



**DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA
CONSTRUCCIÓN**

**CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE
TRABAJO DE TITULACIÓN, PREVIO A LA OBTENCIÓN DEL TÍTULO
DE INGENIERA GEÓGRAFA Y DEL MEDIO AMBIENTE**

**TEMA: MODELAMIENTO DE RADIACIÓN SOLAR A PARTIR DE
VARIABLES METEOROLÓGICAS MEDIANTE LA TRANSFORMADA
DE WAVELET Y REDES NEURONALES ARTIFICIALES
RECURRENTES**

AUTORA: CEPEDA VELASTEGUI, MARILYN VANESA

DIRECTOR: DR. TIERRA CRIOLLO, ALFONSO RODRIGO

SANGOLQUÍ

2019



DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA CONSTRUCCIÓN

CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE

CERTIFICACIÓN

Certifico que el trabajo de titulación, **“MODELAMIENTO DE RADIACIÓN SOLAR A PARTIR DE VARIABLES METEOROLÓGICAS MEDIANTE LA TRANSFORMADA DE WAVELET Y REDES NEURONALES ARTIFICIALES RECURRENTE”** fue realizado por la señorita Cepeda Velastegui, Marilyn Vanesa el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 12 de noviembre del 2019

A handwritten signature in black ink, appearing to read 'Alfonso Rodrigo Tierra Criollo', written over a light blue circular stamp.

Dr. Tierra Criollo, Alfonso Rodrigo

C.C. 0601803273



DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA CONSTRUCCIÓN

CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE

AUTORÍA DE RESPONSABILIDAD

Yo, Cepeda Velastegui, Marilyn Vanesa, declaro que el contenido, ideas y criterios del trabajo de titulación: **“MODELAMIENTO DE RADIACIÓN SOLAR A PARTIR DE VARIABLES METEOROLÓGICAS MEDIANTE LA TRANSFORMADA DE WAVELET Y REDES NEURONALES ARTIFICIALES RECURRENTE”** es de mi autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente el contenido de la investigación mencionada es veraz.

Sangolquí, 12 de noviembre del 2019

A handwritten signature in blue ink, which appears to read 'Marilyn Vanesa Cepeda Velastegui', is positioned above the printed name.

Cepeda Velastegui, Marilyn Vanesa

C.C. 0603943556



DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA CONSTRUCCIÓN
CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE

AUTORIZACIÓN

Yo, Cepeda Velastegui, Marilyn Vanesa autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: “**MODELAMIENTO DE RADIACIÓN SOLAR A PARTIR DE VARIABLES METEOROLÓGICAS MEDIANTE LA TRANSFORMADA DE WAVELET Y REDES NEURONALES ARTIFICIALES RECURRENTE**” en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 12 de noviembre del 2019

A handwritten signature in blue ink, which appears to read 'Marilyn Vanesa Cepeda Velastegui'. The signature is written in a cursive style.

Cepeda Velastegui, Marilyn Vanesa

C.C. 0603943556

DEDICATORIA

A mi ejemplo más grande de amor y generosidad, Carmen Demetría Rodríguez Barba, abuelita de mi vida y mi ángel amado, a Jocelyne Aliaga, amiga querida quien me enseñó a volar y que ahora guía todos mis pasos. Pronto nos encontraremos en la paz del valle.

A mi tesoro más grande que es mi familia, a mis madres Nancy y Carmen Velastegui, por ser quienes me han enseñado las lecciones más importantes de mi vida y el reflejo más grande valentía, a mi hermana Cristina Rodríguez, mi pequeña, la mejor compañera de vida, a mi tío Rodrigo Velastegui, querido compañero de historias y risas, a mi Papi Germán por siempre estar junto a mí con sus mejores consejos y sus palabras de aliento día a día, a mi Papi Lucio que me ha acompañado durante todos los años de mi vida con su cariño y apoyo incondicional.

A todos mis amigos, especialmente a David Benalcázar mi mejor amigo, hermano y compañero de alegrías y tristezas, a Valeria Chimbo, Selena Castro, Kenneth Pachacama y Christian Almendariz, que me han brindado una amistad sincera a lo largo de mi vida y se han convertido en mi segundo hogar. Al profesor Iván Jumbo por brindarme una de las herramientas más hermosas para sobrellevar la vida y hacerla más bella, la música.

A mí porque *“En las profundidades del invierno finalmente aprendí que en mi interior habitaba un verano invencible”* (Camus, 1954).

AGRADECIMIENTO

Agradecer se transforma inevitablemente en un acto injusto, es muy complejo el abarcar a todos quienes de alguna forma hicieron posible directa o indirectamente el presente trabajo. Sin embargo, asumiendo el riesgo deseo aprovechar estas líneas para agradecer de manera especial a mis guías en este proceso de investigación a los doctores Alfonso Tierra y Marco Luna, quienes desde un inicio brindaron su mejor contingente para lograr el objetivo, y esencialmente por su gran calidad de seres humanos. Al ingeniero Raúl Galeas y quienes conforman el Fondo de Protección del Agua (FONAG) por su afabilidad y contribución con la información para realizar la investigación. Al doctor Oswaldo Padilla por su apoyo a lo largo de mi carrera universitaria con aportes académicos y humanos.

ÍNDICE DE CONTENIDOS

CERTIFICACIÓN	i
AUTORÍA DE RESPONSABILIDAD.....	ii
AUTORIZACIÓN.....	iii
DEDICATORIA.....	iv
AGRADECIMIENTO.....	v
ÍNDICE DE CONTENIDOS	vi
ÍNDICE DE TABLAS	x
ÍNDICE DE FIGURAS.....	xi
RESUMEN.....	xiv
ABSTRACT	xv
CAPÍTULO I.....	1
INTRODUCCIÓN.....	1
1.1 Planteamiento del problema	1
1.2 Antecedentes.....	2
1.3 Justificación e importancia	3
1.4 Objetivos.....	3
1.4.1 Objetivo General	3
1.4.2 Objetivos específicos.....	4
1.4.3 Metas	4
CAPÍTULO II	5

MARCO TEÓRICO	5
2.1 Meteorología.....	5
2.1.1 Viento.....	5
2.1.2 Humedad del aire	5
2.1.3 Precipitación	5
2.1.4 Temperatura del aire	5
2.1.5 Humedad del suelo.....	6
2.1.6 Radiación solar.....	6
2.1.6.1 Potencial del aprovechamiento de la radiación solar.....	6
2.1.6.2 Efecto de la radiación solar sobre la salud	6
2.2 Series Temporales.....	7
2.2.1 Componentes de una series temporal.....	8
2.2.2 Clasificación descriptiva de las series temporales	8
2.3 Relleno de datos en series temporales	9
2.3.1 Método Holt – Winters	9
2.4 Transformada de Wavelet.....	10
2.4.1 Daubechies Wavelets	12
2.5 Coeficiente de correlación	13
2.6 Machine Learning.....	13
2.6.1 Deep Learning.....	14
2.6.1.1 Redes neuronales artificiales.....	14
2.6.1.1.1 Perceptrón Multicapa	15
2.6.1.1.2 Feed-back propagation	16
2.6.1.1.3 Levenberg-Marquardt.....	17

2.6.1.1.4 Función de Transferencia	17
2.6.1.2 Redes Neuronales Artificiales Recurrentes	18
2.7 Coeficiente de determinación	19
CAPÍTULO III	20
MARCO METODOLÓGICO	20
3.1 Área de estudio	20
3.2 Recopilación de datos	21
3.3 Generación de series temporales de las variables meteorológicas	22
3.3.1 Relleno de datos en las series temporales y obtención de coeficientes de correlación.	23
3.4 Aplicación de la transformada de wavelet a las series temporales de radiación	24
3.5 Redes neuronales artificiales recurrentes	25
3.5.1 Creación de redes neuronales	25
3.5.2 Entrenamiento de las redes neuronales.....	28
3.5.3 Predicción mediante redes neuronales.....	29
3.5.4 Validación.....	30
CAPÍTULO IV	31
RESULTADOS Y DISCUSIÓN	31
4.1 Resultado de la generación de las series temporales de las variables meteorológicas ...	31
4.1.1 Relleno de series temporales	31
4.1.2 Coeficientes de correlación de las variables meteorológica.....	35
4.2 Aplicación de la transformada de wavelet a series temporales de radiación.....	40
4.3 Redes neuronales artificiales	43
4.3.1 Arquitectura de las redes neuronales artificiales.....	43
4.3.2 Resultados de entrenamiento de las redes neuronales.....	44

4.3.3	Resultados de la predicción.....	49
4.3.4	Validación de la predicción.....	62
CAPITULO V		65
CONCLUSIONES Y RECOMENDACIONES		65
5.1	Conclusiones.....	65
5.2	Recomendaciones	65
BIBLIOGRAFÍA.....		67

ÍNDICE DE TABLAS

Tabla 1 <i>Intensidad y dirección del coeficiente r</i>	13
Tabla 2 <i>Estaciones meteorológicas y variables para generación de series de tiempo</i>	22
Tabla 3 <i>Paquetes empleados en el software R</i>	23
Tabla 4 <i>Criterios seleccionados para creación de las redes neuronales artificiales</i>	27
Tabla 5 <i>Coefficientes de correlación entre variables meteorológicas de la estación Miranda</i>	35
Tabla 6 <i>Coefficientes de correlación entre variables meteorológicas de la estación</i>	36
Tabla 7 <i>Coefficientes de correlación entre variables meteorológicas de la estación Chumillos</i> .	37
Tabla 8 <i>Coefficientes de correlación entre variables meteorológicas de la estación Puntas</i>	38
Tabla 9 <i>Coefficientes de correlación entre variables meteorológicas de la estación Itulcachi</i> ...	39
Tabla 10 <i>Resumen de los datos predichos mediante redes neuronales</i>	50
Tabla 11 <i>Coefficientes de determinación de las predicciones</i>	62
Tabla 12 <i>Valores máximos, mínimos, promedio y desviación estándar de las predicciones, valores reales y errores absolutos de radiación solar de las estaciones en las horas de mayor incidencia</i>	63

ÍNDICE DE FIGURAS

Figura 1. Índice UV.....	7
Figura 2. Algoritmo wavelet	11
Figura 3. Descomposición de la señal mediante wavelet.....	12
Figura 4. Árbol de descomposición wavelet	12
Figura 5. Arquitectura de una red neuronal típica	16
Figura 6. Función de transferencia Tangente Sigmoidal Hiperbólica.....	18
Figura 7. Función de transferencia lineal.....	18
Figura 8. Red UV INAMHI	20
Figura 9. Estaciones con información de radiación solar FONAG.....	21
Figura 10. Interfaz de nntool para Redes Neuronales en MATLAB R2015a	26
Figura 11. Interfaz para creación de variables en MATLAB R2015a	26
Figura 12. Interfaz de la ventana de entrenamiento	28
Figura 13. Interfaz de la ventana de resultados del entrenamiento	29
Figura 14. Interfaz de la ventana de Simulación.....	30
Figura 15. Serie temporal de radiación solar (datos de entrenamiento y validación)	30
Figura 16. Gráfico de la serie temporal de la variable dirección del viento– Miranda.....	32
Figura 17. Gráfico de la serie temporal de la variable humedad del aire– Miranda	33
Figura 18. Gráfico de la serie temporal de la variable humedad del suelo – Miranda.....	33
Figura 19. Gráfico de la serie temporal de la variable precipitación – Miranda.....	33
Figura 20. Gráfico de la serie temporal de la variable radiación - Miranda	34
Figura 21. Gráfico de la serie temporal de la variable temperatura – Miranda	34
Figura 22. Gráfico de la serie temporal de la variable velocidad del viento – Miranda	34
Figura 23. Análisis de radiación solar mediante la transformada de wavelet – Miranda	40
Figura 24. Análisis de radiación solar mediante la transformada de wavelet – El Carmen.....	41
Figura 25. Análisis de radiación solar mediante la transformada de wavelet – Chumillos	41
Figura 26. Análisis de radiación solar mediante la transformada de wavelet – Puntas	42
Figura 27. Análisis de radiación solar mediante la transformada de wavelet – Itulcachi	42

Figura 28. Arquitectura de la red neuronal artificial.....	44
Figura 29. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Miranda a) 10H00, b) 12H00, c) 14H00 y d) 16H00.....	45
Figura 30. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación El Carmen a) 10H00, b)12H00, c)14H00 y d) 16H00.....	46
Figura 31. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Chumillos a) 10H00, b)12H00, c)14H00 y d) 16H00.....	47
Figura 32. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Puntas a) 10H00, b)12H00, c)14H00 y d) 16H00.....	48
Figura 33. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Itulcachi a) 10H00, b)12H00, c)14H00 y d) 16H00.....	49
Figura 34. Modelo radiación solar 10H00 - Miranda.....	50
Figura 35. Modelo de radiación solar 12H00 - Miranda.....	51
Figura 36. Modelo de radiación solar 14H00 - Miranda.....	51
Figura 37. Modelo de radiación solar 16H00 - Miranda.....	52
Figura 38. Modelo de radiación solar 10H00 - El Carmen.....	53
Figura 39. Modelo de radiación solar 12H00 - El Carmen.....	53
Figura 40. Modelo de radiación solar 14H00 - El Carmen.....	54
Figura 41. Modelo de radiación solar 16H00 - El Carmen.....	54
Figura 42. Modelo de radiación solar 10H00 - Chumillos.....	55
Figura 43. Modelo de radiación solar 12H00 - Chumillos.....	56
Figura 44. Modelo de radiación solar 14H00 - Chumillos.....	56
Figura 45. Modelo de radiación solar 16H00 - Chumillos.....	57
Figura 46. Modelo de radiación solar 10H00- Puntas.....	57
Figura 47. Modelo de radiación solar 12H00- Puntas.....	58
Figura 48. Modelo de radiación solar 14H00- Puntas.....	58
Figura 49. Modelo de radiación solar 16H00- Puntas.....	59
Figura 50. Modelo de radiación solar 10H00- Itulcachi.....	60
Figura 51. Modelo de radiación solar 12H00- Itulcachi.....	60
Figura 52. Modelo de radiación solar 14H00- Itulcachi.....	61

Figura 53. Modelo de radiación solar 16H00- Itulcachi 61

RESUMEN

El estudio de la radiación solar ha sido un tema de gran importancia e interés a lo largo de los años; sin embargo, su comportamiento al igual que de otras variables meteorológicas han implicado cierto nivel de dificultad por la gran fluctuación de los datos. Existen diversas técnicas que proponen el modelamiento y predicción de radiación solar y de otras variables meteorológicas, en muchos de los casos se supone naturaleza lineal de la información o se emplea modelos ajenos a la zona local, lo cual induce a que los resultados no brinden veracidad y permanezcan alejados de la realidad. El proyecto de investigación analizó las series de tiempo de las variables: dirección del viento, humedad del aire, humedad del suelo, precipitación, radiación solar, temperatura del aire y velocidad del viento, de las estaciones Miranda, El Carmen, Chumillos, Puntas e Itulcachi, las cuales pertenecen al Fondo de Protección del Agua (FONAG). Se creó las series de tiempo de las diferentes variables y estaciones en función de las horas de mayor incidencia de radiación (10H00, 12H00, 14H00 y 16H00 - hora local) por un periodo de 4 años y seis meses con datos diarios. Se realizó el relleno de datos de las series de tiempo mediante la técnica Holt-Winters en el software R Studio. Se aplicó la transformada de wavelet familia Daubechies 4 nivel de discretización 3 con el fin de eliminar el ruido presente en las series temporales de la radiación solar en el software MatLab. Se entrenó las redes neuronales artificiales recurrentes empleando como variables de entrada las variables meteorológicas y como variable de salida específicamente la radiación solar. Se analizó el correcto entrenamiento de las redes y se procedió a la predicción de la radiación solar. Finalmente se validó la predicción mediante el coeficiente de determinación como criterio estadístico.

PALABRAS CLAVE:

- **SERIES TEMPORALES**
- **METEOROLOGÍA**
- **HOLT-WINTERS**
- **PREDICCIÓN**

ABSTRACT

The study of solar radiation has been a subject of great importance and interest over the years; however, its behavior, like other meteorological variables, has implied a certain level of difficulty due to the large fluctuation of the data. There are several techniques that propose the modeling and forecast of solar radiation and other meteorological variables, in many cases the linear nature of the information is assumed or models outside the local area are used, which leads to the results without truthfulness. The research project analyzed the time series of the variables: wind direction, air humidity, soil humidity, precipitation, solar radiation, air temperature and wind speed, from Miranda, El Carmen, Chumillos, Puntas e Itulcachi climatological stations, which belong to Fondo de Protección del Agua (FONAG). The time series of the different variables and stations was created based on the hours with the highest incidence of radiation (10H00, 12H00, 14H00 and 16H00 - local time) for a period of 4 years and six months with daily data. Time series data was filled using the Holt-Winters technique in the R Studio software. The transformation of wavelet family Daubechies 4 level of discretization 3 was applied in order to eliminate the noise present in the solar radiation time series in MatLab software. Artificial neural networks were trained using meteorological variables as input variables and solar radiation as a specific output variable. The correct training of the networks was analyzed and the solar radiation was predicted. Finally, the prediction was validated using the coefficient of determination as a statistical criterion.

KEYWORDS:

- **TIME SERIES**
- **METEOROLOGY**
- **HOLT-WINTERS**
- **FORECAST**

CAPÍTULO I

INTRODUCCIÓN

1.1 Planteamiento del problema

A nivel mundial el aprovechamiento del potencial de la energía solar es un tema de gran importancia, debido a que corresponde a una fuente significativa de energía alternativa ante las energías convencionales, que en muchos casos generan impactos ambientales negativos (Doncel, 2018).

Ecuador posee una ubicación adecuada para obtener y aprovechar la energía solar, y a la vez está propenso a índices muy altos de radiación solar, incluso se encuentra entre las zonas que han registrado mayores valores a nivel mundial, informó el Diario la Hora (2018) según el Instituto Nacional de Meteorología e Hidrometeorología (INAMHI). Existe mayor radiación al encontrarse más cerca del ecuador y a su vez a mayor altitud la atmosfera es más delgada, por lo que según la Organización Mundial de la Salud (2003) con cada 1000 metros de incremento de altitud, la intensidad de la radiación aumenta 10 a 12%.

Sin embargo, no existen numerosos aportes que permitan obtener una visión clara de la dinámica de la radiación solar. Los insumos que han sido utilizados no corresponden a datos actuales o no cuentan con precisión adecuada para el país, debido a que se basan en la realidad de otras zonas.

Es por ello que existe la necesidad de generar análisis que permitan conocer el comportamiento de la radiación solar para lograr el beneficio de la misma. Además es indispensable contar con información para la contingencia y respuesta cuando la radiación solar pueda ser perjudicial tanto para el ser humano, como para posibles escenarios de sequía entre otras problemáticas.

Por este motivo se plantea la predicción de datos de radiación solar mediante el análisis de las series de tiempo con la transformada de wavelet y la predicción a través de redes neuronales, lo cual busca constituir un apoyo para determinar medidas en el aprovechamiento y la prevención de la energía solar.

1.2 Antecedentes

El estudio de la energía solar es de gran importancia para diversas áreas, tanto para el campo de estudio de las energías renovables, la agricultura, cambio climático, salud entre otras (Arla S. , et al., 2017). A escala mundial se han desarrollado diversos sistemas para cuantificar la energía solar, entre ellos existen los métodos terrestres los cuales utilizan herramientas como pirómetros, heliógrafos y estaciones meteorológicas de medición, mediante los cuales se pueden recopilar datos de radiación solar, que posteriormente pueden ser analizados (Doncel, 2018).

En el Ecuador existe un Atlas Solar con fines de generación eléctrica, para el cual se utilizaron métodos satelitales como es el caso de imágenes GOES, con una resolución espacial de 1 Km x 1 Km y se incorporaron datos de estaciones en tierra. La información del documento corresponde al período entre el 1 de Enero de 1985 y el 31 de Diciembre de 1991, y fue publicado en Julio del 2006 (CONELEC, 2008).

Además la red UV del INAMHI posee 9 estaciones que miden radiación ultravioleta, las cuales se encuentran ubicadas en diferentes puntos del Ecuador continental. Los pronósticos para el Ecuador son elaborados por la Dirección de Pronóstico y Alertas Hidrometeorológicas del INAMHI, mediante la información del satélite GOES-16 y modelo numérico del ECMWF (European Centre for Medium-Range Weather Forecasts) y WRF (Weather Research and Forecasting) (INAMHI, 2018).

La predicción de la radiación solar se ha convertido en un tema indispensable para el aprovechamiento y contingencia de la misma, es por ello que existen diversas técnicas que permiten obtener un panorama futuro en función de los datos que se han ido almacenando históricamente.

La transformada de wavelet y las redes neuronales son herramientas muy útiles para este tipo de análisis, debido a que logran obtener pronósticos a través de valores del pasado para generar valores futuros (Pratap Yadav & Laxmidhar, 2014).

1.3 Justificación e importancia

El beneficio que otorga el análisis del comportamiento de la radiación solar, constituye un aspecto indispensable para mensurar la disposición de sistemas energéticos solares, tomando en consideración la parte ambiental. Además es fundamental para la investigación de efectos de cambio climático y aspectos de salud de la colectividad referente a la interacción con la radiación solar y sus consecuencias (Arla S. , et al., 2017).

En el país no se cuenta con información actualizada y la predicción se encuentra basada en modelos que están adaptados para zonas ajenas a Ecuador, por lo que los resultados no se encuentran acoplados a nuestra realidad en varios de los casos (Arla S. , et al., 2017).

Las redes neuronales utilizan valores pasados para generar una retroalimentación en ella, es decir esta red tiene un tipo de elemento de memoria en base a los entrenamientos asignados, por lo cual la convierten en una herramienta muy útil para este tipo de estudios de series de tiempo con la variable de radiación solar y para obtener datos de salida futuros (Pratap Yadav & Laxmidhar, 2014).

El presente proyecto se alinea al objetivo 7 de los Objetivo de Desarrollo Sostenible, en el cual se manifiestan temas de energía asequible y no contaminante. El presente estudio brinda una visión de la dinámica de la energía solar para que los tomadores de decisiones puedan adoptar las medidas en busca del aprovechamiento de este tipo de energía no contaminante (UNDP, 2019).

1.4 Objetivos

1.4.1 Objetivo General

Modelar series temporales de radiación solar a partir de variables meteorológicas mediante la transformada de wavelet y redes neuronales artificiales, en la zona sur-este del Distrito Metropolitano de Quito, para obtener buenas predicciones de radiación solar en un tiempo no mayor a seis meses.

1.4.2 Objetivos específicos

- Analizar las variables meteorológicas y la radiación solar para determinar la correlación entre ellas a través de la obtención coeficiente de correlación.
- Aplicar la transformada de wavelet en las series temporales de radiación para eliminación de ruido mediante el análisis con la familia wavelet db4 nivel 3.
- Predecir los valores de radiación solar mediante el entrenamiento de redes neuronales artificiales para evaluar la generalización de las mismas.

1.4.3 Metas

- Una base de datos de la información de las variables meteorológicas de las estaciones de FONAG ubicadas en la zona sur-este del DMQ.
- Un análisis de la correlación de las diferentes variables meteorológicas con la radiación solar.
- Cinco series temporales de radiación solar sin ruido.
- Redes neuronales entrenadas en función de las series temporales (10H00, 12H00, 14H00, 16H00) de los parámetros meteorológicos y la radiación solar.
- Valores predichos de radiación solar.

CAPÍTULO II

MARCO TEÓRICO

2.1 Meteorología

Es la ciencia que abarca el estudio de la atmósfera y los fenómenos que se producen en la misma. Se basa en magnitudes o variables meteorológicas que corresponden a presión, temperatura, humedad, presión atmosférica entre otras (Rodríguez, Capa, & Portela, 2004).

2.1.1 Viento

Es el aire en movimiento de una zona a otra. Se origina por la diferencia de presión o temperatura entre dos puntos (Rodríguez, Capa, & Portela, 2004). La velocidad del viento es medida en metros por segundo mediante una herramienta llamada anemómetro de cazoletas (Rodríguez, Capa, & Portela, 2004). La dirección del viento se mide a través de veletas las cuales indican la procedencia geográfica del viento (Rodríguez, Capa, & Portela, 2004).

2.1.2 Humedad del aire

Es resultado del vapor de agua presente en la atmosfera debido a la evaporación de lagos, ríos, océanos, sublimación de hielo entre otros (Jaramillo, 2005).

2.1.3 Precipitación

Es la cantidad de agua que proviene de la caída de lluvia, llovizna (estado líquido) o en forma de granizo o nieve (estado sólido), hacia la superficie terrestre. Se mide mediante un pluviómetro (INAMHI, 2014).

2.1.4 Temperatura del aire

Es una magnitud que está relacionada con la rapidez del movimiento de las partículas que forman la materia. Mientras exista mayor agitación por ende habrá mayor temperatura (Rodríguez,

Capa, & Portela, 2004). La temperatura del aire es aquella que no se encuentra expuesta a la radiación del sol directamente y es señalada por un termómetro al aire (INAMHI, 2014).

2.1.5 Humedad del suelo

La humedad del suelo corresponde a la masa de agua que se encuentra por unidad de masa de suelo (Flores & Alcalá, 2010). La humedad del suelo no corresponde netamente a una variable meteorológica, sino una propiedad física del suelo.

2.1.6 Radiación solar

La radiación solar se define como la medida incidente de la cantidad total de radiación solar transmitida a la superficie de la atmósfera de la Tierra en una unidad de tiempo determinada. Se mide en vatios por metro cuadrado (W/m^2) (Pratap Yadav & Laxmidhar, 2014).

La proporción de la radiación solar en las regiones del espectro es: Ultravioleta 7%, luz visible 43%, Infrarrojo 49 % y 1% restante (Agencia Estatal de Meteorología, 2014).

2.1.6.1 Potencial del aprovechamiento de la radiación solar

El Sol emite intermitentemente aproximadamente 62.000 kilowatts de potencia por metro cuadrado de su superficie. En un periodo de dos días la tierra obtiene una cantidad equivalente de energía a todas las reservas conocidas de gas, carbón y petróleo (Arancibia & Best, 2010).

Sin duda existe un antagonismo entre las fuentes de energía de los combustibles fósiles y la energía solar, debido a que en el primer caso existen fuertes impactos ambientales mientras que en el segundo respectivamente se recibe de forma natural y no implica la destrucción o modificación del entorno (Arancibia & Best, 2010).

2.1.6.2 Efecto de la radiación solar sobre la salud

La radiación Ultravioleta representa al 7% del total de radiación del sol, sin embargo provoca efectos sobre el ambiente y los seres que habitan en él, convirtiendo esa porción relativamente pequeña en un factor importante (Agencia Estatal de Meteorología, 2014).

Los valores altos de radiación tienen una influencia importante sobre la salud de las personas, la exposición prolongada genera efectos crónicos en la piel, ojos y el sistema inmune, incluso produciendo cáncer y cataratas (OMS, 2003).

Existe un índice de radiación ultravioleta que permite comunicar a la ciudadanía las diferentes condiciones de la misma de forma sencilla. En la *Figura 1* se indica el rango, nivel de riesgo y recomendación en cada caso respectivamente.



Figura 1. Índice UV

Fuente: (Secretaría del Ambiente, 2018)

2.2 Series Temporales

Una serie temporal es una secuencia de N observaciones ordenadas cronológicamente sobre una característica o sobre varias características de la unidad observable en diferentes momentos. Se puede representar la series univariantes de forma matemática como indica la ecuación (1). Donde X_i es la observación en el instante i de la serie y N corresponde al número de observaciones (Mauricio, 2007).

$$(X_i)_{i=1}^N; X_1, X_2, \dots, X_N \quad (1)$$

2.2.1 Componentes de una series temporal

Tendencia: Es el cambio a largo plazo en relación a la media de la serie (Villavicencio, 2011).

Estacionalidad: representa la periodicidad que se puede dar en variaciones anuales, mensuales, semestrales entre otras (Villavicencio, 2011).

Aleatoriedad: Resulta de elementos aleatorios, es decir no son parte de los patrones de comportamiento (Villavicencio, 2011).

La descomposición aditiva de series temporales se expresa como la ecuación (2), donde X_t es el valor observado, T_t representa la tendencia, E_t representa a la estacionalidad e I_t es la componente irregular.

La componente aleatoria o irregular corresponde a fluctuaciones aleatorias de innovaciones no explicadas por los otros componentes (Justel & Cayuela, 2015).

$$X_t = T_t + E_t + I_t \quad (2)$$

2.2.2 Clasificación descriptiva de las series temporales

Estacionarias: La serie corresponde a estacionaria cuando la media y la variabilidad permanecen constantes a través del tiempo. Es decir, los valores de la serie oscilan sobre una media constante y la variación con respecto a dicha media es constante en el tiempo (Marin, 2017).

No estacionarias: En este caso la serie no fluctúa sobre un valor constante, la tendencia cambia a lo largo del tiempo (Marin, 2017).

2.3 Relleno de datos en series temporales

Existen diversos métodos para el relleno de datos de series temporales de variables meteorológicas, entre las más usadas es mediante regresiones lineales, suponiendo relación lineal entre las variables, también se aplica estaciones auxiliares mediante una ecuación entre dos variables (Carrera et al., 2016). Esta metodología es adecuada para la estimación de datos mensuales y anuales (Pizarro et al., 2009).

En el caso de las variables meteorológicas de la región andina existe escasez de información por lo que se ha propuesto un umbral de 25% de datos faltantes en las series temporales (Campozano et al., 2015).

Varios de los procedimientos asumen linealidad entre las variables (Justel & Cayuela, 2015), por lo que en el caso de la información meteorológica no siempre cumple con este parámetro, motivo por el cual se han implementado otros métodos que permitan solventar estas condiciones.

2.3.1 Método Holt – Winters

Los métodos de suavizado exponencial han permitido resolver varias limitaciones de modelos deterministas. Es una forma simple de extrapolar de manera en que los últimos puntos de datos de la serie tienen más peso en el pronóstico que los valores anteriores (Justel & Cayuela, 2015).

El método Holt-Winters es una herramienta para pronóstico que posee fácil adaptabilidad a medida que la información real incrementa. Existen los métodos Holt-Winters doble exponente suavizante (doble suavizado exponencial) y Holt-Winters triple exponente suavizante (triple suavizado exponencial). Se tienen 3 constantes que son la media, tendencia y variación estacional respectivamente, en el caso del doble suavizado exponencial no se considera la tercera variable (Izar, 2019).

El suavizado exponencial debe actualizar dos componentes en cada período: nivel y tendencia. El nivel es una estimación suavizada del valor de los datos al final de cada período. La tendencia es una estimación uniforme del crecimiento promedio al final de cada período (Prajakta, 2004).

El método consta de tres fases e interactúa con tres conjuntos de datos diferentes. El primer grupo de datos es para iniciar el modelamiento, en el que se determina los indicadores de nivel, tendencia y estacionalidad. A través del segundo conjunto de datos se obtiene la calibración de los índices de suavización Alfa (α), Beta (β) y Gamma (γ). El tercer grupo de valores permiten pronosticar y evaluar el modelo.

El método Holt-Winters doble exponente suavizante o doble suavizado exponencial se realiza a partir de la ecuación (3) y (4), donde m_t corresponde al nivel promedio, x_t es el valor observado en tiempo t y b_t es la tendencia.

$$m_t = \alpha x_t + (1-\alpha)(m_{t-1} + b_{t-1}) \quad (3)$$

$$b_t = \beta(m_t - m_{t-1}) + (1 - \beta)b_{t-1} \quad (4)$$

$$0 \leq \alpha \leq 1$$

$$0 \leq \beta \leq 1$$

$\alpha \approx 0$ y $\beta \approx 0$ Cuando la evolución es casi constante.

$\beta \approx 0$ Cuando la pendiente es casi constante.

Por lo general, los parámetros se eligen para minimizar el error cuadrático medio (MSE) y el algoritmo comienza con $m_1 = x_1$, $b_1 = x_2 - x_1$ y finalmente las predicciones (\hat{x}_{t+h}) están dados por la ecuación (5), donde h corresponde al número de periodos que serán proyectados (Justel & Cayuela, 2015).

$$\hat{x}_{t+h} = m_t + b_t h \quad (5)$$

2.4 Transformada de Wavelet

Wavelet es una señal oscilatoria, de corta duración y energía finita concentrada en un intervalo de tiempo alrededor de un punto (Nieto & Orozco, 2008; Renza, 2010; Palacios & Cepeda, 2018). Las wavelets son funciones básicas de la Transformada Wavelet madre, que representa una señal en versiones traducidas y dilatadas de una onda finita, no solo es local en tiempo, sino también en

frecuencia (Santamaría, Cortés, & Román, 2012). La transformada de wavelet es expresada de forma matemática como se indica en la ecuación (6).

$$WT(f(x)) = f(x) * \Psi(x) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \Psi\left(\frac{x-t}{a}\right) dt \quad (6)$$

Donde "a" es el factor de escala (dilatación), "t" es el tiempo y "x" es la posición. La función ψ se llama "wavelet madre"; primero, wavelet porque es de naturaleza oscilante y de duración finita (soporte compacto) y se llama madre por servir como base para la generación de las funciones de ventana restantes (Butler & Bagriyanik, 2003).

Existen varias familias wavelet, entre las más conocidas se encuentran: Haar, Daubechies, Coiflets, Symlets, Biorthogonal, Meyer, Mexican hat, Shannon y Morlet entre otras (Yamasaki, 2008; Butler & Bagriyanik, 2003).

El primer paso para la descomposición de una señal "s" de longitud N es generar dos conjuntos de coeficientes: coeficientes de aproximación cA_1 y coeficientes de detalle cD_1 . Para ello se realiza la convolución a "s" con el filtro de paso bajo LoD (low) y el filtro de paso alto HiD (high), seguido de la decimación diádica (disminución de resolución) y se obtiene como resultado los coeficientes de aproximación y detalle respectivamente como se indica en la *Figura 2*.

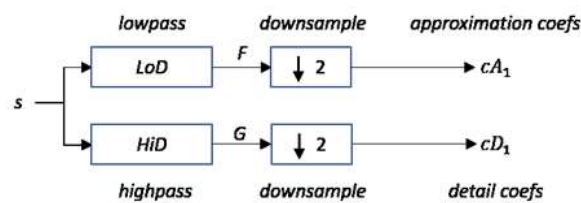


Figura 2. Algoritmo wavelet

Fuente: MathWorks, 2019

La longitud de cada filtro es igual a $2n$ (n corresponde al nivel de descomposición). Si $N =$ longitud de "s", las señales F y G son de longitud $N + 2n - 1$ y los coeficientes cA_1 y cD_1 son de longitud $\left(\frac{N-1}{2}\right) + n$.

El siguiente paso divide los coeficientes de aproximación cA_1 en dos partes usando el mismo esquema, reemplazando “s” por cA_1 y produciendo cA_2 y cD_2 , y así sucesivamente como se muestra en la *Figura 3*.

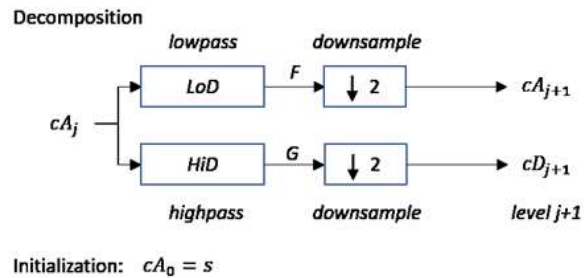


Figura 3. Descomposición de la señal mediante wavelet
Fuente: MathWorks, 2019

La descomposición wavelet de la señal “s” analizada en el nivel j tiene la siguiente estructura: $[cA_j, cD_j, \dots, cD_1]$. Esta estructura contiene, para $j = 3$, los nodos terminales que se indica en la *Figura 4* (MathWorks, 2019).

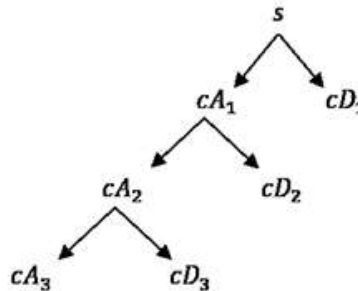


Figura 4. Árbol de descomposición wavelet
Fuente: MathWorks, 2019

2.4.1 Daubechies Wavelets

Una base de wavelet está aproximadamente localizada tanto en el dominio de frecuencia como en el espacial. En la fórmula Daubechies, estas propiedades se logran a través del proceso de recursión. Las funciones de escala y las wavelets se definen de forma recursiva mediante el uso de

ecuaciones de dilatación. La ecuación de dilatación básica es una ecuación de dos escalas o diferencia diádica (7) (Rowe & Abbott, 1995).

$$\varphi(x) = \sqrt{2} \sum_{k=0}^{N-1} c_k \theta(2x - k) \quad (7)$$

Donde: $\theta(x)$ es la función de escala, N representa el orden del sistema.

2.5 Coeficiente de correlación

El coeficiente de Pearson o de correlación r , corresponde a un indicador de correlación perfecta, en un rango de -1 a 1. En la tabla 1 se presenta la intensidad y la pendiente del coeficiente r (Lind, et al., 2004).

Tabla 1

Intensidad y dirección del coeficiente r

Coeficiente de correlación r	Intensidad de relación lineal
$r = -1$	Correlación negativa perfecta
$-1 < r < -0,5$	Correlación negativa intensa
$r = -0,5$	Correlación negativa moderada
$-0,5 < r < 0$	Correlación negativa débil
$r = 0$	Sin correlación
$0 < r < 0,5$	Correlación positiva débil
$r = 0,5$	Correlación positiva moderada
$0,5 < r < 1$	Correlación positiva intensa
$r = 1$	Correlación positiva perfecta

Fuente: (Lind, et al., 2004)

2.6 Machine Learning

Es una técnica de análisis de datos que enseña a las computadoras a hacer lo que es natural para los humanos y los animales: aprender de la experiencia. Los algoritmos de aprendizaje automático utilizan métodos computacionales para "aprender" la información directamente de los datos sin

depender de una ecuación predeterminada como modelo. Los algoritmos mejoran su rendimiento de forma adaptativa a medida que aumenta el número de muestras disponibles para el aprendizaje. El aprendizaje profundo es una forma especializada de aprendizaje automático (MathWorks, 2019).

Los algoritmos de aprendizaje automático encuentran patrones naturales en los datos que generan información y lo ayudan a tomar mejores decisiones y predicciones. Se usan todos los días para tomar decisiones críticas en el diagnóstico médico, el comercio de acciones, el pronóstico de carga de energía y más.

Por ejemplo, los sitios de medios confían en el aprendizaje automático para examinar millones de opciones para darle recomendaciones de canciones o películas. Los minoristas lo utilizan para obtener información sobre el comportamiento de compra de sus clientes (MathWorks, 2019).

2.6.1 Deep Learning

Es una técnica de “Machine Learning” o aprendizaje automático que indica a los computadores a aprender mediante ejemplos. Pueden tener gran precisión y se entrenan mediante un gran conjunto de datos y redes neuronales con múltiples capas (MathWorks, 2019).

El Deep Learning alcanza grandes niveles de precisión, sin embargo requiere grandes cantidades de datos etiquetados y la potencia de cálculo importante (MathWorks, 2019).

La mayoría de los métodos de aprendizaje utilizan redes neuronales que se dominan profundas. La expresión profunda, se refiere a las capas ocultas presentes en la red neuronal (MathWorks, 2019).

2.6.1.1 Redes neuronales artificiales

Son modelos de computación que presentan una estructura similar a la de una red de neuronas interconectadas en el cerebro. Una red neuronal aprende a través de datos, por lo que se realiza un entrenamiento para el reconocimiento de patrones, catalogación de datos y prediga o pronostique sucesos futuros (MathWorks, 2019).

Las entradas son sometidas a descomposición con el objetivo de lograr diferentes capas de abstracción. El comportamiento es definido por la manera en que se encuentran conectados los elementos individuales y por la ponderación de las conexiones. A lo largo del entrenamiento los pesos se ajustan de forma automática y la regla de aprendizaje específica, hasta que el objetivo se ha logrado correctamente (MathWorks, 2019).

Existen varias técnicas de machine learning para el uso de redes neuronales entre ellas se encuentran el aprendizaje supervisado y no supervisado, regresiones, clasificación, clustering entre otros (MathWorks, 2019).

El aprendizaje supervisado en redes neuronales consiste en producir las salidas que se desean con respecto a las entradas. Este tipo de redes es adecuado para modelar y controlar sistemas dinámicos, tratar datos con ruido y pronosticar datos del futuro. Este tipo de redes son: base radial, dinámicas, de cuantificación vectorial de aprendizaje y feedforward (MathWorks, 2019).

Por otro lado el aprendizaje no supervisado consiste en que en el entrenamiento se ajuste automáticamente con las entradas nuevas. Se utilizan para lograr deducir datos mediante conjuntos que constan de datos sin respuestas. Se emplean para encontrar categorías y las relaciones que existen entre ellas y también para deducir las distribuciones naturales (MathWorks, 2019).

2.6.1.1.1 Perceptrón Multicapa

El perceptrón multicapa es la arquitectura de una red típica, la cual consta de un número de neuronas en la capa de entrada, que está ligada de la cantidad de componentes del vector de entrada, del número de capas ocultas y del número de neuronas en ellas y de las neuronas en la capa de la salida, que a su vez tiene relación con el número de componentes del vector de salida (Yuleidys et al., 2013).

El funcionamiento de una red neuronal combina varias capas de procesos que operan en paralelo y emplea componentes simples. Está formada por una capa de entrada, capas ocultas y una de salida. Los nodos son los encargados de interconectar las capas y cada una de ellas usa la salida de la anterior como entrada (*Figura 5*) (MathWorks, 2019).

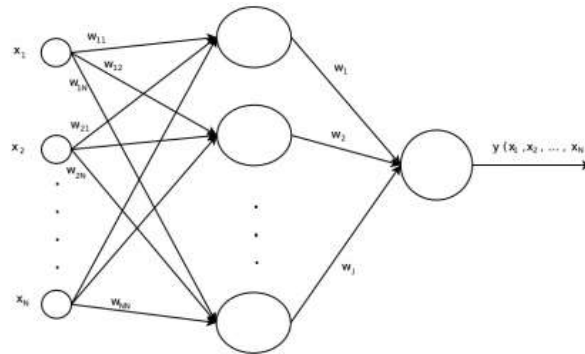


Figura 5. Arquitectura de una red neuronal típica
Fuente: MathWorks, 2019

Por lo tanto la salida de la red neuronal artificial se expresa como la ecuación (8).

$$y(x_1, x_2, \dots, x_N) = \sum_{j=1}^J w_j \sigma_j \left(\sum_{i=1}^N w_{ij} x_i + \theta_j \right) \quad (8)$$

Donde J es el número de neuronas de la capa oculta, σ_j una función de transferencia, w_j el peso sináptico entre la neurona j y la neurona de la última capa y w_{ij} el peso sináptico entre la entrada i y la neurona j de la capa oculta.

2.6.1.1.2 Feed-back propagation

Las redes neuronales recurrentes poseen varios caminos de retroalimentación entre todos los elementos, es decir, una neurona está conectada a las neuronas posteriores en la capa siguiente (Oropeza, 2007).

Las redes de avance consisten en una serie de capas, en el caso de la primera capa, existe una conexión desde la entrada de red. Cada capa posterior tiene una conexión desde la capa anterior. La capa final produce la salida de la red (MathWorks, 2019).

Una red de alimentación directa con una capa oculta y suficientes neuronas en las capas ocultas puede adaptarse a cualquier problema de mapeo de entrada-salida finito. La información que es cambiada en función de los pesos de cada neurona, se transmite por la red por las diferentes capas ocultas y que al final es procesada por la capa de salida (MathWorks, 2019).

2.6.1.1.3 Levenberg-Marquardt

El algoritmo Levenberg-Marquardt fue creado para acercarse a la velocidad de entrenamiento de segundo orden sin necesidad de calcular la matriz de Hesse.

Cuando la función de rendimiento tiene la forma de una suma de cuadrados (como es típico en el entrenamiento de redes de alimentación directa), la matriz de Hesse se puede aproximar como $H = J^T J$ y el gradiente puede ser calculado como $g = J^T e$ donde J es la matriz jacobiana que contiene las primeras derivadas de los errores de red con respecto a los pesos y sesgos, y “e” es un vector de errores de red.

La matriz jacobiana se puede calcular a través de una técnica estándar de retropropagación que es mucho menos compleja que calcular la matriz de Hesse. El algoritmo de Levenberg-Marquardt utiliza esta aproximación a la matriz de Hesse en la siguiente actualización de tipo Newton como se indica en la ecuación (8) (Hagan & Menhaj, 1999).

$$X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e \quad (8)$$

2.6.1.1.4 Función de Transferencia

Una función de transferencia es un modelo matemático que relaciona la señal de salida con una señal de entrada o excitación (también modelada). Constantemente se usan las funciones de transferencia para relacionar las entradas y las salidas a través de ecuaciones diferenciales lineales e invariantes en el tiempo (Perez, Perez, & Perez, 2008).

Función de Transferencia Tangente Hiperbólica (tansig)

Es una función empleada habitualmente en redes neuronales debido a que es flexible y posee un rango amplio de resultados (Campos, Torres, & Paz, 2014). La función tangente sigmoidea hiperbólica presenta resultados de 1 y -1, donde corresponde $a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$, como se indica en la *Figura 6*.

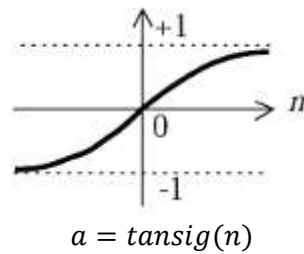


Figura 6. Función de transferencia Tangente Sigmoidal Hiperbólica

Fuente: (Campos, Torres, & Paz, 2014)

Función de Transferencia Lineal

La función de transferencia lineal presenta como resultado la salida igual a su entrada como se indica en la *Figura 7*, donde $a = n$ (Campos, Torres, & Paz, 2014).

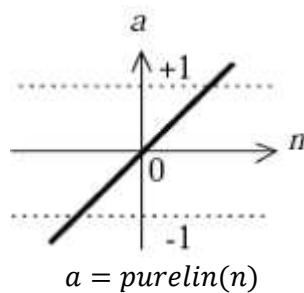


Figura 7. Función de transferencia lineal

Fuente: (Campos, Torres, & Paz, 2014)

2.6.1.2 Redes Neuronales Artificiales Recurrentes

Las redes neuronales artificiales recurrentes son aquellas que están constituidas por realimentación, lo cual permite que mediante bucles la información persista durante las diferentes épocas de entrenamiento y que las capas de salida se interconecten nuevamente con los datos de entrada, es decir, recuerdan las salidas como entrada (Rivas & Mazón, 2018). Las redes que cuentan con propagación hacia atrás, que presentan lazos cerrados se denominan sistemas recurrentes (Rivas & Mazón, 2018).

Existen dos metodologías básicas de entrenamiento de redes recurrentes; la retropropagación a través del tiempo y aprendizaje recurrente al tiempo real (Gómez, 2008).

2.7 Coeficiente de determinación

Es un indicador o estadístico empleado ampliamente en la determinación la calidad de un modelo específico. Permite medir el grado de asociación entre dos variables. Se expresa como muestra la ecuación (9) (Novales, 2010).

$$R^2 = \frac{\sigma^2_{xy}}{\sigma^2_x \sigma^2_y} \quad (9)$$

Donde σ_{xy} es la covarianza de las variables x e y, σ^2_x corresponde a la varianza de x y σ^2_y es la varianza de y.

El R^2 alcanza valores de 0 a 1. En los casos que el valor sea próximo a 1, se traduce que el modelo posee un elevado ajuste a los datos. En caso contrario si el valor del coeficiente de terminación se acerca a cero, indica que los valores estimados no logran explicar las variaciones de la variable, por lo tanto el modelo no representa gran confianza (Novales, 2010).

No obstante, el algunas ocasiones la muestra presenta pocas observaciones, por lo que si existen escasos residuos elevados pueden inducir un R^2 bajo (cerca a cero), dando la idea de que el modelo no es bueno, a pesar que exceptuando los residuos mencionados el ajuste puede ser óptimo. Y viceversa si existen pocos valores observados y no existen residuos altos, el coeficiente de determinación tendrá un valor alto, que indicaría un resultado erróneo de un buen ajuste, pero realmente constituye a un indicador de insuficiente información (Novales, 2010).

CAPÍTULO III

MARCO METODOLÓGICO

3.1 Área de estudio

El Distrito Metropolitano de Quito (DMQ), está localizado en la provincia de Pichincha, limitado al norte por la provincia de Imbabura, al sur con los cantones Rumiñahui y Mejía, al este por los cantones Pedro Moncayo, Cayambe y provincia de Napo y al oeste por los cantones Pedro Vicente Maldonado, Los Bancos y la provincia de Santo Domingo de los Tsáchilas. Consta de 9 administraciones zonales, las cuales poseen 32 y 33 parroquias urbanas y rurales respectivamente. Tiene una superficie aproximada de 4.183 Km² (Gobierno Provincial de Pichincha, 2017). Su población corresponde a 2'239.191 habitantes (INEC, 2010).

El Instituto Nacional de Meteorología e Hidrología (INAMHI) es el servicio encargado de informar sobre el tiempo, clima y los recursos hídricos del Ecuador, incluyendo el monitoreo y pronóstico meteorológico. Posee una red UV que consta de 9 estaciones como se muestra en la *Figura 8* en puntos color marrón (INAMHI, 2019).



Figura 8. Red UV INAMHI
Fuente: (INAMHI, 2017)

En la jurisdicción del DMQ existen 5 estaciones con información de radiación solar que pertenecen al Fondo para la Protección del Agua (FONAG). FONAG es un ente encargado de la protección de las cuencas hídricas que proveen de agua al Distrito Metropolitano de Quito, mediante la cooperación, la investigación y uso de tecnología (FONAG, 2019). En la *Figura 9* se indica la red de estaciones que cuentan con registro de radiación solar.

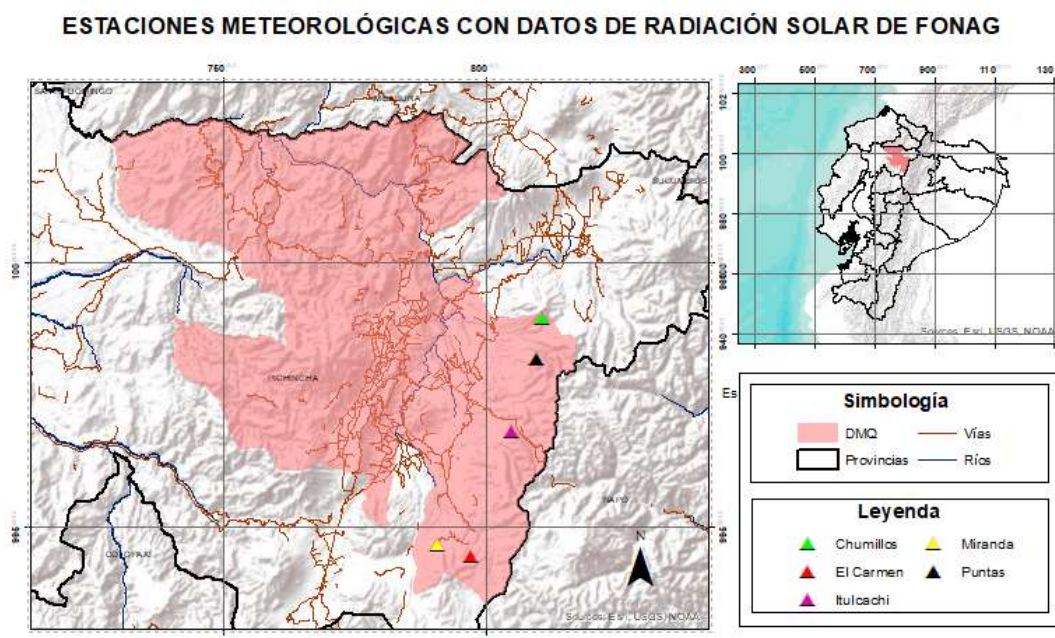


Figura 9. Estaciones con información de radiación solar FONAG

Fuente: Elaboración propia

3.2 Recopilación de datos

Los datos se obtuvieron a través de FONAG, de las estaciones Miranda (M5028), El Carmen (M5029), Chumillos (M5031), Puntas (M5074) e Itulcachi (M5075). Las variables que fueron empleadas para generar las series de tiempo en cada una de las estaciones son las siguientes: dirección del viento, humedad del aire, precipitación, radiación solar, temperatura del aire, velocidad del viento y humedad del suelo. La humedad del suelo no corresponde netamente a una variable meteorológica, sin embargo se ha tomado en consideración por ser un aspecto importante al mostrar una condición significativa del entorno y del comportamiento de la radiación.

La información fue entregada en formato *.xls* para cada estación con sus respectivas variables en un periodo de 5 años (varía por estación) con datos cada hora.

3.3 Generación de series temporales de las variables meteorológicas

Las series temporales se realizaron en función de las horas de mayor incidencia de radiación solar, debido a que corresponde al periodo en el cual se requiere énfasis, tanto para el aprovechamiento y la prevención para la población. Las horas más relevantes en incidencia corresponden de 10 de la mañana hasta las 4 de la tarde a nivel mundial, según la Agencia de Protección Ambiental de los Estados Unidos (EPA, 2001). Se realizaron las series temporales para las 10H00, 12H00, 14H00 y 16H00 (hora local, UTC-5) para cada estación con su respectiva variable como se muestra en la Tabla 2.

Es decir cada variable supone un total de 1643 datos por cada estación y por cada hora en análisis, criterio recomendado por Ayala (2017) donde menciona la importancia de trabajar con datos diarios.

Tabla 2

Estaciones meteorológicas y variables para generación de series de tiempo

Estaciones	Variables	Unidad	Inicio	Fin	Horas
	Dirección del viento	°			
<i>Miranda (M5028)</i>	Humedad del aire	%			
<i>El Carmen (M5029)</i>	Humedad del suelo	%	1 de enero	1 de julio	10
<i>Chumillos (M5031)</i>	Precipitación	mm	del 2014	del 2018	12
<i>Puntas (5074)</i>	Radiación solar	W/m ²			14
<i>Itulcachi (5075)</i>	Temperatura del aire	° C			16
	Velocidad del viento	m/s			

3.3.1 Relleno de datos en las series temporales y obtención de coeficientes de correlación

Una vez que se obtuvieron las series temporales para cada estación con sus respectivas variables a las horas determinadas, se identificaron vacíos de información, conocidos como “gaps”. En el caso de la zona andina del Ecuador existe escasez de información meteorológica por lo cual se toma el umbral planteado por Campozano et al. (2015) en su estudio, el cual corresponde a la aceptación de hasta 25% de datos faltantes.

La totalidad de las series en análisis cumplieron con el requisito planteado, por lo cual se empleó la información de las 5 estaciones y el conjunto total de variables sin excepción alguna.

Existen varios métodos que permiten obtener los datos faltantes, en el presente estudio se aplicó el método Holt-Winters doble suavizado exponencial, a través del uso del software libre *RStudio* y el código R 3.4.4.. Para proceder al uso de los diversos beneficios que brinda R, fue necesario cargar los paquetes que ofrecen al usuario la optimización de recursos, tiempo de análisis y procesamiento. Se emplearon los paquetes que se detallan en la Tabla 3.

Tabla 3

Paquetes empleados en el software R

Paquete	Descripción	Autor (es)
<i>tseries</i>	El análisis de series de tiempo y finanzas computacionales	Adrian Trapletti, Kurt Hornik, Blake LeBaron
<i>TSA</i>	Contiene funciones R y conjuntos de datos	Kung-Sik Chan, Brian Ripley
<i>forecast</i>	Métodos y herramientas para mostrar y analizar pronósticos de series de tiempo univariantes que incluyen suavizado exponencial a través de modelos de espacio de estado y modelado automático ARIMA.	Rob Hyndman, George Athanasopoulos, Christoph Bergmeir, Gabriel Caceres, Leanne Chhay, Mitchell O'Hara-Wild, Fotios Petropoulos, Slava Razbash

Fuente: (R Core Team, 2019)

Una vez que los paquetes han sido instalados, para obtener los datos faltantes en las diferentes series temporales, se procedió a ejecutar el algoritmo que se presentan a continuación:

1. Leer de los datos mediante la función “*read.table*”
2. Convertir los datos en una serie temporal a partir de la función “*ts*”
3. Aplicar la función “*HoltWinters*” a la serie temporal para el modelamiento de la misma
4. Predecir los valores mediante la función “*predict*”
5. Plotear los resultados a través de la función “*plot*”
6. Obtener los resultados y colocar los datos obtenidos en los gaps de la base de datos

El algoritmo se aplicó para cada de una de las variables, para todas las estaciones y para las 4 horas establecidas, es decir un total de 140 series temporales que fueron analizadas y rellenadas a través del método mencionado.

Finalmente se aplicó el coeficiente de correlación para cada estación con su respectiva hora mediante la herramienta de análisis de datos en Excel, haciendo énfasis en la correlación que existe entre las variables meteorológicas y la radiación solar.

3.4 Aplicación de la transformada de wavelet a las series temporales de radiación

Una vez que las series han sido rellenadas, se procedió a aplicar la transformada de wavelet con el objetivo de eliminar el ruido de las series temporales de radiación solar mediante la familia db4 nivel de discretización 3, criterio que ha sido recomendado en el estudio realizado por Pratap Yadav & Laxmidhar (2014). Para llevar a cabo el propósito, se utilizó el software *MatLab R2015a* con licencia estudiantil en lenguaje M, propio del programa.

En primer lugar se exportó los datos en formato *.txt*, y mediante la función “*load*” se importó la información al espacio de trabajo, se asignó una variable N para que contenga los datos de la serie, después se colocó la transformada de la familia daubechies que corresponde a “*db4*”, nivel 3.

A continuación se generó una matriz en la cual mediante la función “*wavedec*” se obtiene la descomposición wavelet de la señal en función del nivel y tipo de wavelet que en este caso corresponde a 3 y db4 respectivamente.

Una vez obtenida la estructura de descomposición wavelet, se procede a extraer los coeficientes a través de la función “*detcoef*” y luego se obtiene la desviación estándar de dichos coeficientes mediante la función “*std*”.

Se calculó el umbral de la transformada, para ello se empleó la fórmula que traducida en código corresponde a $umbral = desv_{tip} * sqrt(2 * log(length(N)))$. También es necesaria la función “*wthresh*” que retornó el umbral otorgado un suavizado o crudo, en este caso se usará umbral duro o “hard” que en código corresponde a “*h*” debido a que no se demanda suavizado en este caso por la naturaleza de los datos.

Finalmente se aplicó la función “*waverec*” que realiza una reconstrucción wavelet unidimensional multinivel y es así como se obtiene la serie final sin ruido. Este proceso se aplicó a las 5 series temporales de radiación solar.

3.5 Redes neuronales artificiales recurrentes

3.5.1 Creación de redes neuronales

Para generar las redes neuronales, se utilizó el software *MalLab R2015a* con licencia estudiantil. Los paquetes que posee el software dan facilidades al usuario mediante una interfaz amigable y que permite optimizar tiempo y recursos. En este caso se usó la herramienta “*nntool*”, la cual llama a una nueva ventana donde se sintetiza los elementos más importantes y permite trabajar de manera ordenada y eficiente como se indica en la *Figura 10*.

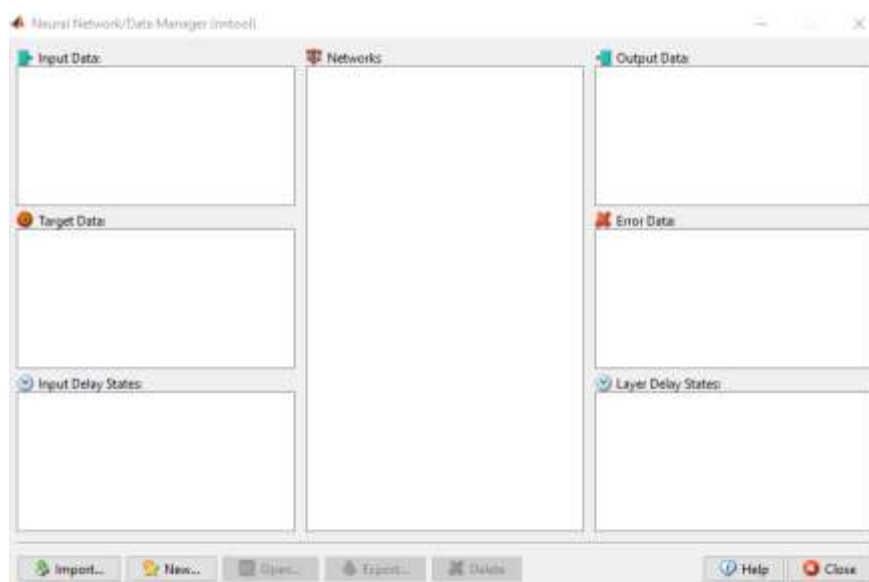


Figura 10. Interfaz de *nntool* para Redes Neuronales en MATLAB R2015a

Para importar la información, se creó variables dentro del espacio de trabajo como se indica en la *Figura 11*, para que tengan el formato requerido para el uso en la herramienta “*nntool*”.

	1	2	3	4	5	6	7	8	9	10	11	12	13
red10new	1944	1845	1948	1647	1648	1848	1830	3031	1852	1653	1834	1855	3038
red10new	290.9940	179.5303	167.8267	170.1800	185.8300	152.7300	157.5600	155.8900	155.7967	151.9467	153.3833	156.1133	154.7033
pred_entrada10new	35.8340	47.1888	51.8188	55.1433	53.3057	57.5595	50.8180	52.8278	50.2555	52.2138	51.2742	50.6190	52.9487
red10new	34.7000	33.9050	33.3100	32.7800	32.1283	31.6767	31.2317	30.8000	30.4000	30	29.5717	29.1600	28.7000
red10new	0	0	0	0	0	0	0	0	0	0	0	0	0
red10new	14.8168	14.2987	12.8752	11.6683	13.0842	12.5205	13.7630	12.6825	12.8873	12.8487	12.5617	12.5678	12.4785
red10new	0.2711	2.5486	4.7528	8.6618	2.3263	9.2393	7.2189	11.1886	10.8273	11.0887	8.4362	12.4250	8.5303

Figura 11. Interfaz para creación de variables en MATLAB R2015a

Se generó una nueva red neuronal en la herramienta “*nntool*” botón “New” con las especificaciones mostradas en la Tabla 4. Cabe recalcar que se generaron 4 redes neuronales por cada estación.

Tabla 4*Criterios seleccionados para creación de las redes neuronales artificiales*

Criterio	Selección	Descripción
<i>Tipo de red</i>	Feed-forward backprop	Tipo de red neuronal artificial que usa retroalimentación para el aprendizaje
<i>Input data</i>	Variables: tiempo, dirección del viento, humedad del aire, humedad del suelo, precipitación, temperatura y velocidad del viento	Matriz de variables meteorológicas incluida la variable tiempo con valor de 7x1643, los datos corresponde a valor diarios de 4 años y 6 meses
<i>Target data</i>	Variable radiación	Matriz de variable radiación con valor de 7x1643, los datos corresponde a valor diarios de 4 años y 6 meses
<i>Función de entrenamiento</i>	TRAINLM	Función de entrenamiento que actualiza los valores de peso y sesgo según la optimización de Levenberg-Marquardt, es el algoritmo más rápido para este tipo de redes, pero usa mayor memoria (MathWorks, 2019)
<i>Función de aprendizaje adaptativa</i>	LEARNGDM	Calcula el cambio de peso para una neurona dada a partir de la entrada de la neurona y el error, el peso, sesgo, la tasa de aprendizaje y la constante de momento, según el descenso del gradiente con momento (MathWorks, 2019)
<i>Rendimiento de la función</i>	MSE	Mide el rendimiento de la red de acuerdo con el error cuadrático medio

(Continúa)

<i>Número de capas</i>		2	El número de capas presentes en la arquitectura de la red
<i>Propiedades capa 1</i>	Número de neuronas	15	Número de neuronas presentes en la primera capa que corresponde a la capa oculta
	Función de transferencia	TANSIG	Función de transferencia sigmoidea tangente hiperbólica que calcula la salida de una capa a partir de su entrada neta.
<i>Propiedades capa 2</i>	Número de neuronas	1	Número de neuronas presentes en la segunda capa
	Función de transferencia	PURELIN	Función de transferencia lineal de la capa 2 que corresponde a las salidas

3.5.2 Entrenamiento de las redes neuronales

Una vez que las redes fueron creadas se procedió a el entrenamiento mediante la función “*nntraintool*”, en la cual se puede observar el progreso del entrenamiento en función de los parámetros establecidos como se visualiza en la *Figura 12*.

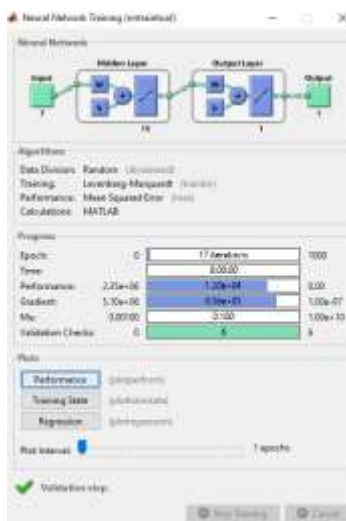


Figura 12. Interfaz de la ventana de entrenamiento

Cabe recalcar que los datos de entrada fueron divididos en 3 grupos mediante la función “*dividerand*”, en la que el primer grupo de datos corresponden a los valores que permitieron el entrenamiento de la red (70%), el segundo grupo fueron los que dieron paso a la verificación en el entrenamiento actual (15%) y el conjunto final restante que se empleó para realizar las pruebas de la predicción de la red (15%).

El entrenamiento general de las redes se aceptó una vez que fueron analizados tanto el entrenamiento, validación y el test de prueba en función del R que muestra en la regresión de los resultados (*Figura 13*). Se admitió el criterio de una correlación alta en función del rango expuesto por Lind et al. (2004).

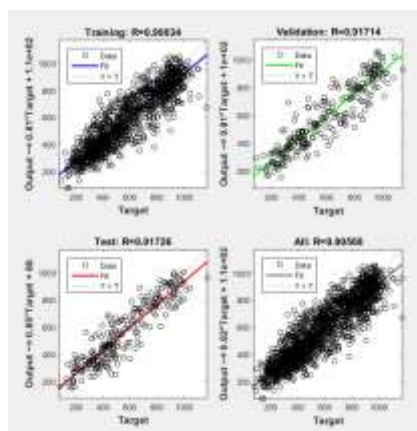


Figura 13. Interfaz de la ventana de resultados del entrenamiento

3.5.3 Predicción mediante redes neuronales

Una vez que las redes neuronales fueron entrenadas y cumplieron con el criterio establecido para la aceptación, se procedió a predecir los valores de radiación. Para ello se usó nuevamente la herramienta “*nntool*”, en la cual se ingresó los valores de entrada o “inputs” que corresponden al tiempo, dirección del viento, humedad del aire, humedad del suelo, precipitación, temperatura y velocidad del viento, en donde el t_0 es 1644, es decir la continuación del tiempo t con el que la red fue entrenada.

Luego se ubicó la red entrenada en la sección “Networks” y en la selección de “Simulate” como se observa en la *Figura 14*, se indica los “inputs” que ya fueron cargados inicialmente

(variables) y en la sección “outputs” se colocó el nombre de los valores de salida, que en este caso corresponde a la radiación que ha sido predicha.

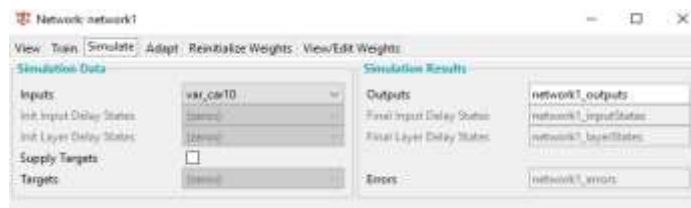


Figura 14. Interfaz de la ventana de Simulación

Una vez que el software indicó que se generó los valores, se procede a exportar la información y ubicarla en la base de datos.

3.5.4 Validación

La validación de las predicciones de las redes entrenadas para las estaciones con sus respectivas horas, se realizó mediante la comparación de los datos reales con los predichos de radiación. Se destinó un porcentaje de datos de las series temporales de radiación real menor a 6 meses (varía por estación) como se indica en la *Figura 15*. Se analizó el coeficiente de determinación entre los datos reales y los pronosticados para valorar la predicción.

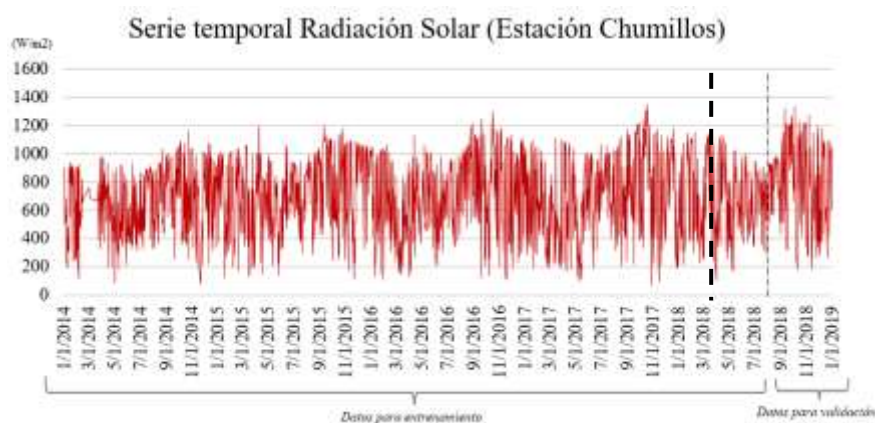


Figura 15. Serie temporal de radiación solar (datos de entrenamiento y validación)

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1 Resultado de la generación de las series temporales de las variables meteorológicas

4.1.1 Relleno de series temporales

Las series temporales fueron rellanadas mediante el método Holt-Winters de manera satisfactoria, existen 5 estaciones con 7 variables y 4 horas analizadas por lo que se tiene un total de 140 series temporales.

A manera de ejemplo se presenta el resultado visual de la estación Miranda (M5028) en el tiempo 10H00, el cual corresponde al tratamiento que se realizó para todas las series de tiempo. La base de datos de las series temporales completas se encuentran en los entregables del estudio, clasificadas por estación, variable y hora en formato .xlsx.

En las *Figuras 16, 17, 18, 19, 20, 21 y 22* se puede observar gráficamente como trabaja el método Holt-Winters doble suavizado exponencial, en color negro se puede visualizar los valores observados, es decir los valores que presenta la serie, mientras que en rojo se pueden observar los valores predichos o modelados (*Figuras* fueron desarrolladas en el software *RStudio*).

Este método es recomendado para series temporales que no presenta variación estacional y con tendencia, además permite resolver las limitaciones de los metodologías deterministas (Justel & Cayuela, 2015), como este caso fue usado para relleno de vacíos y a una hora específica de datos diarios, es adecuado para el propósito. Se usó los periodos con mayor cantidad de datos para predecir los datos faltantes los cuales no sobrepasaron el 25% del total.

En diversas ocasiones cuando se trabajan con variables cuantitativas continuas se asume que los procesos corresponden a una distribución normal o gaussiana. En el caso de variables meteorológicas no siempre obedecen a este tipo de distribución.

Las variables meteorológicas poseen sin número de estudios e investigaciones en busca de comprender su dinámica por su naturaleza fluctuante. En estudios como los de Ayala (2017) se infiere sobre la dinámica caótica para series temporales de variables meteorológicas. Es por ello que es importante analizar los diversos componentes que posee la misma, tanto para la predicción como para el relleno de datos ausentes.

Carrera et al. (2016) realizaron en su estudio el relleno de series anuales de datos meteorológicos mediante regresión lineal y la razón normal que permite calcular los valores faltantes en función de valores promedio de estaciones de referencia.

El método Holt-Winters tiene la posibilidad de adaptarse al cambio del comportamiento y de los parámetros del modelo de la serie temporal (Prajakta, 2004), por lo que en variables meteorológicas representa una técnica acertada. Normalmente se usa este método para la predicción propiamente; sin embargo, para el relleno de datos representa una herramienta importante porque desde un inicio se analiza los elementos de la serie.

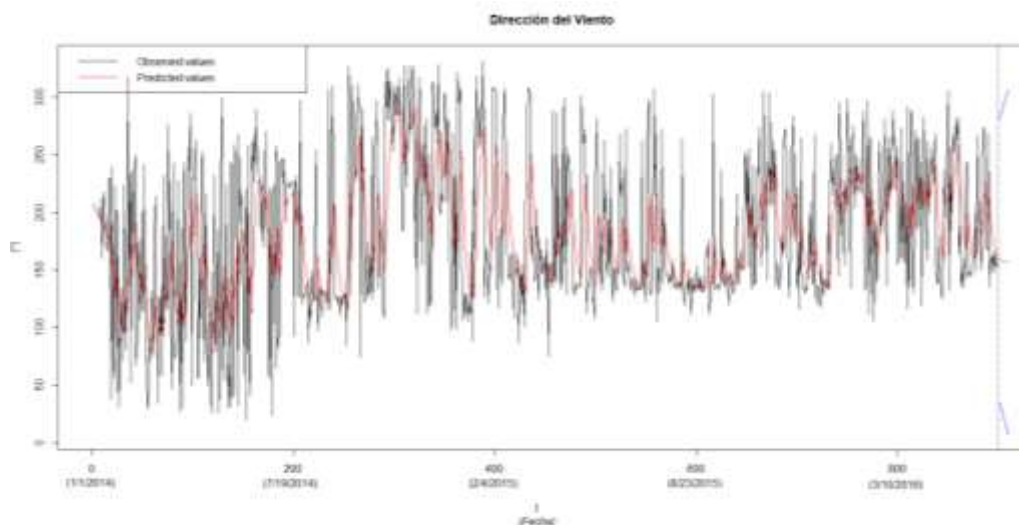


Figura 16. Gráfico de la serie temporal de la variable dirección del viento– Miranda

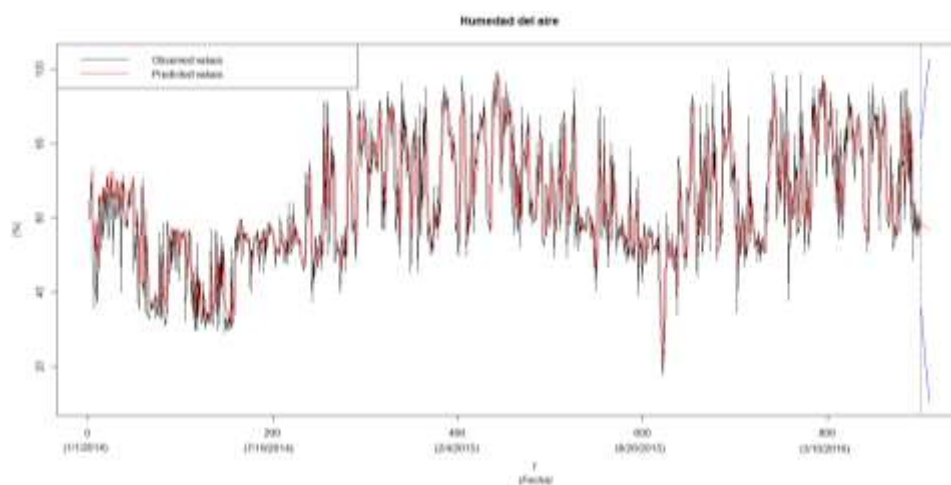


Figura 17. Gráfico de la serie temporal de la variable humedad del aire– Miranda

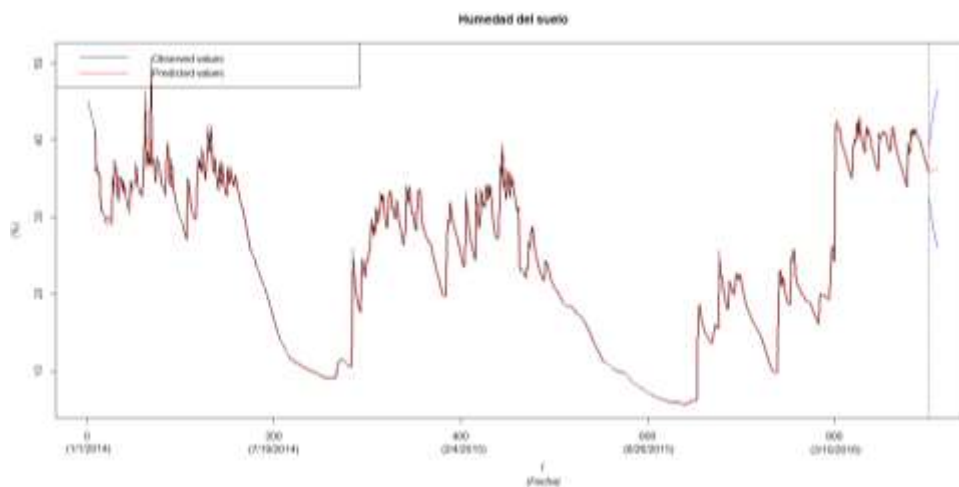


Figura 18. Gráfico de la serie temporal de la variable humedad del suelo – Miranda

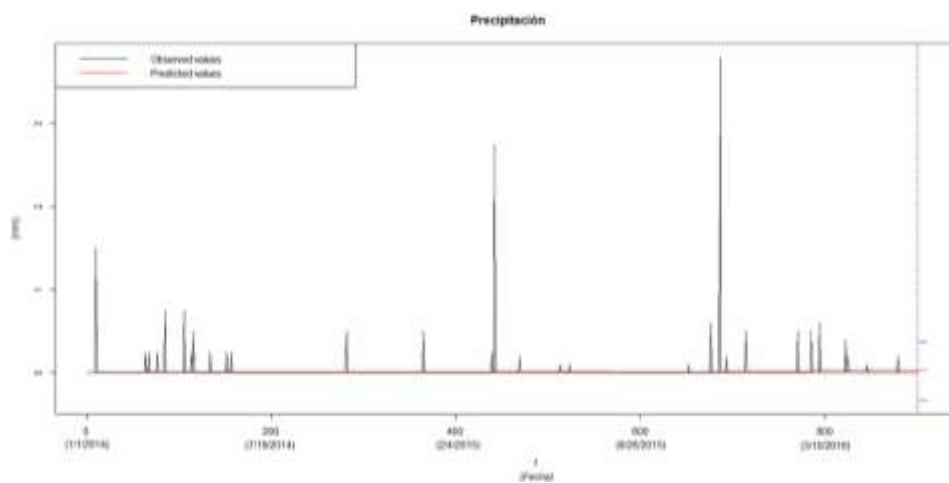


Figura 19. Gráfico de la serie temporal de la variable precipitación – Miranda

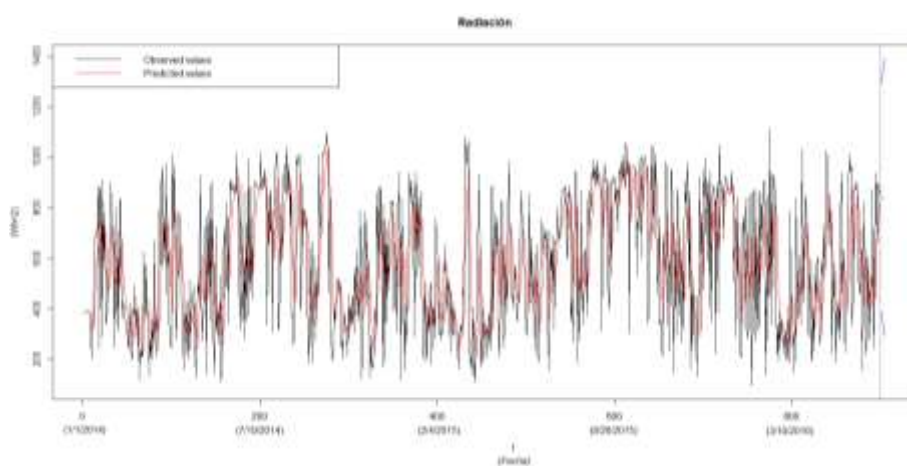


Figura 20. Gráfico de la serie temporal de la variable radiación - Miranda

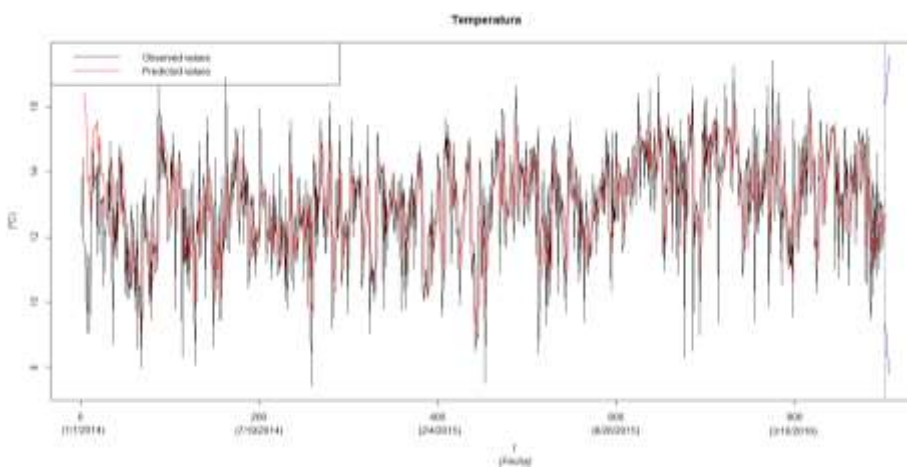


Figura 21. Gráfico de la serie temporal de la variable temperatura – Miranda

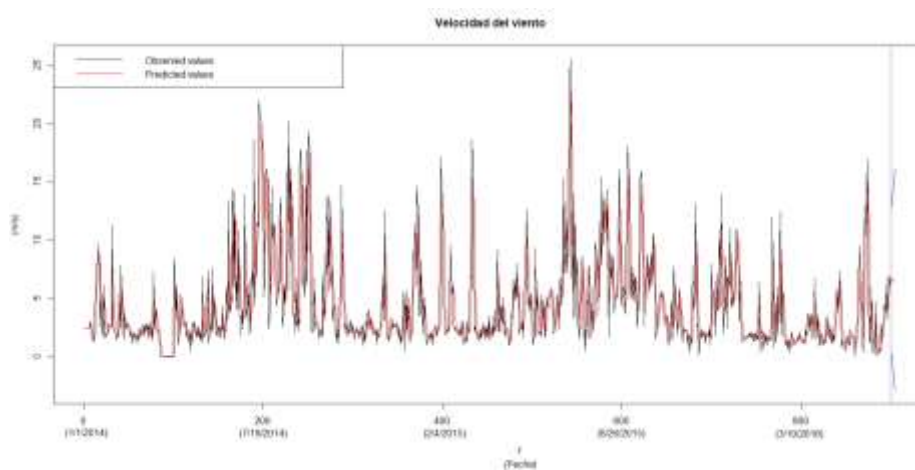


Figura 22. Gráfico de la serie temporal de la variable velocidad del viento – Miranda

4.1.2 Coeficientes de correlación de las variables meteorológica

Se obtuvo los coeficientes de correlación de la totalidad de variables en sus respectivas horas (10H00, 12H00, 14H00 y 16H00) y estaciones. Para identificar la intensidad y dirección de los coeficientes se empleó los criterios de Lind et al. (2004). Donde DV: dirección del viento; HA: humedad del aire; HS: humedad del suelo; P: precipitación; T: temperatura; VV: velocidad del viento y R: radiación solar.

Tabla 5

Coeficientes de correlación entre variables meteorológicas de la estación Miranda

Miranda	Hora	DV	HA	HS	P	T	VV	R
DV	10	1	0,4	0,05	-0,02	-0,06	-0,25	-0,20
	12	1	0,34	0,05	0,02	-0,02	-0,25	-0,23
	14	1	0,31	0,15	0,06	-0,06	-0,27	-0,20
	16	1	0,03	0,08	0,00	0,00	-0,14	-0,02
HA	10		1	0,17	0,16	-0,44	-0,46	-0,57
	12		1	0,21	0,11	-0,51	-0,47	-0,57
	14		1	0,19	0,09	-0,58	-0,46	-0,61
	16		1	0,13	0,23	-0,52	-0,50	-0,53
HS	10			1	0,03	-0,12	-0,35	-0,32
	12			1	0,05	-0,12	-0,37	-0,38
	14			1	0,07	-0,15	-0,37	-0,31
	16			1	0,07	-0,18	-0,36	-0,30
P	10				1	-0,27	-0,10	-0,17
	12				1	-0,05	-0,09	-0,09
	14				1	-0,04	-0,11	-0,10
	16				1	-0,32	-0,14	-0,27
T	10					1	0,09	0,61
	12					1	0,08	0,48
	14					1	0,16	0,58
	16					1	0,30	0,59
VV	10						1	0,55
	12						1	0,62
	14						1	0,61
	16						1	0,51
R	10							1
	12							1
	14							1
	16							1

Los resultados obtenidos para la estación Miranda se muestran sintetizados en la Tabla 5. El valor más alto de correlación corresponde a 0,62 entre la variable radiación solar y velocidad del viento y el valor más bajo de -0,02 entre la radiación solar y la dirección del viento.

Tabla 6

Coefficientes de correlación entre variables meteorológicas de la estación El Carmen

El Carmen	Hora	DV	HA	HS	P	T	VV	R
DV	10	1	0,01	0,02	0,01	-0,01	-0,21	-0,04
	12	1	0,04	0,00	0,08	-0,05	-0,21	-0,20
	14	1	0,07	0,03	0,07	-0,07	-0,22	-0,21
	16	1	0,19	0,04	0,03	-0,11	-0,22	-0,22
HA	10		1	0,12	0,21	-0,63	0,09	-0,62
	12		1	0,14	0,26	-0,66	-0,02	-0,41
	14		1	0,17	0,25	-0,67	-0,10	-0,49
	16		1	0,25	0,18	-0,54	-0,23	-0,55
HS	10			1	0,01	0,03	-0,04	-0,06
	12			1	0,03	0,01	-0,04	-0,07
	14			1	0,05	0,01	-0,05	-0,03
	16			1	0,07	-0,01	-0,04	-0,04
P	10				1	-0,28	-0,08	-0,19
	12				1	-0,27	-0,20	-0,30
	14				1	-0,31	-0,23	-0,34
	16				1	-0,20	-0,20	-0,19
T	10					1	0,09	0,61
	12					1	-0,15	0,36
	14					1	-0,06	0,40
	16					1	-0,08	0,30
VV	10						1	0,23
	12						1	0,47
	14						1	0,55
	16						1	0,59
R	10							1
	12							1
	14							1
	16							1

En la Tabla 6 se puede observar los coeficientes de correlación a las horas indicadas para la estación El Carmen. El coeficiente más alto con respecto a la radiación solar corresponde a -0,62 con la variable humedad del aire a las 10H00 y el más bajo corresponde a -0,03 con la variable humedad del suelo a las 14H00.

Tabla 7

Coefficientes de correlación entre variables meteorológicas de la estación Chumillos

Chumillos	Hora	DV	HA	HS	P	T	VV	R
DV	10	1	0,12	-0,08	-0,04	0,02	-0,06	-0,09
	12	1	0,11	-0,07	0,00	0,01	-0,06	-0,10
	14	1	0,15	-0,04	0,01	-0,03	-0,05	-0,17
	16	1	0,18	-0,01	0,04	-0,09	-0,06	-0,13
HA	10		1	0,24	0,21	-0,57	-0,48	-0,73
	12		1	0,25	0,29	-0,45	-0,50	-0,70
	14		1	0,27	0,32	-0,71	-0,55	-0,73
	16		1	0,25	0,29	-0,69	-0,60	-0,78
HS	10			1	0,07	-0,04	-0,24	-0,12
	12			1	0,03	-0,05	-0,23	-0,14
	14			1	0,16	-0,13	-0,24	-0,19
	16			1	0,07	-0,13	-0,21	-0,18
P	10				1	-0,27	-0,13	-0,21
	12				1	-0,14	-0,21	-0,27
	14				1	-0,37	-0,19	-0,33
	16				1	-0,33	-0,19	-0,30
T	10					1	0,02	0,64
	12					1	-0,03	0,27
	14					1	0,24	0,59
	16					1	0,27	0,66
VV	10						1	0,51
	12						1	0,64
	14						1	0,68
	16						1	0,63
R	10							1
	12							1
	14							1
	16							1

El valor más alto de correlación de la variable de interés que corresponde a la radiación solar fue -0,78 con la humedad del aire a las 16H00 y el valor mínimo fue de -0,09 con la variable dirección del viento a las 10H00. Los resultados totales de los coeficientes de correlación en la estación Chumillos se presentan a la Tabla 7.

Tabla 8

Coefficientes de correlación entre variables meteorológicas de la estación Puntas

Puntas	Hora	DV	HA	HS	P	T	VV	R
DV	10	1	0,08	0,18	-0,05	-0,13	-0,08	-0,02
	12	1	0,10	0,15	0,03	-0,12	-0,14	-0,04
	14	1	0,13	0,12	0,06	-0,04	-0,18	-0,04
	16	1	0,13	0,13	0,07	-0,11	-0,15	-0,12
HA	10		1	0,17	0,21	-0,65	-0,02	-0,49
	12		1	0,18	0,20	-0,59	-0,04	-0,45
	14		1	0,21	0,23	-0,45	-0,13	-0,45
	16		1	0,21	0,20	-0,59	-0,12	-0,61
HS	10			1	0,03	-0,15	0,03	-0,03
	12			1	0,04	-0,15	-0,02	-0,01
	14			1	0,07	-0,15	0,02	-0,03
	16			1	0,09	-0,15	-0,01	-0,08
P	10				1	-0,25	-0,09	-0,26
	12				1	-0,16	-0,12	-0,25
	14				1	-0,05	-0,17	-0,23
	16				1	-0,19	-0,14	-0,27
T	10					1	0,02	0,64
	12					1	-0,11	0,30
	14					1	-0,30	0,32
	16					1	-0,32	0,30
VV	10						1	0,46
	12						1	0,45
	14						1	0,35
	16						1	0,47
R	10							1
	12							1
	14							1
	16							1

La Tabla 8 muestra los coeficientes de correlación de la estación Puntas donde el valor más alto de correlación de la variable de interés que corresponde a la radiación solar fue 0,64 con la temperatura a las 10H00 y el valor mínimo fue de -0,01 con la variable humedad del suelo a las 12H00.

Tabla 9

Coefficientes de correlación entre variables meteorológicas de la estación Itulcachi

Itulcachi	Hora	DV	HA	HS	P	T	VV	R
DV	10	1	0,05	0,07	-0,02	-0,02	-0,13	-0,02
	12	1	0,01	0,06	0,00	0,00	-0,15	0,00
	14	1	0,03	0,07	-0,01	-0,01	-0,09	-0,02
	16	1	0,05	0,08	0,01	-0,01	-0,06	-0,02
HA	10		1	0,12	0,16	-0,72	0,07	-0,65
	12		1	0,12	0,19	-0,78	0,03	-0,60
	14		1	0,15	0,18	-0,80	-0,05	-0,58
	16		1	0,18	0,18	-0,73	-0,14	-0,59
HS	10			1	0,01	-0,09	0,03	-0,02
	12			1	-0,01	-0,10	0,04	0,00
	14			1	0,02	-0,12	0,03	-0,04
	16			1	0,02	-0,11	-0,02	-0,03
P	10				1	-0,23	-0,07	-0,20
	12				1	-0,23	-0,14	-0,25
	14				1	-0,20	-0,17	-0,27
	16				1	-0,20	-0,15	-0,24
T	10					1	-0,35	0,60
	12					1	-0,24	0,58
	14					1	-0,17	0,53
	16					1	-0,17	0,46
VV	10						1	0,33
	12						1	0,44
	14						1	0,56
	16						1	0,55
R	10							1
	12							1
	14							1
	16							1

El valor más alto de los coeficientes relacionados con la variable radiación solar de la estación Itulcachi fue de -0,65 a las 10H00 y el menor valor de correlación fue de 0,00 en la variable dirección del viento a las 12H00 como se indica en la Tabla 9.

La correlación más alta de la totalidad de estaciones fue -0,78, que corresponde a la humedad del aire a las 16H00 en la estación Chumillos, la cual indica correlación negativa intensa. Por otra parte, la más baja fue de 0,00 en la variable dirección del viento a las 12H00 de la estación Itulcachi, lo cual indica una relación nula o sin correlación. El 68,90 % de los coeficientes en la totalidad correspondieron una correlación débil y el 31,1% fue una correlación intensa. Las variables con correlación intensa en forma general fueron la temperatura y la humedad del aire.

4.2 Aplicación de la transformada de wavelet a series temporales de radiación

Las series temporales de radiación solar de las 5 estaciones fueron sometidas al análisis mediante la transformada de wavelet de la familia db4 nivel de discretización con el objetivo de eliminar el ruido presente (las Figuras fueron desarrolladas en el software *MatLab R2015a*).

En la *Figura 23* se indica en color rojo la serie temporal de la variable radiación solar de la estación Miranda, en color azul se identificó la serie sin ruido, es decir, la serie temporal resultante de la aplicación de la transformada de wavelet y en color negro se puede observar la información considerada ruido. El porcentaje de energía retenida del análisis corresponde a 84,43% de la información.

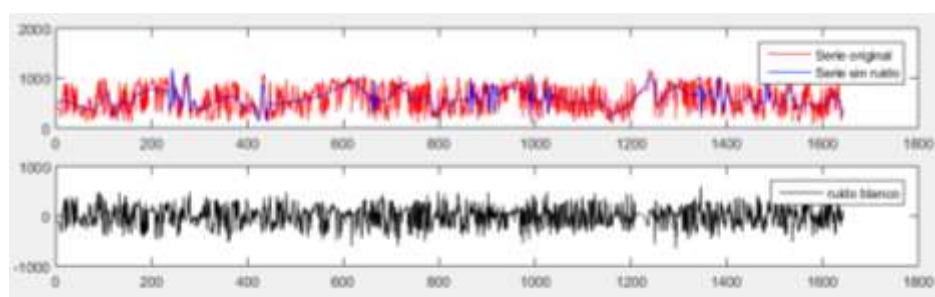


Figura 23. Análisis de radiación solar mediante la transformada de wavelet – Miranda

Para el caso de la variable de radiación solar de la estación El Carmen se obtuvo los resultados que se observan en la *Figura 24*, color rojo la serie temporal original de la variable

radiación solar, en color azul se identificó la serie sin ruido, es decir, la serie temporal resultante de la aplicación de la transformada de wavelet y en color negro se puede observar la información considerada ruido. El porcentaje de energía retenida es de 88,07%.

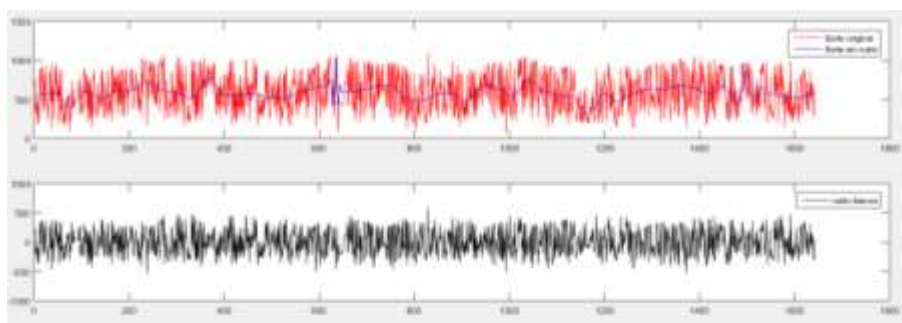


Figura 24. Análisis de radiación solar mediante la transformada de wavelet – El Carmen

En la *Figura 25* se puede observar en color rojo la serie temporal original de la variable radiación solar de la estación Chumillos, en color azul se identificó la serie sin ruido, es decir, la serie temporal resultante de la aplicación de la transformada de wavelet y en color negro se puede observar la información considerada ruido. El porcentaje de energía retenida es de 88,85%.

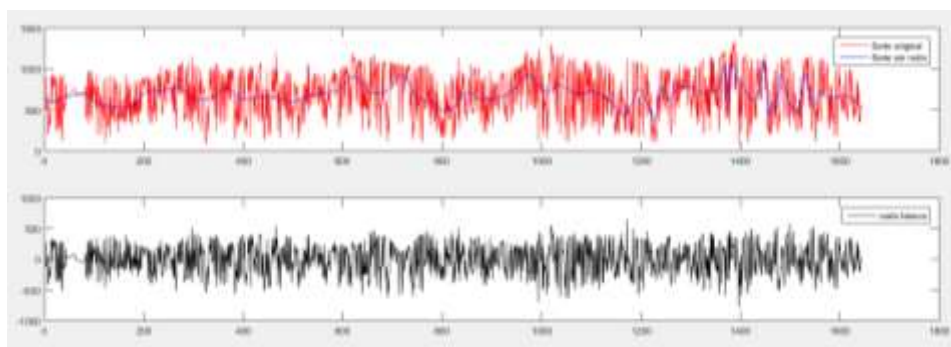


Figura 25. Análisis de radiación solar mediante la transformada de wavelet – Chumillos

Para el caso de la variable de radiación solar de la estación Puntas se obtuvo los resultados que se observan en la *Figura 26*, color rojo la serie temporal original de la variable radiación solar, en color azul se identificó la serie sin ruido, es decir, la serie temporal resultante de la aplicación de la transformada de wavelet y en color negro se puede observar la información considerada ruido. El porcentaje de energía retenida es de 90,42%.

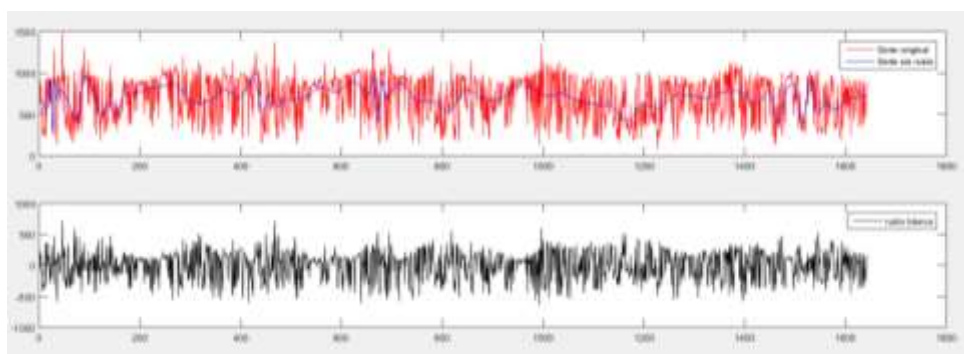


Figura 26. Análisis de radiación solar mediante la transformada de wavelet – Puntas

En la *Figura 27* se puede observar en color rojo la serie temporal original de la variable radiación solar de la estación Itulcachi, en color azul se identificó la serie sin ruido, es decir, la serie temporal resultante de la aplicación de la transformada de wavelet y en color negro se puede observar la información considerada ruido. El porcentaje de energía retenida es de 90,01%.

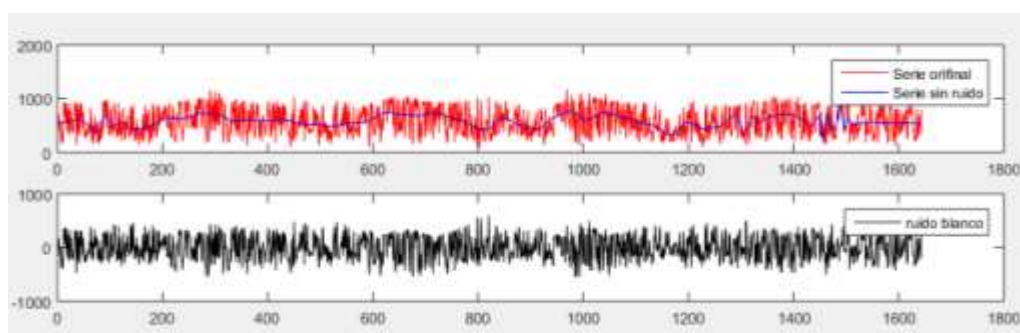


Figura 27. Análisis de radiación solar mediante la transformada de wavelet – Itulcachi

En el estudio realizado por Guarnizo (2008), se empleó wavelet para el tratamiento de señales electroencefalográficas debido a la susceptibilidad al ruido y se obtuvo una mejora del 1% en los porcentajes de acierto, por lo que el filtrado no proporcionó una gran diferencia. Por otro lado, en los resultados obtenidos por Pérez et al. (2012) indica que existe mejoría en los resultados en su estudio de modelos términos una vez que ha sido aplicada la transformada de wavelet. Según Pratap Yadav & Laxmidhar (2014) con la aplicación de la transformada de wavelet db4 nivel 3 proporcionó una disminución en los errores para la predicción de radiación solar. El criterio de los autores fue utilizado para el presente trabajo.

El resultado que se obtuvo de la aplicación de wavelet db4 nivel 3 indica en los 5 casos un porcentaje de energía retenida moderado. El nivel de descomposición fue 3, el cual es recomendado por Daubechies (1992) creadora de las wavelet db. El nivel 3 es bajo, debido a que busca garantizar que no exista eliminación de parte importante de la información. Sin embargo, indica en su investigación que debe considerar un nivel de acuerdo con la naturaleza de la señal a estudiar.

No obstante, como se muestra en las *Figuras 23, 24, 25, 26 y 27* se puede observar un suavizado excesivo en las series, a pesar que en la programación se especificó un umbral duro el cual presenta los datos crudos y sin un mayor suavizado. La dinámica de los datos de radiación y meteorológicos presenta fluctuaciones fuertes (Pratap Yadav & Laxmidhar, 2014) e incluso por este motivo se analiza su naturaleza caótica (Ayala, 2017), por lo que en este caso es posible que el filtro wavelet esté eliminando información importante para abstraer el comportamiento de las variables de radiación solar.

4.3 Redes neuronales artificiales

4.3.1 Arquitectura de las redes neuronales artificiales

Una vez que se especificaron los campos requeridos se obtiene como resultado la arquitectura que se indica en la *Figura 28*, donde se puede observar la entrada que corresponde a las 7 variables de entrada (tiempo, dirección del viento, humedad del aire, humedad del suelo, precipitación, temperatura y velocidad del viento), la capa oculta con 15 neuronas y con su función de transferencia que corresponde a sigmoidea tangente hiperbólica y la capa de salida con 1 neurona y su función de transferencia es lineal.

W corresponde a los pesos sinápticos y da mayor valor a ciertas conexiones, los cuales se modifican a lo largo del entrenamiento para minimizar el error con respecto al valor real (Muñoz & Seijas, 2015). El “bias” o sesgo que corresponde a un umbral en la cual se lleva a cabo a través la retropropagación (1000 épocas) y usa el gradiente descendente (Muñoz & Seijas, 2015). Es importante recalcar que el número de variables de entrada corresponde al número de neuronas en la capa de entrada, al igual que en el caso de la capa de salida.

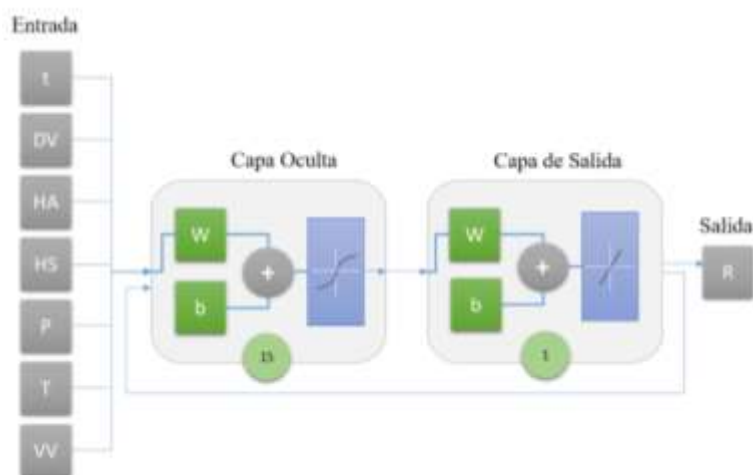


Figura 28. Arquitectura de la red neuronal artificial

Fuente: Elaboración propia

Para la arquitectura se aplicó el criterio del Instituto Asimove (2016) con respecto a redes Feed-forward en la cual se encuentra compuesta por la capa de entrada, una capa oculta y capa de salida. En el presente estudio se aplicó 15 neuronas en la capa oculta, debido a que después de varias pruebas con un número mayor de neuronas el resultado convergía a partir de 15, al igual que el estudio de Muñoz & Seijas (2015).

4.3.2 Resultados de entrenamiento de las redes neuronales

Estación Miranda

Los resultados del entrenamiento de las redes neuronales para la estación Miranda corresponden a los valores de R del ajuste, entrenamiento, validación, prueba y total. Los productos corresponden a las horas 10H00, 12H00, 14H00 y 16H00 respectivamente, como se indica en la *Figura 29*. El mayor valor de correlación en función del total corresponde a la red de las 10H00 con 0,87604, mientras que el valor más bajo corresponde a la red de las 16H00 con 0,80832. Los valores poseen una correlación intensa (Lind, et al., 2004)

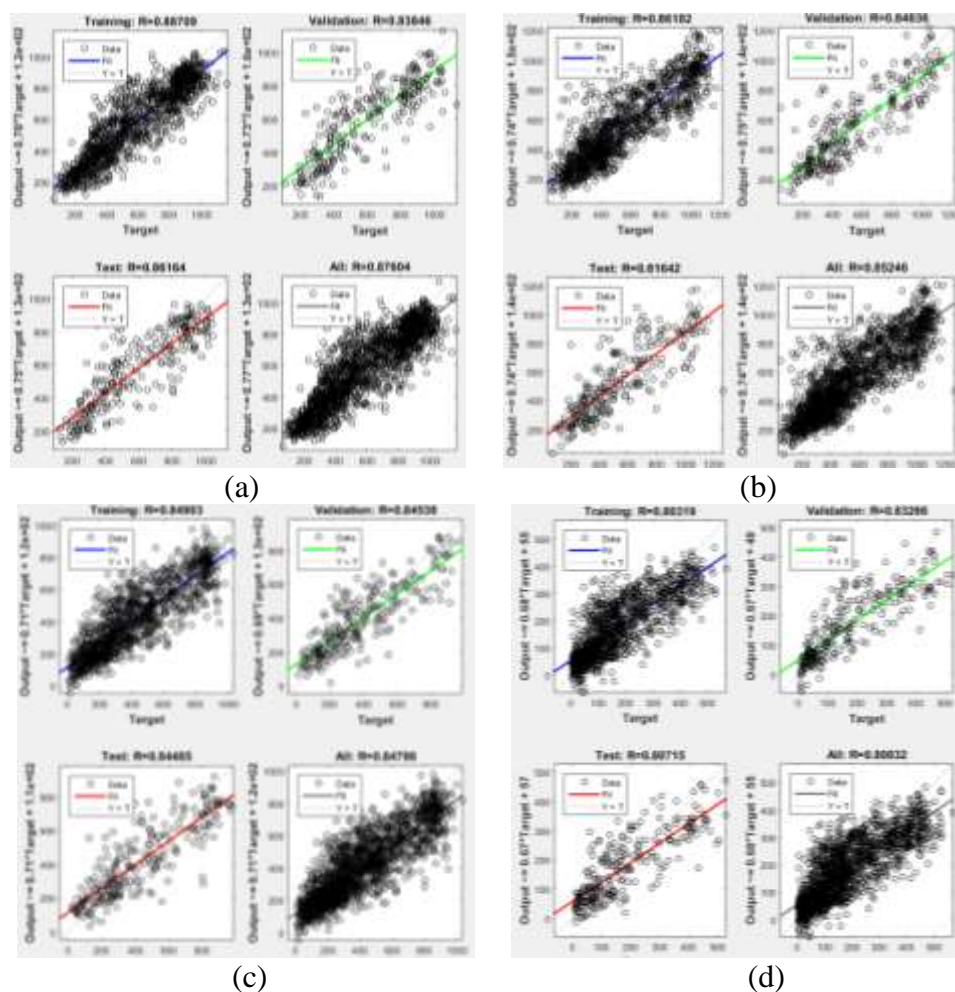


Figura 29. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Miranda a) 10H00, b) 12H00, c) 14H00 y d) 16H00

Estación El Carmen

Los resultados del entrenamiento de las redes neuronales para la estación El Carmen corresponden a los valores de R del ajuste, entrenamiento, validación, prueba y total. Los productos corresponden a las horas 10H00, 12H00, 14H00 y 16H00 respectivamente, como se indica en la *Figura 30*.

El mayor valor de correlación en función del total corresponde a la red de las 16H00 con 0,80385, mientras que el valor más bajo corresponde a la red de las 10H00 con 0,74257. Los valores poseen una correlación intensa (Lind, et al., 2004)

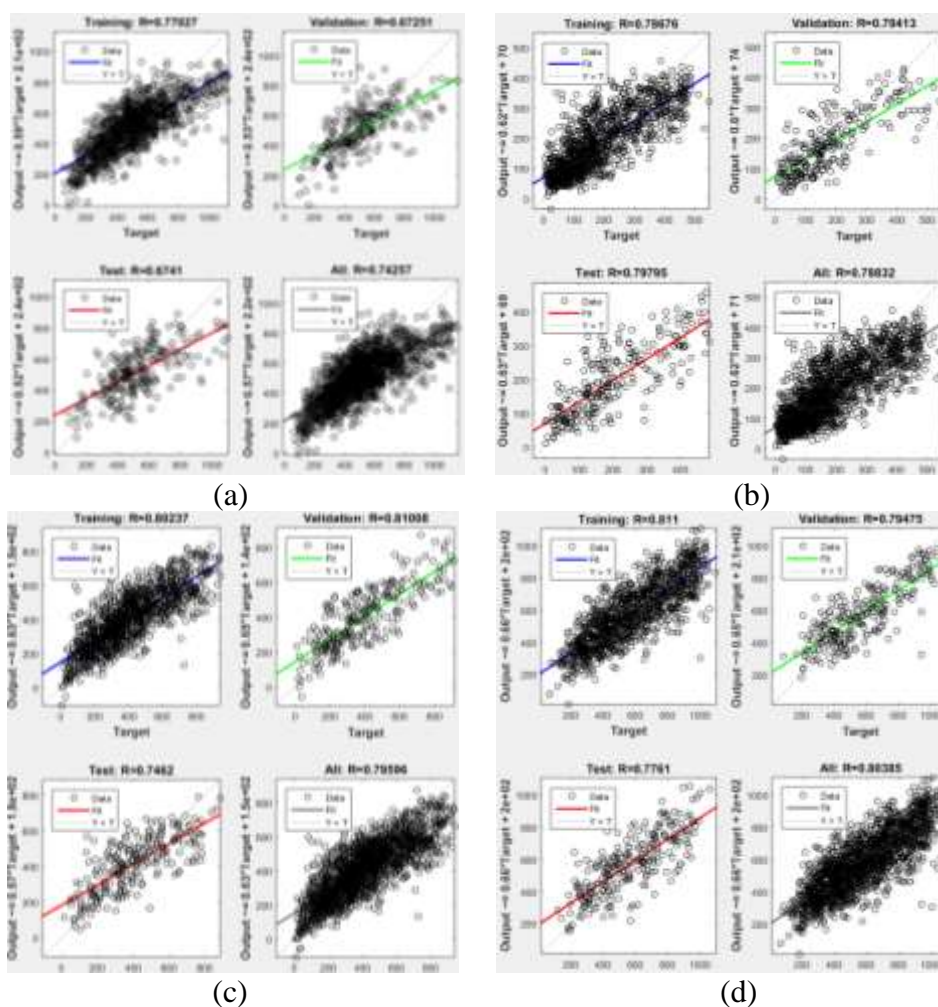


Figura 30. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación El Carmen a) 10H00, b)12H00, c)14H00 y d) 16H00

Estación Chumillos

Los resultados del entrenamiento de las redes neuronales para la estación Chumillos corresponden a los valores de R del ajuste, entrenamiento, validación, prueba y total. Los productos corresponden a las horas 10H00, 12H00, 14H00 y 16H00 respectivamente, como se indica en la *Figura 31*.

El mayor valor de correlación en función del total corresponde a la red de las 14H00 con 0,89096, mientras que el valor más bajo corresponde a la red de las 12H00 con 0,81856. Los valores poseen una correlación intensa (Lind, et al., 2004)

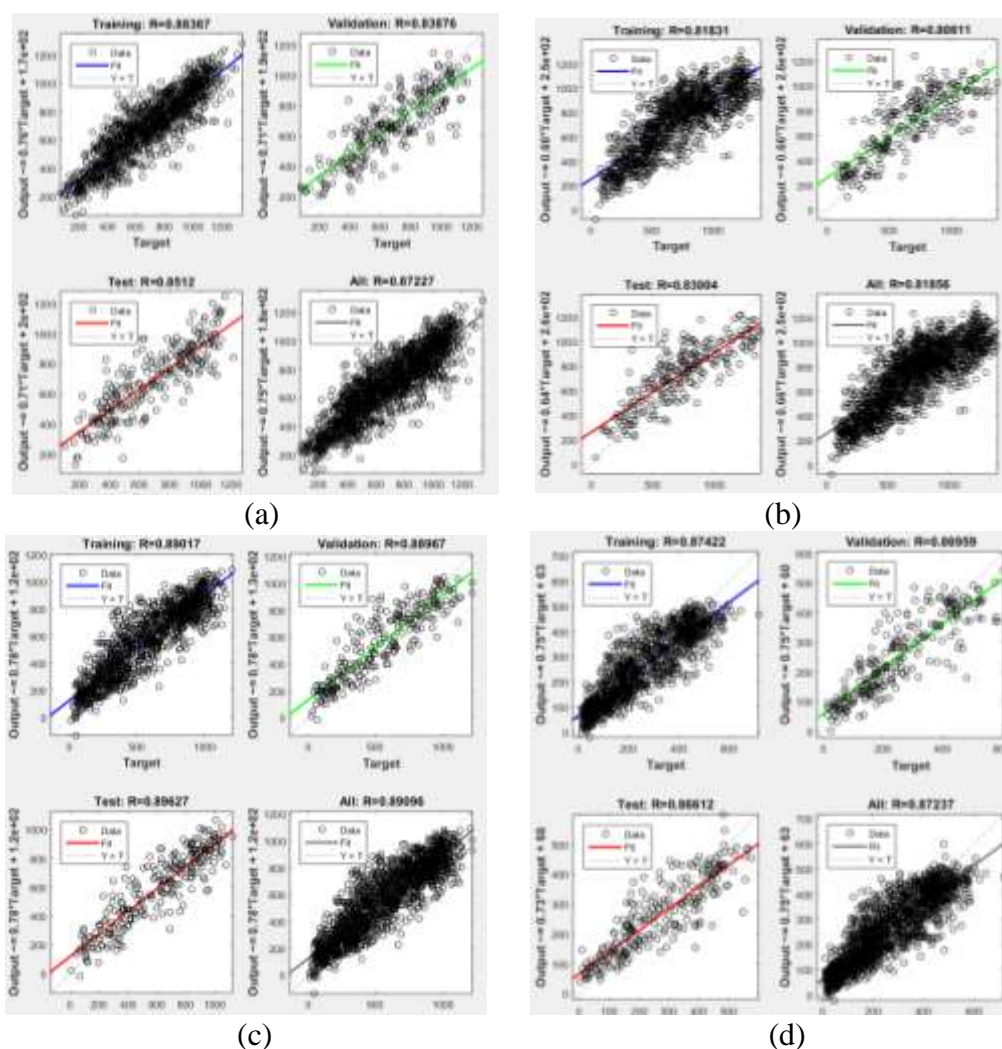


Figura 31. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Chumillos a) 10H00, b) 12H00, c) 14H00 y d) 16H00

Estación Puntas

Los resultados del entrenamiento de las redes neuronales para la estación Puntas corresponden a los valores de R del ajuste, entrenamiento, validación, prueba y total. Los productos corresponden a las horas 10H00, 12H00, 14H00 y 16H00 respectivamente, como se indica en la *Figura 32*.

El mayor valor de correlación en función del total corresponde a la red de las 10H00 con 0,83749, mientras que el valor más bajo corresponde a la red de las 16H00 con 0,70316. Los valores poseen una correlación intensa (Lind, et al., 2004)

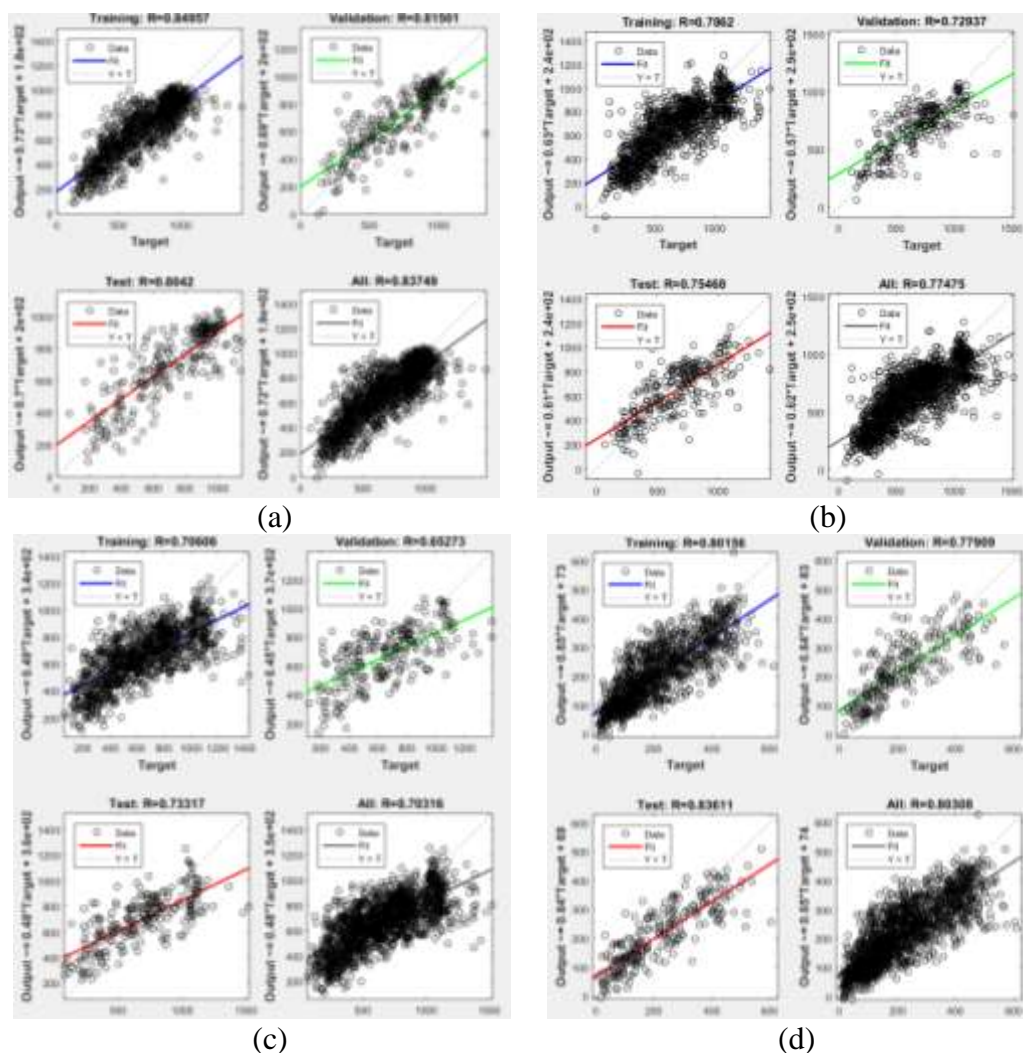


Figura 32. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Puntas a) 10H00, b) 12H00, c) 14H00 y d) 16H00

Estación Itulcachi

Los resultados del entrenamiento de las redes neuronales para la estación Itulcachi corresponden a los valores de R del ajuste, entrenamiento, validación, prueba y total. Los productos corresponden a las horas 10H00, 12H00, 14H00 y 16H00 respectivamente, como se indica en la *Figura 33*.

El mayor valor de correlación en función del total corresponde a la red de las 10H00 con 0,90568, mientras que el valor más bajo corresponde a la red de las 16H00 con 0,82646. Los valores poseen una correlación intensa (Lind, et al., 2004)

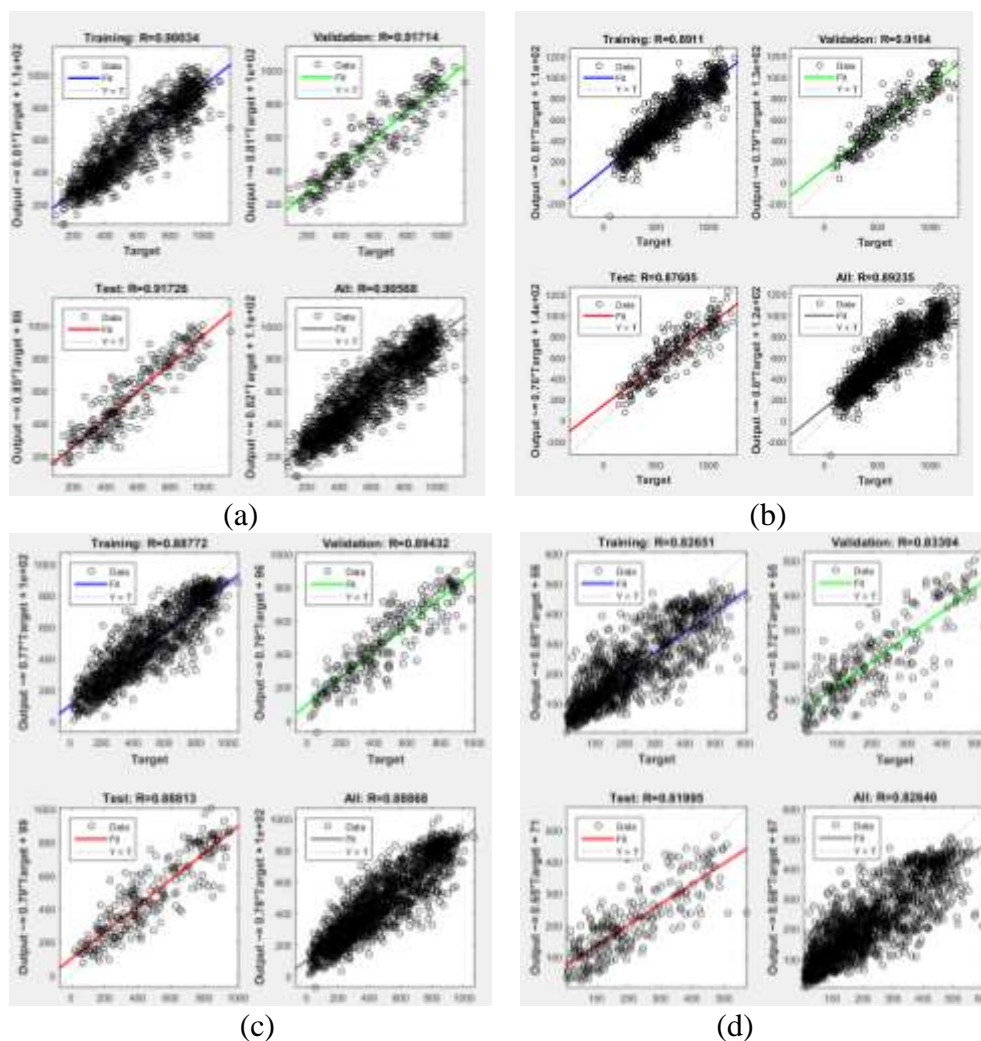


Figura 33. Correlación de ajuste del entrenamiento, validación y prueba de las redes estación Itulcachi a) 10H00, b)12H00, c)14H00 y d) 16H00

4.3.3 Resultados de la predicción

El resumen de la cantidad de datos predichos por estación y por fecha se indica en la Tabla 10, dando un total de 2232 por las 5 estaciones en análisis, en las 4 horas de mayor interés.

Tabla 10

Resumen de los datos predichos mediante redes neuronales

Estación	Inicio predicción	Final predicción	Horas	Cantidad de datos predichos
Miranda	07/02/2018	09/26/2018	10	(84x4) 336
El Carmen		08/26/2018	12	(56x4) 224
Chumillos		12/23/2018	14	(175x4) 700
Puntas		12/09/2018	16	(73x4) 292
Itulcachi		12/18/2018		(170x4) 680
Total				2232

Estación Miranda

Se obtuvieron los resultados de la predicción de las redes neuronales para la estación Miranda a las 10H00, 12H00, 14H00 y 16H00 respectivamente. Los valores de la predicción se indican en el *Anexo A*.

En la *Figura 34* se puede observar gráficamente el modelo de radiación solar de las 10H00, en la cual se muestra un buen ajuste en el periodo de 16 de julio del 2014 al 23 de julio del 2014. Mientras que en el intervalo de tiempo del mes de agosto existe un mayor error.

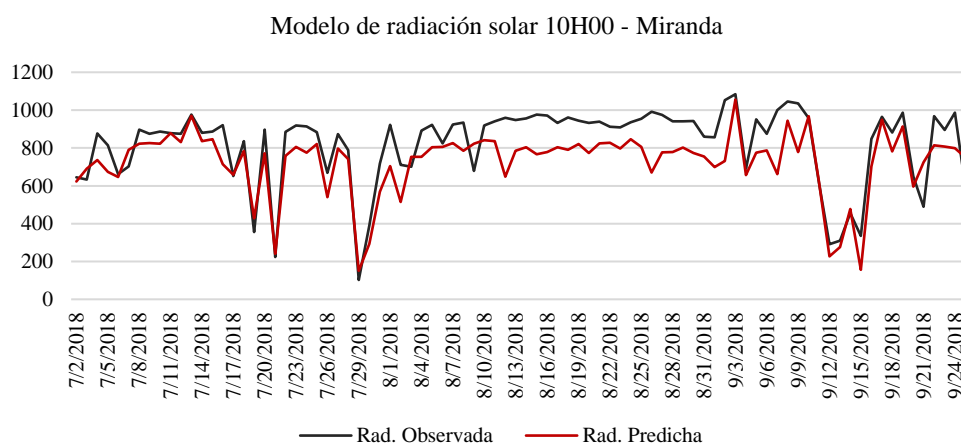


Figura 34. Modelo radiación solar 10H00 - Miranda

Los resultados de la predicción de las redes neuronales para la estación Miranda a las 12H00 se muestra en la *Figura 35*. El mejor ajuste se puede visualizar en el periodo del mes de septiembre del 2018, mientras que el mayor existe en el intervalo de tiempo de 30 de julio al 5 de agosto del 2018.

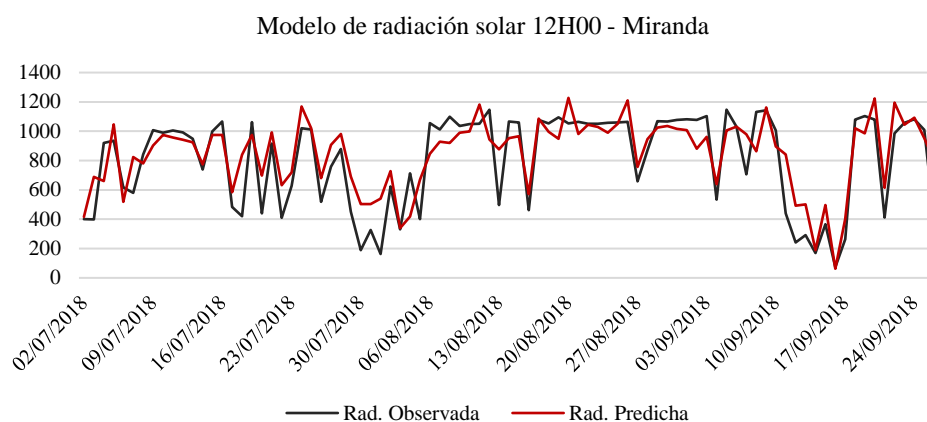


Figura 35. Modelo de radiación solar 12H00 - Miranda

Los resultados de la predicción de las redes neuronales para la estación Miranda a las 14H00 se muestra en la *Figura 36*, en donde se puede observar que el mejor ajuste se presenta en el intervalo del 30 de julio del 2018 al 12 de agosto del 2018. Sin embargo, el mayor error en el periodo del 16 de julio del 2018 al 23 de julio del 2018.

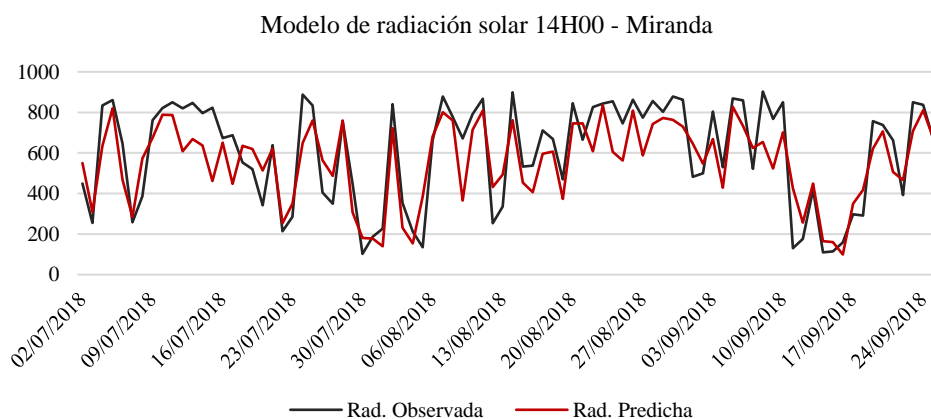


Figura 36. Modelo de radiación solar 14H00 - Miranda

Los resultados de la predicción de las redes neuronales para la estación Miranda a las 16H00 se muestra en la *Figura 37*. El mejor ajuste se puede visualizar en el periodo del 2 de julio del 2018 al 8 de julio del 2018. El mayor error se puede observar en el intervalo del mes de 20 de agosto del 2018 hasta aproximadamente el 3 de septiembre del 2018.

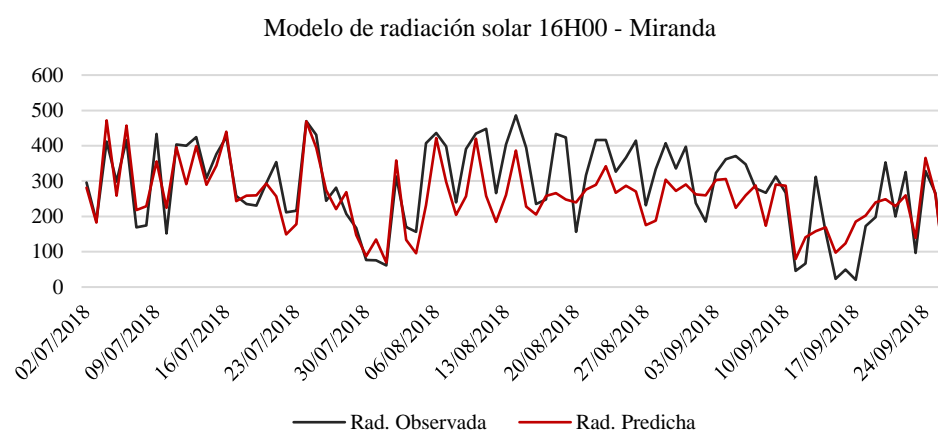


Figura 37. Modelo de radiación solar 16H00 – Miranda

Estación El Carmen

Se obtuvieron los resultados de la predicción de las redes neuronales para la estación El Carmen a las 10H00, 12H00, 14H00 y 16H00 respectivamente. Los valores de la predicción se indican en el *Anexo B*.

Los resultados de la predicción de las redes neuronales para la estación El Carmen a las 10H00 se muestran en la *Figura 38*, en donde se puede observar que el mejor ajuste se presente en el intervalo de 2 de julio del 2018 al 8 de julio del 2018, mientras que el mayor error se observa en el intervalo de 18 de julio del 2018 al 22 de julio del 2018 y en el periodo del 3 de agosto al 7 de agosto del 2018.

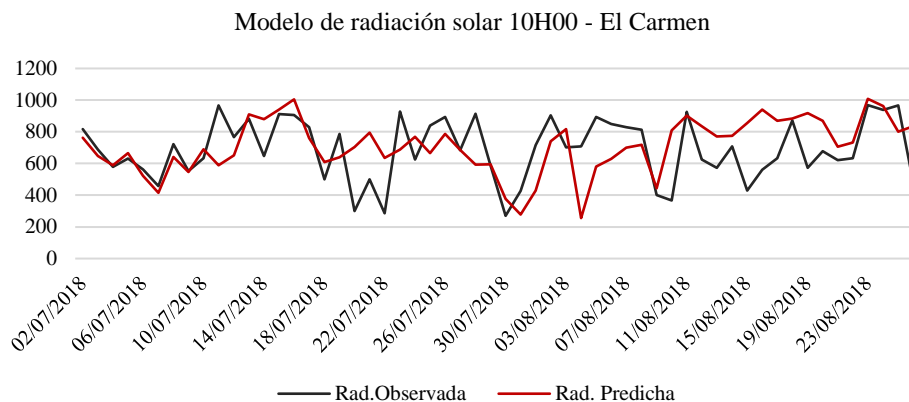


Figura 38. Modelo de radiación solar 10H00 - El Carmen

Los resultados de la predicción de las redes neuronales para la estación El Carmen a las 12H00 se muestra en la *Figura 39*. En el periodo del 2 de julio al 8 de julio presenta el mejor ajuste, mientras que en el intervalo de 23 de julio al 29 de julio del 2018 indica mayor error en las predicciones, al igual que en el tramo del 3 de agosto al 7 de agosto.

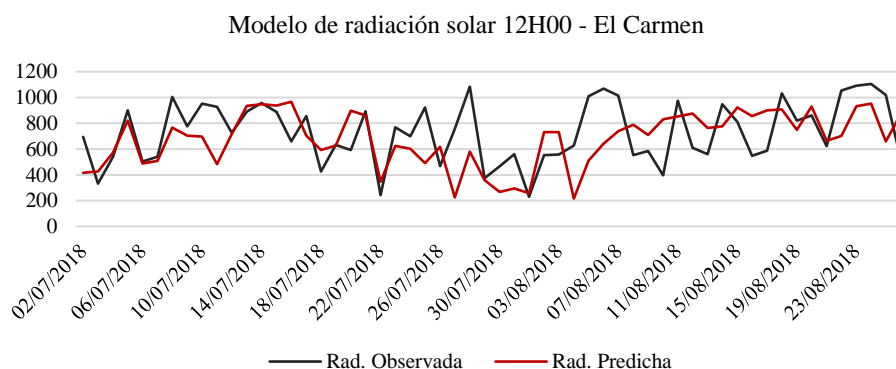


Figura 39. Modelo de radiación solar 12H00 - El Carmen

Los resultados de la predicción de las redes neuronales para la estación El Carmen a las 14H00 se muestra en la *Figura 40*. Se puede identificar que el menor error en el periodo del 3 de agosto al 5 de agosto del 2018, mientras que el mayor error se presenta en el intervalo del 22 de julio 1 de agosto del 2018.

En este caso la predicción no muestra un buen ajuste en el análisis general como se puede observar y corroborar en criterio estadístico que se muestra más adelante la sección Validación de resultados.

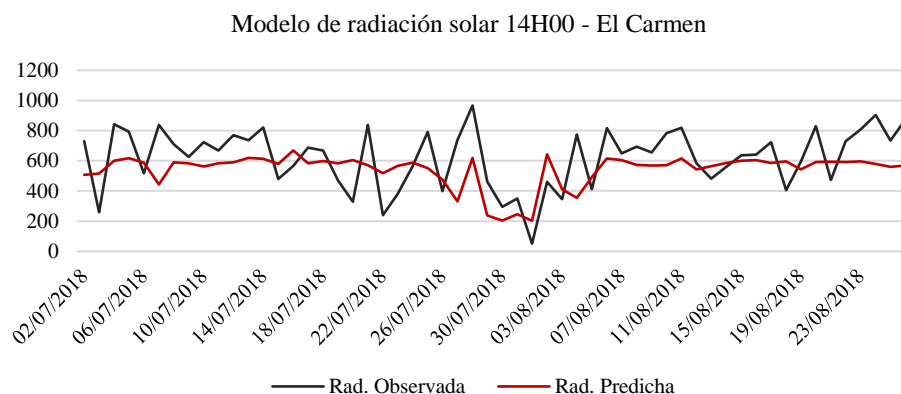


Figura 40. Modelo de radiación solar 14H00 - El Carmen

Los resultados de la predicción de las redes neuronales para la estación El Carmen a las 16H00 se muestra en la *Figura 41*. El mejor ajuste se puede visualizar en el intervalo del 1 de agosto al 6 de agosto del 2018, mientras que en el periodo del 26 de julio al 31 de julio del 2018 indica mayor error.

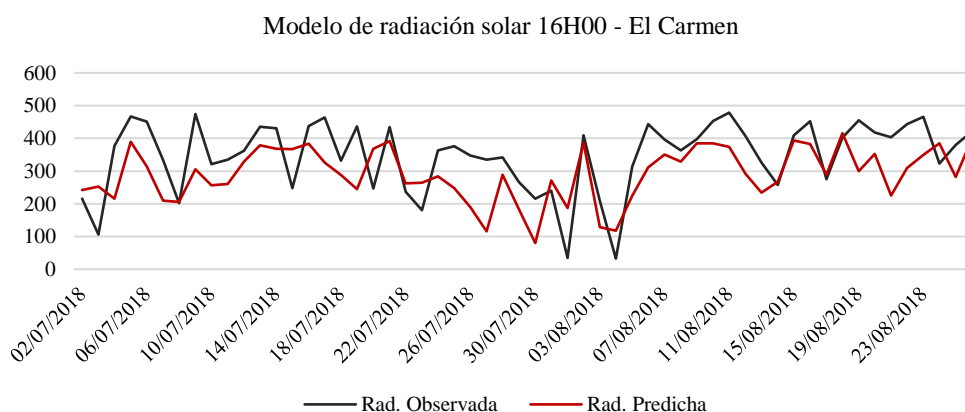


Figura 41. Modelo de radiación solar 16H00 - El Carmen

Estación Chumillos

Se obtuvieron los resultados de la predicción de las redes neuronales para la estación Chumillos a las 10H00, 12H00, 14H00 y 16H00 respectivamente. Los valores de la predicción se indican en el *Anexo C*.

Los resultados de la predicción de las redes neuronales para la estación Chumillos a las 10H00 se muestran en la *Figura 42*. En la cual se muestra en ajuste adecuado a nivel general, sin embargo en el periodo del 2 de septiembre al 27 de septiembre existe la mayor presencia de errores.

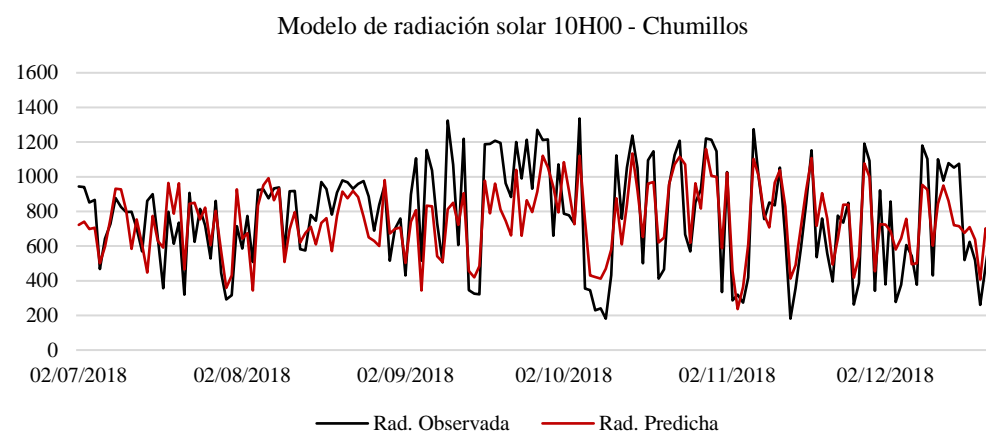


Figura 42. Modelo de radiación solar 10H00 - Chumillos

Los resultados de la predicción de las redes neuronales para la estación Chumillos a las 12H00 se muestran en la *Figura 43*. El periodo del 2 de julio al 31 de julio del 2018 al igual que el intervalo del 5 de agosto al 28 de agosto presentan los mayores errores del modelo.

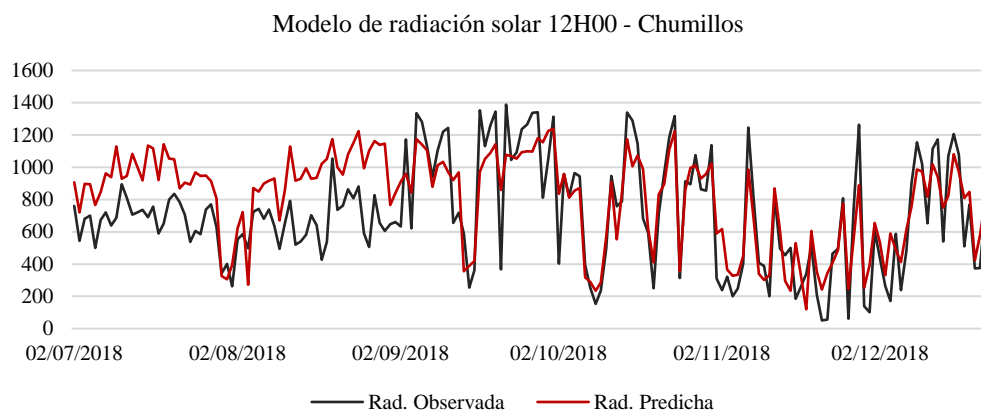


Figura 43. Modelo de radiación solar 12H00 - Chumillos

Los resultados de la predicción de las redes neuronales para la estación Chumillos a las 14H00 se muestran en la *Figura 44*. El modelo presenta un buen ajuste de forma general, sin embargo, en el periodo del 2 de septiembre al 10 de septiembre del 2018 presenta mayores errores.

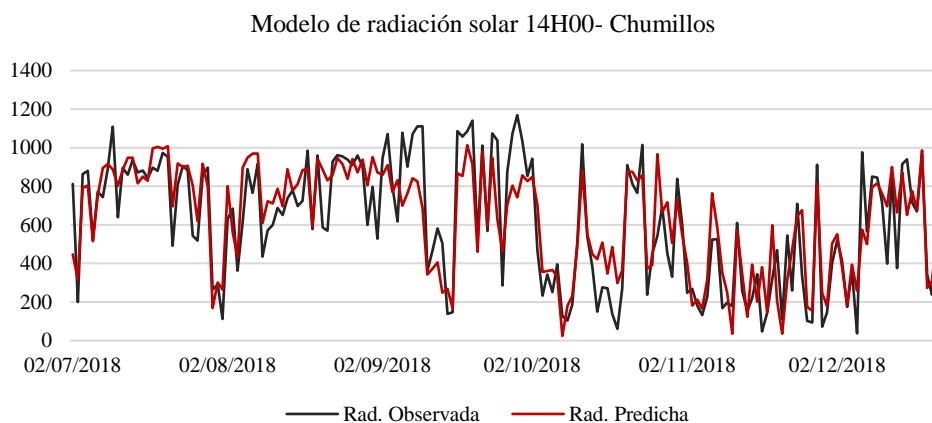


Figura 44. Modelo de radiación solar 14H00 - Chumillos

Los resultados de la predicción de las redes neuronales para la estación Chumillos a las 16H00 se muestran en la *Figura 45*. El intervalo que presenta mayores errores corresponde al intervalo de 25 de julio al 2 de agosto del 2018 como se puede visualizar.

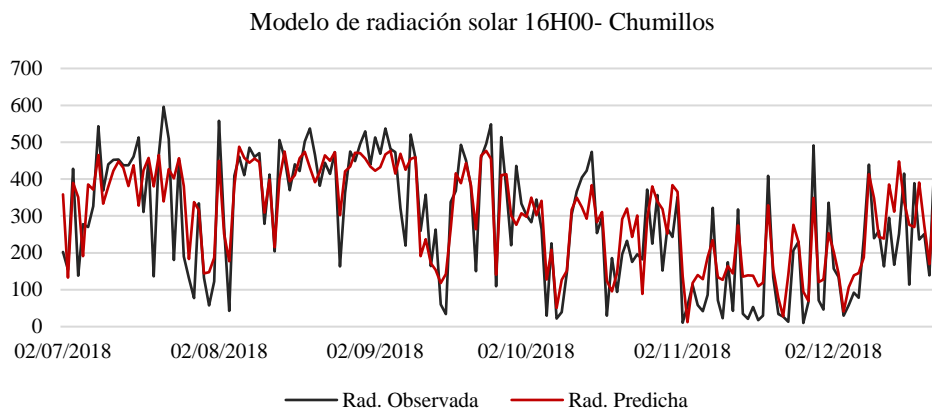


Figura 45. Modelo de radiación solar 16H00 – Chumillos

Estación Puntas

Se obtuvieron los resultados de la predicción de las redes neuronales para la estación Puntas a las 10H00, 12H00, 14H00 y 16H00 respectivamente. Los valores de la predicción se indican en el *Anexo D*.

Los resultados de la predicción de las redes neuronales para la estación Puntas a las 10H00 se muestran en la *Figura 46*. El mayor error se encuentra en el intervalo del 7 de septiembre del 2018 al 9 de septiembre del 2018. El mejor ajuste se puede visualizar en el periodo del 27 de agosto del 2018 al 1 de septiembre del 2018.

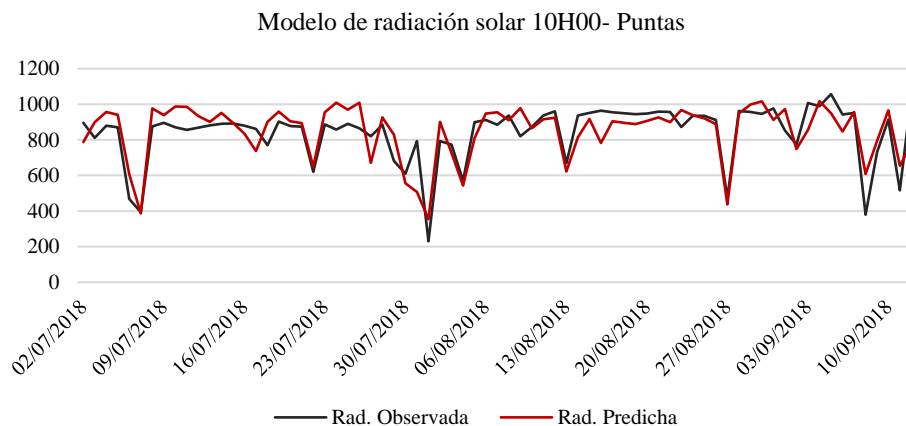


Figura 46. Modelo de radiación solar 10H00- Puntas

Los resultados de la predicción de las redes neuronales para la estación Puntas a las 12H00 se muestran en la *Figura 47*. El mayor error se encuentra en el intervalo del 7 de septiembre del 2018 al 15 de septiembre del 2018, al igual que en el transcurso del 20 de agosto al 25 de agosto del 2018. El mejor ajuste se puede visualizar en el periodo del 27 de julio al 2 de agosto del 2018.

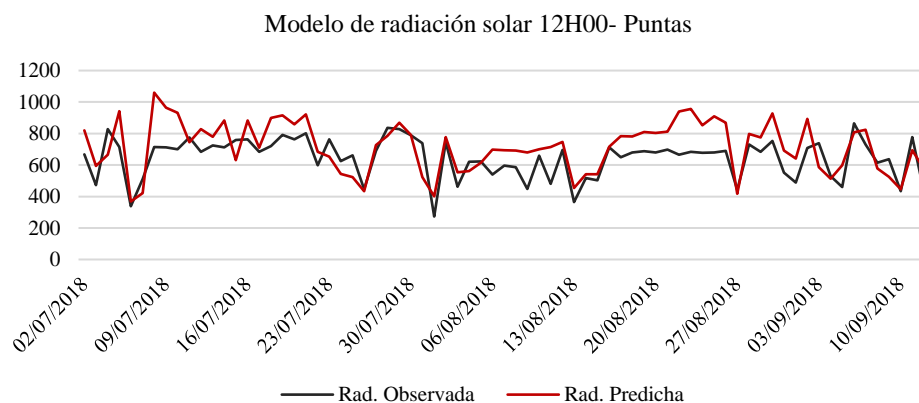


Figura 47. Modelo de radiación solar 12H00- Puntas

Los resultados de la predicción de las redes neuronales para la estación Puntas a las 14H00 se muestran en la *Figura 48*. Los intervalos que presenta mayor error corresponden al del 2 de julio al 7 de julio del 2018 y el del 2 de agosto al 6 de agosto del 2018. El mejor ajuste se muestra en el transcurso del 8 de agosto al 13 de agosto del 2018.

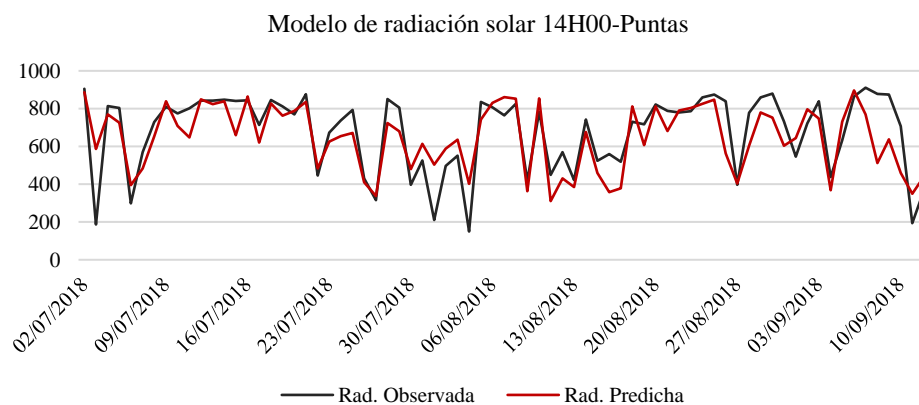


Figura 48. Modelo de radiación solar 14H00- Puntas

Los resultados de la predicción de las redes neuronales para la estación Puntas a las 16H00 se muestran en la *Figura 49*. El mayor error se encuentra en el intervalo del 9 de julio del 2018 al 16 de julio del 2018, al igual que en el transcurso del 30 de julio al 2 de agosto del 2018. El mejor ajuste se puede visualizar en el periodo del 13 de agosto al 20 de agosto del 2018.

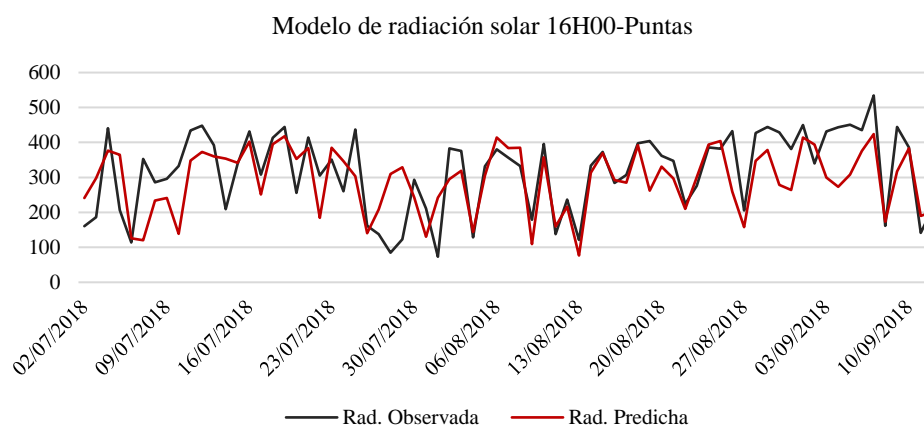


Figura 49. Modelo de radiación solar 16H00- Puntas

Estación Itulcachi

Se obtuvieron los resultados de la predicción de las redes neuronales para la estación Itulcachi a las 10H00, 12H00, 14H00 y 16H00 respectivamente. Los valores de la predicción se indican en el *Anexo E*.

Los resultados de la predicción de las redes neuronales para la estación Itulcachi a las 10H00 se muestran en la *Figura 50*. Los intervalos que presenta mayor error corresponden al del 20 de agosto al 2 de septiembre del 2018 y el del 4 de noviembre al 28 de noviembre del 2018.

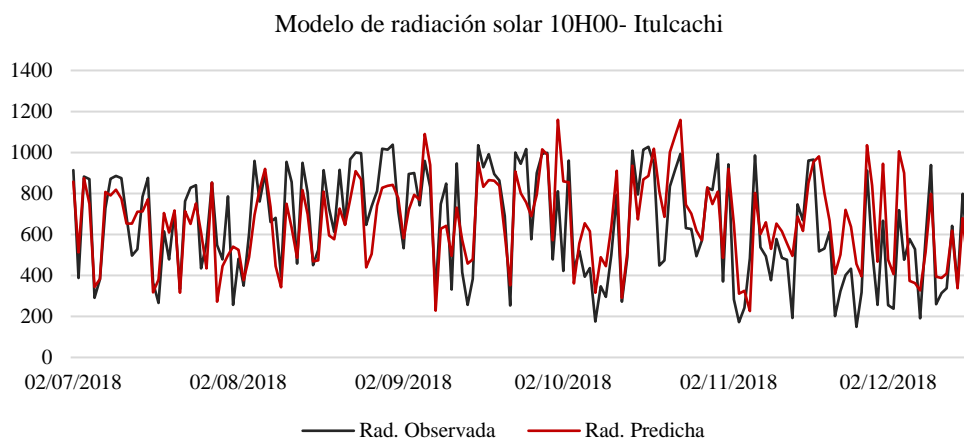


Figura 50. Modelo de radiación solar 10H00- Itulcachi

Los resultados de la predicción de las redes neuronales para la estación Itulcachi a las 12H00 se muestran en la *Figura 51*. El modelo de radiación solar a las 12H00 presenta un buen ajuste, sin embargo el intervalo que presenta mayores errores corresponde al del 10 de septiembre al 28 de septiembre del 2018, al igual que el periodo del 15 de noviembre al 25 de noviembre del 2018.

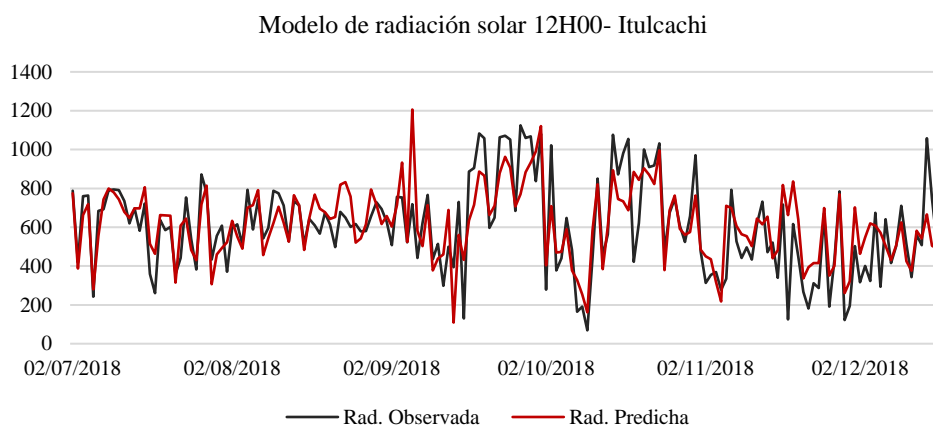


Figura 51. Modelo de radiación solar 12H00- Itulcachi

Los resultados de la predicción de las redes neuronales para la estación Itulcachi a las 14H00 se muestran en la *Figura 52*. El modelo de radiación solar a las 14H00 presenta un buen ajuste de forma general. El intervalo que presenta mayores errores corresponde al del 29 de agosto al 2 de septiembre del 2018, al igual que el periodo del 1 de octubre al 6 de octubre del 2018.

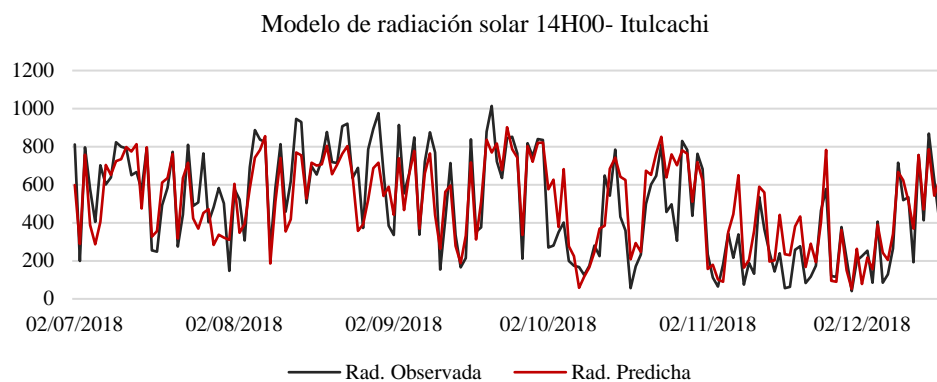


Figura 52. Modelo de radiación solar 14H00- Itulcachi

Los resultados de la predicción de las redes neuronales para la estación Itulcachi a las 16H00 se muestran en la *Figura 53*. Los intervalos que presenta mayor error corresponden al del 15 de julio 20 de julio del 2018 y el del 4 de noviembre al 20 de noviembre del 2018.

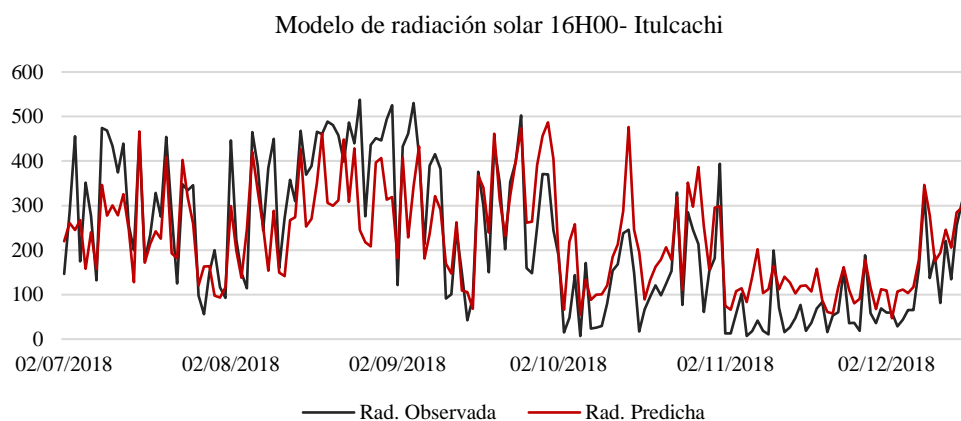


Figura 53. Modelo de radiación solar 16H00- Itulcachi

4.3.4 Validación de la predicción

El resumen de los valores de los coeficientes de determinación (R^2) se presenta en la Tabla 11. El valor más alto corresponde a 0,75 de la estación Miranda en horario de 10H00, mientras que el valor más bajo pertenece la estación El Carmen en el horario de 10H00 con 0,12 de coeficiente de determinación.

El 70 % de las estimaciones corresponden a valores mayores a 0,5, lo que indica un ajuste adecuado, mientras que el 30% restante indican un ajuste aparentemente no fiable. No obstante, cabe recalcar que los valores bajos del indicador estadístico R^2 son pertenecientes a las estaciones El Carmen y Puntas.

Es posible que el número de datos que se utilizaron para el entrenamiento de las dos estaciones especificadas anteriormente, no sean suficientes para que la red pueda comprender la dinámica de la radiación solar; sin embargo para las estaciones restantes se obtuvieron resultados adecuados considerando la complejidad que conlleva la predicción de este tipo de variables.

Tabla 11
Coeficientes de determinación de las predicciones

Estación	Hora	Coef. de determinación (R^2)
<i>Miranda</i>	10	0,75
	12	0,74
	14	0,74
	16	0,64
<i>El Carmen</i>	10	0,12
	12	0,17
	14	0,21
	16	0,38
<i>Chumillos</i>	10	0,65
	12	0,61
	14	0,74
	16	0,74
<i>Puntas</i>	10	0,65
	12	0,48
	14	0,64
	16	0,41
<i>Itulcachi</i>	10	0,61
	12	0,60
	14	0,73
	16	0,64

Tabla 12

Valores máximos, mínimos, promedio y desviación estándar de las predicciones, valores reales y errores absolutos de radiación solar de las estaciones en las horas de mayor incidencia (W/m^2)

Estaciones	Hora	Máximo			Mínimo			Promedio			Desv. Estandar		
		Valor Predicción	Valor Real	Error abs.	Valor Predicción	Valor Real	Error abs.	Valor Predicción	Valor Real	Error abs.	Valor Predicción	Valor Real	Error abs.
Miranda	10	1055,53	1084,23	338,31	150,53	102,36	0,90	726,71	819,76	114,15	171,75	203,15	76,17
	12	1227,51	1146,33	418,26	62,31	72,92	0,36	848,23	805,79	131,12	240,25	314,89	104,87
	14	834,49	903,15	361,03	99,39	102,57	3,75	555,56	597,41	114,18	194,39	254,25	76,91
	16	471,79	485,84	189,74	24,81	20,52	1,10	253,23	283,73	62,54	92,39	121,70	49,06
El Carmen	10	959,49	968,27	451,27	267,36	269,79	0,23	607,33	686,78	113,62	188,69	190,85	98,39
	12	1066,74	1105,16	531,36	169,33	229,82	7,10	611,25	728,31	197,07	245,34	235,61	148,14
	14	669,04	966,96	420,88	202,34	50,64	21,96	540,79	619,58	163,89	108,73	197,36	100,79
	16	415,56	478,50	216,26	80,20	32,54	3,39	291,74	346,08	86,28	81,65	107,67	53,32
Chumillos	10	1160,49	1336,11	510,99	236,45	181,22	0,81	752,06	778,19	153,59	196,62	293,16	96,69
	12	1237,87	1388,39	596,11	119,61	50,96	0,13	808,12	713,47	190,71	286,03	332,11	129,14
	14	1056,19	1168,99	425,98	24,31	37,31	1,04	644,46	622,44	125,07	270,96	313,97	104,20
	16	497,41	596,07	259,57	12,18	9,92	0,78	307,75	291,85	68,03	129,32	166,61	54,55
Puntas	10	970,78	1056,83	377,55	199,04	229,99	1,50	739,29	840,73	129,52	155,63	161,13	82,86
	12	1036,36	864,655	428,87	309,61	272,89	9,39	744,95	643,49	155,25	148,91	128,04	98,36
	14	883,67	910,79	660,01	95,42	149,58	0,69	643,66	676,99	106,02	161,39	202,46	109,46
	16	413,68	534,42	251,54	132,47	73,14	1,01	270,35	314,55	95,99	71,25	113,75	65,61
Itulcachi	10	1159,29	1037,78	464,05	226,24	149,07	1,72	660,95	642,27	129,97	207,51	254,87	92,74
	12	1205,98	1124,57	536,77	108,83	68,78	0,24	616,22	604,33	111,80	183,32	233,19	96,27
	14	902,17	1014,13	396,68	51,44	40,89	0,78	505,61	498,17	107,81	226,49	265,71	85,09
	16	486,91	537,72	292,40	47,00	6,97	0,48	226,54	223,09	74,56	111,60	157,60	59,42

En la Tabla 12 se presentan los valores máximos, mínimos, promedios y desviaciones estándar de las predicciones, valores reales y errores absolutos de radiación solar de las estaciones en las horas de mayor incidencia (10H00, 12H00, 14H00 y 16H00). Como se puede observar existe gran oscilación entre los valores máximos y mínimos, a modo de ejemplo en la estación el Carmen a las 16H00 indica un valor máximo de $596,07 \text{ W/m}^2$ y un valor mínimo de $9,92 \text{ W/m}^2$, existe una diferencia de 586.15 W/m^2 en una hora específica. La desviación estándar que presentan los datos de la radiación nos muestra la gran dispersión que existe entre los mismos con respecto a la media. La dinámica de la radiación solar indica un comportamiento complejo al igual que otras variables meteorológicas, incluso indagando sobre su comportamiento caótico (Ayala, 2017).

CAPITULO V

CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

En el presente análisis las variables meteorológicas que presentaron correlación intensa con la radiación solar son la temperatura y la humedad del aire.

El relleno de datos con Holt-Winters analizó la media, tendencia y variación estacional de las series temporales de las variables meteorológicas, lo que permitió que el relleno de datos sea adecuado al comportamiento de las mismas.

La naturaleza de la radiación solar presenta valores con gran fluctuación, por lo que en este caso el filtro wavelet db4 nivel de discretización 3, no fue óptimo para el análisis debido a que podría considerar ruido a valores que son genuinos en las series temporales.

Existe una relación directamente proporcional entre los valores del coeficiente del ajuste del entrenamiento y el R^2 como indicador de las predicciones. Las redes neuronales que presentaron valores altos de coeficiente de ajuste del entrenamiento mostraron a su vez coeficientes de determinación elevados.

El coeficiente de correlación más alto fue de 0,75 de la estación Miranda en horario de 10H00, mientras que el valor más bajo pertenece la estación El Carmen en el horario de 10H00 con 0,12.

Las redes neuronales artificiales recurrentes obtuvieron el 70% de coeficientes de determinación altos en la validación de las predicciones, lo cual indica la calidad del modelo.

5.2 Recomendaciones

Por la gran cantidad de vacíos que presenta la información meteorológica en las series temporales de las zonas andinas, se recomienda utilizar métodos adecuados para el relleno de datos, en los cuales se analicen los elementos principales de una serie temporal como es la media,

tendencia y estacionalidad como es el caso de la herramienta aplicada en el presente estudio, debido a que corresponde a una de las etapas más importantes para alcanzar el objetivo del trabajo.

Para el entrenamiento de las redes neuronales artificiales es importante contar con mayor cantidad de datos de observación para que la red pueda comprender de mejor manera el comportamiento de la variable en estudio y generar resultados más cercanos a los reales.

Se sugiere interactuar con una familia wavelet diferente o a su vez un diferente orden o nivel de descritización, para la eliminación de ruido de las series temporales de variables meteorológicas.

Se recomienda combinar la metodología empleada en el presente trabajo y los diferentes procedimientos para obtener el índice UV.

BIBLIOGRAFÍA

- Agencia Estatal de Meteorología. (2014). La radiación solar. España: Ministerio del Medio Ambiente y Medio Rural y Marino.
- Arancibia, C., & Best, R. (2010). *Energía del sol. Revista Ciencia*, 10-17.
- Arla, S., Tapia, M., Guasumba, J., Martínez, J., Asitimbat, J., & Tapia, E. (2017). Validación del Recurso Solar en el Ecuador para Aplicaciones de Media y Alta Temperatura. *INNOVA Research Journal*, 2(7), 34-45.
- Arla, S., Tapia, M., Guasumba, J., Martinez, J., Asitimbay, J., & Tapia, E. (2017). Validación del Recurso Solar en el Ecuador para Aplicaciones de Media y Alta temperatura. *INNOVA Research Journal*, 2(7), 34-45.
- Ayala, M. (2017). Análisis de la dinámica caótica para las series temporales de variables meteorológicas en la estación climatológica de Chone. Sangolquí, Ecuador: Universidad de las Fuerzas Armadas ESPE.
- Butler, K., & Bagriyanik, M. (2003). Characterization of transients in transformers using discrete wavelet transforms. *IEEE Transactions on Power Systems* 18, 648-656. doi:10.1109/TPWRS.2003.810979
- Campos, J., Torres, Y., & Paz, H. (2014). Recognition of alphabetic characters using functions Logsig and Tagsig in the Matlab environment.
- Campozano, L., Sánchez, E., Aviles, A., & Samaniego, E. (2015). Evaluation of infilling methods for times series of daily precipitation and temperature: The case of the Ecuadorian Andes. *Maskana*, 5(1), 99-115.
- Camus, A. (1954). *L'Été*. Gallimard.
- Carrera-Villacrés, D., Guevara-García, P., Maya-Carrillo, M. T.-B.-A., Narváez-Rivera, C., & Morocho-López, D. (2016). Relleno de series anuales de datos meteorológicos mediante métodos estadísticos en la zona costanera e interandina del Ecuador, y cálculo de precipitación media. *IDESIA*, 34(3), 81-90.
- CONELEC. (2008). *Corporación para la investigación energética*. Recuperado el 2019, de <http://energia.org.ec/cie/wp-content/uploads/2017/09/AtlasSolar.pdf>
- Daubechies, I. (1992). Ten Lectures on Wavelet. *Society for Industrial and Applied Mathematics*, 198, 254-257.

- Diario La Hora. (2018). Ecuador: Hoy, los índices de radiación solar más altos en la Sierra. Recuperado el 9 Abril de 2019, de <https://lahora.com.ec/noticia/1102142037/ecuador-hoy-los-indices-de-radiacion-solar-mas-altos-en-la-sierra>
- Doncel, D. (2018). Estimación de irradiancia solar aplicando el algoritmo HELIOSAT 1 con imágenes satelitales GOES en la región Cundiboyacense para el año 2011. Colombia: Universidad Distrital Francisco de Caldas.
- EPA. (2001). Guía SunWise para la exposición al sol. Recuperado el 2 de Junio de 2019, de https://www.epa.gov/sites/production/files/documents/sunuvu_spanish.pdf
- Flores, L., & Alcalá, J. (2010). Manual de Procedimientos Analíticos . México: Universidad Nacional Autónoma de México.
- FONAG. (2019). El Fondo. Recuperado el 1 de Junio de 2019, de http://www.fonag.org.ec/web/?page_id=77
- Gobierno Provincial de Pichincha. (2017). Distrito Metropolitano de Quito. Recuperado el 2 de Junio de 2019 , de <https://www.pichincha.gob.ec/cantones/distrito-metropolitano-de-quito>
- Gómez, M. (2008). Tópicos Avanzados: Redes Neuronales Artificiales Recurrentes. INAOE.
- Guarnizo, C. (2008). Análisis de reducción de ruido en señales EEG orientado al reconocimiento de patrones. *Revista Tecnológicas*, 67-80.
- Hagan, M., & Menhaj, M. (1999). Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989-993.
- INAMHI. (2014). Anuario Meteorológico 2011. Carlos Naranjo.
- INAMHI. (2017). Red de estaciones meteorológicas e hidrológicas. doi:<http://186.42.174.236/IndiceUV2/>
- INAMHI. (2 de Octubre de 2018). *Instituto Nacional de Meteorología en Hidrología*. Recuperado el 3 de Mayo de 2019, de <http://www.serviciometeorologico.gob.ec/>
- INAMHI. (2019). Misión y Visión INAMHI. Recuperado el 1 de Junio de 2019, de <http://www.serviciometeorologico.gob.ec/valores-mision-vision/>
- INEC. (2010). Censo de Población y Vivienda. Recuperado el 1 de Junio de 2019, de <https://www.ecuadorencifras.gob.ec/base-de-datos-censo-de-poblacion-y-vivienda/>
- Izar, J. (2019). Método de Winter. Universidad del Centro de México.
- Jaramillo, A. (2005). Clima Andino y Café en Colombia. Colombia: Centro Nacional de Investigaciones de Café CENICAFÉ.

- Justel, A., & Cayuela, L. (2015). *TIME SERIES*. Madrid: Universidad Politécnica de Madrid.
- Lind, D., Mason, W., Lind, R., Marchal, W., & Manson, R. (2004). *Estadística para administración y economía*. Alfaomega Grupo Editor.
- Marin, J. (2017). *Series Temporales*. Universidad Carlos III de Madrid. Obtenido de <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/EDescrip/tema7.pdf>
- MathWorks. (2019). Recuperado el Julio de 2019, de <https://es.mathworks.com/discovery/deep-learning.html>
- MathWorks. (2019). *MathWorks*. Recuperado el Julio de 2019, de <https://www.mathworks.com/help/wavelet/ref/wavedec.html>
- MathWorks. (2019). *MathWorks*. Recuperado el Julio de 2019, de <https://www.mathworks.com/discovery/machine-learning.html>
- MathWorks. (2019). *MathWorks*. Recuperado el Julio de 2019, de <https://es.mathworks.com/discovery/neural-network.html>
- MathWorks. (2019). *MathWorks*. Recuperado el Julio de 2019, de <https://nl.mathworks.com/help/deeplearning/ref/feedforwardnet.html;jsessionid=058cd9936e68fb08a637adf9476d>
- Mauricio, J. (2007). *Análisis de series temporales*. Madrid: Universidad Complutense de Madrid.
- Muñoz, E., & Seijas, C. (2015). Análisis comparativo de Ajuste en Entrenamiento de Redes Neuronales artificiales a partir de las Librerías Open NN y ALGLIB. *La Granja: Revista de Ciencias de la Vida*, 21, 49-60.
- Nieto, N., & Orozco, D. (2008). The use of the discrete Wavelet transform in the reconstruction of sinusoidal signals. *Scientia et Technica Año XIV*(38), 381-387.
- Novales, A. (2010). *Análisis de Regresión*. Universidad Complutense.
- OMS. (2003). *Índice uv solar mundial*. Guía práctica.
- Oropeza, C. (2007). *Modelado y Simulación de un Sistema de Detección de Intrusos Utilizando Redes Neuronales Recurrentes*. México: Universidad de las Américas Puebla.
- Palacios, I., & Cepeda, M. (2018). Elimination of White Noise from Time Series through the Wavelet Transform. *Modern Environmental Science and Engineerin*, 4(7), 681-685.
- Perez, M., Perez, A., & Perez, E. (2008). *Introducción a los sistemas de control y modelo matemático para sistemas lineales*. Universidad Nacional de San Juan.
- Pérez, R., Matos, E., Vásquez, C., & Fernández, S. (2012). Procesamiento de datos mediante Wavelet para la modelación térmica de transformadores de potencia. *Ingeniería Energética*, 33(1), 16-26.

- Pizarro, R., Ausensi, P., Aravena, D., Sanguensa, C., & Balocchi, F. (2009). Evaluación de métodos hidrológicos para la completación de datos faltantes de precipitación en estaciones de la región del Maule, Chile. *Aqua-LAC: revista del Programa Hidrológico Internacional para América Latina y el Caribe*, 1(2), 172-184.
- Prajakta, S. (2004). Time series Forecasting using Holt-Winters Exponential Smoothing. Kanwal Rekhi School of Information Technology.
- Pratap Yadav, A., & Laxmidhar, B. (2014). Solar Radiation forecasting using neural networks and Wavelet Transform. *IFAC*, 890 -896.
- R Core Team. (2019). R: A language and environment for statistical computing. Viena: R Foundation for Statistical Computing. Obtenido de <http://www.r-project.org/>
- Renza, D. (2010). Implementación y evaluación de un filtro digital para la transformada Wavelet compleja de doble árbol (DT-CWT) aplicada a fusión de imágenes bidimensionales en escala de gris. Bogotá: Universidad Nacional de Colombi.
- Rivas, W., & Mazón, B. (2018). *Redes neuronales artificiales aplicadas al reconocimiento de patrones*. Machala: UTMACH.
- Rodríguez, R., Capa, A., & Portela, A. (2004). Meteorología y Climatología. Fundación Española para la ciencia y tecnología.
- Rowe, A., & Abbott, P. (1995). Daubechies wavelets and Mathematica. 9, 635-648.
- Santamaría, F., Cortés, C., & Román, F. (2012). Uso de la Transformada de Ondeletas (Wavelet Transform) en la Reducción de Ruidos en las Señales de Campo Eléctrico producidas por Rayos. *Información Tecnológica* 23(1), 65-78. doi: 10.4067/ S0718-07642012000100008
- Secretaria del Ambiente. (2018). Índice Ultravioleta. Recuperado el Mayo de 2019, de <http://www.quitoambiente.gob.ec/ambiente/index.php/radiacion-ultravioleta-app>
- The Asimove institute. (2016). The neural network zoo. Recuperado el Julio de 2019, de <http://www.asimovinstitute.org/neural-network-zoo/>
- UNDP. (2019). *Programa de las Naciones Unidas para el Desarrollo*. Obtenido de <http://www.undp.org/content/undp/es/home/sustainable-development-goals/goal-7-affordable-and-clean-energy.html>
- Villavicencio, J. (2011). *Introducción a las series de tiempo*. Obtenido de http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D
- Yamasaki, L. (2008). Aplicaciones de la transformada wavelet: Eliminación de ruido. *Mural UV-Universidad de Valencia*, 1-7.

Yuleidys, C., Carrasco, R., Ochoa, I., & Moreno, E. (2013). Funciones de transferencia en el perceptrón multicapa: efecto de su combinación en entrenamiento local y distribuido. *RCIM [online]*, 5(2), 186-199. doi:1684-1859