales. Gracias por su ayuda y orientación han sido fundamentales para comprender y aplicar de manera efectiva este enfoque en la investigación.

A cada uno de ustedes, agradezco sinceramente el tiempo y esfuerzo dedicados a revisar y evaluar cada aspecto de mi investigación. Sus contribuciones han elevado la calidad y el alcance de este trabajo de título.

# Bibliografía

- [1] A. Duran, “Estructura típica de una neurona,” <https://www.msdmanuals.com/es/hogar/multimedia/figure/estructura-t%C3%ADpica-de-una-neurona>, 2012.
- [2] N. commonswiki, “Cloud forest,” [https://es.wikipedia.org/wiki/Humedad\\_relativa#/media/Archivo:Cloud\\_forest\\_mount\\_kinabalu.jpg](https://es.wikipedia.org/wiki/Humedad_relativa#/media/Archivo:Cloud_forest_mount_kinabalu.jpg), 2005.
- [3] G. Portillo, “Qué es el rocío,” <https://www.meteorologiaenred.com/que-es-el-rocio.html>.
- [4] “Elementos del clima nubosidad,” [https://www.educapplus.org/climatic/04\\_elem\\_nubosidad.html](https://www.educapplus.org/climatic/04_elem_nubosidad.html), 2022.
- [5] L. Muñoz and S. Alfaro, “Centro de estudios avanzados de zonas áridas (ceaza),” [http://www.ceazamet.cl/index.php?pag=mod\\_mapa&p\\_cod=ceazamet](http://www.ceazamet.cl/index.php?pag=mod_mapa&p_cod=ceazamet), 2019.
- [6] J. A. Rodrigo, “Redes neuronales con python,” <https://cienciadedatos.net/documentos/py35-redes-neuronales-python>, 2021.
- [7] F. S. Q, “El cambio climático y los recursos hídricos de chile,” [https://opia.fia.cl/601/articles-91835\\_archivo\\_01.pdf/](https://opia.fia.cl/601/articles-91835_archivo_01.pdf/), 2016.
- [8] S. Alhattab, “Uno de cada cinco niños en el mundo carece de agua suficiente para satisfacer sus necesidades diarias,” <https://www.unicef.org/es/comunicados-prensa/uno-de-cada-cinco-ninos-en-el-mundo-carece-de-agua-suficiente-para-satisfacer/>, 2021.
- [9] E. Arias, G. Tolentino, and M. Jacobo, “Captación de agua atmosférica sin consumo de energía,” <https://contactos.itz.uam.mx/index.php/contactos/article/view/221>, 2022.
- [10] D. Beysens, “Estimating dew yield worldwide from a few meteo data,” *Atmospheric Research*, vol. 167, pp. 146–155, 2016.
- [11] I. Lekouch, K. Lekouch, M. Muselli, A. Mongruel, B. Kabbachi, and D. Beysens, “Rooftop dew, fog rain collection in southwest morocco and predictive dew modeling

- using neural networks,” *Journal of Hydrology*, vol. 448-449, pp. 60–72, 2012. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0022169412002673>
- [12] G. Shrivastava, S. Karmakar, M. K. Kowar, and P. Guhathakurta, “Application of artificial neural networks in weather forecasting: a comprehensive literature review,” *International Journal of Computer Applications*, vol. 51, no. 18, 2012.
- [13] D. Carvajal, J.-G. Minonzio, E. Casanga, J. Muñoz, A. Aracena, S. Montecinos, and D. Beysens, “Roof-integrated dew water harvesting in combarbalá, chile,” *Journal of Water Supply: Research and Technology—Aqua*, vol. 67, p. jws2018174, 05 2018.
- [14] D. Calvo, “Función de activación – redes neuronales,” <https://www.diegocalvo.es/funcion-de-activacion-redes-neuronales/>, 2018.
- [15] F. Ritter, M. Berkelhammer, and D. Beysens, “Dew frequency across the us from a network of in situ radiometers,” *Hydrology and Earth System Sciences*, vol. 23, no. 2, pp. 1179–1197, 2019. [Online]. Available: <https://hess.copernicus.org/articles/23/1179/2019/>
- [16] M. Kars, “Application note dew-point calculation,” [http://irtfweb.ifa.hawaii.edu/~tcs3/tcs3/Misc/Dewpoint\\_Calculation\\_Humidity\\_Sensor\\_E.pdf](http://irtfweb.ifa.hawaii.edu/~tcs3/tcs3/Misc/Dewpoint_Calculation_Humidity_Sensor_E.pdf), 2006.
- [17] “Temperatura del aire,” <https://content.meteoblue.com/es/research-education/educational-resources/meteoscool/el-tiempo/temperatura>, 2007.
- [18] J. Furest and R. Moreno, “Temperatura del aire atmosférico,” [http://www.inia.org.uy/disciplinas/agroclima/le/cc\\_tempe.htm](http://www.inia.org.uy/disciplinas/agroclima/le/cc_tempe.htm).
- [19] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, “Predicting solar generation from weather forecasts using machine learning,” pp. 528–533, 2011.
- [20] S. Corlay, P. Ivanov, J. Mabille, I. PresedoFloyd, M. Ragan-Kelley, Z. Sailer, R. Wagner, and C. Willing, “Jupyter,” <https://docs.jupyter.org/en>, 2014.
- [21] “Docker,” <https://docs.docker.com/>, 2013.
- [22] “Github,” <https://docs.github.com/es/get-started/quickstart/hello-world>, 2008.
- [23] J. Ellingwood, “Información sobre algoritmos de selección de bloques de servidores y ubicación de nginx,” <https://www.digitalocean.com/community/tutorials/understanding-nginx-server-and-location-block-selection-algorithms-es>, 2021.
- [24] M. Drake, “Cómo instalar mysql en ubuntu 20.04,” <https://www.digitalocean.com/community/tutorials/how-to-install-mysql-on-ubuntu-20-04-es>, 2020.