

**UNIVERSIDAD NACIONAL AGRARIA**

**LA MOLINA**

**FACULTAD DE CIENCIAS**



**“PRONÓSTICO DE VISIBILIDAD HORIZONTAL Y TEMPERATURAS  
PARA EL AEROPUERTO DE TRUJILLO MEDIANTE EL USO  
DE REDES NEURONALES ARTIFICIALES  
DE RETROPROPAGACIÓN”**

Presentada por:

**BIBY ANN TENAUD JO**

Tesis para Optar el Título Profesional de:

**INGENIERA METEORÓLOGA**

Lima – Perú

**2023**

---

La UNALM es la titular de los derechos patrimoniales de la presente investigación  
(Art. 24. Reglamento de Propiedad Intelectual)

## Document Information

---

Analyzed document	TESIS_TENAUDBIBY_FINAL.pdf (D151530400)
Submitted	11/30/2022 2:56:00 PM
Submitted by	EUSEBIO IDELMO CISNEROS TARMEÑO
Submitter email	ecisneros@lamolina.edu.pe
Similarity	1%
Analysis address	ecisneros.unalm@analysis.arkund.com

## Sources included in the report

---

Universidad Nacional Agraria La Molina / TSP -FELICIANO CABEZAS G 14 OCT-LINEAS-DECCTE uñtimo del 17 octubre.docx

**SA**

Document TSP -FELICIANO CABEZAS G 14 OCT-LINEAS-DE-CCTE uñtimo del 17 octubre.docx (D146720289)

 4

Submitted by: ecisneros@lamolina.edu.pe

Receiver: ecisneros.unalm@analysis.arkund.com

**SA**

Proyecto de Tesis JMBS Final V F5.0.docx

Document Proyecto de Tesis JMBS Final V F5.0.docx (D50721755)

 1

**SA**

ENSAYO CARLOS CAÑAR.pdf

Document ENSAYO CARLOS CAÑAR.pdf (D16563295)

 3

**SA**

2283-12824-1-SM.pdf

Document 2283-12824-1-SM.pdf (D36002067)

 2

## Entire Document

---

1  
UNIVERSIDAD NACIONAL AGRARIA LA MOLINA FACULTAD DE CIENCIAS “PRONÓSTICO DE VISIBILIDAD HORIZONTAL Y TEMPERATURAS PARA EL AEROPUERTO DE TRUJILLO MEDIANTE EL USO DE REDES NEURONALES ARTIFICIALES DE RETROPROPAGACIÓN” Presentada por BIBY ANN TENAUD JO TESIS PARA OPTAR EL TÍTULO DE INGENIERA METEORÓLOGA Lima – Perú 2022

1 DEDICATORIA A mis padres, por todo el apoyo incondicional que me brindaron a lo largo de mis estudios y ser mi motivación cada día para seguir lográndome personal y profesionalmente.

2 AGRADECIMIENTO Muchas gracias al profesor Eusebio Cisneros por la asesoría brindada y su constante ánimo hasta la finalización de la presente tesis. Muchas gracias al Ing. Wilmer Pulache por haberme impulsado en realizar la presente tesis y su guía para el tratamiento de la información climatológica, y al Ing. Fernando Allasi por inculcarme el interés en el estudio del ámbito marino-costero. Mi más grande agradecimiento a la Corporación Peruana de Aeropuertos y Aviación Comercial, CORPAC, por su colaboración con respecto a la información meteorológica proporcionada. De igual forma, un agradecimiento especial a mis colegas

**UNIVERSIDAD NACIONAL AGRARIA  
LA MOLINA**

**FACULTAD DE CIENCIAS**

**“PRONÓSTICO DE VISIBILIDAD HORIZONTAL Y TEMPERATURAS  
PARA EL AEROPUERTO DE TRUJILLO MEDIANTE EL USO  
DE REDES NEURONALES ARTIFICIALES  
DE RETROPROPAGACIÓN”**

Presentada por:

**BIBY ANN TENAUD JO**

Tesis para Optar el Título Profesional de:

**INGENIERA METEORÓLOGA**

Sustentada y aprobada por el siguiente jurado:

---

Mg. Sc. Jerónimo García Villanueva  
PRESIDENTE

---

Ing. Franklin Delio Unsihuay Tovar  
MIEMBRO

---

Maest. Cs. Alessandri Canchoa Quispe  
MIEMBRO

---

Mg.Sc. Eusebio Idelmo Cisneros Tarmeño  
ASESOR

## **DEDICATORIA**

*A mis padres, por todo el apoyo incondicional que me brindaron a lo largo de mis estudios y ser mi motivación cada día para seguir lográndome personal y profesionalmente.*

## **AGRADECIMIENTOS**

*Muchas gracias al profesor Eusebio Cisneros por la asesoría brindada y su constante ánimo hasta la finalización de la presente tesis.*

*Muchas gracias al Ing. Wilmer Pulache por haberme impulsado en realizar la presente tesis y su guía para el tratamiento de la información climatológica, y al Ing. Fernando Allasi por inculcarme el interés en el estudio del ámbito marino-costero.*

*Mi más grande agradecimiento a la Corporación Peruana de Aeropuertos y Aviación Comercial, CORPAC, por su colaboración con respecto a la información meteorológica proporcionada. De igual forma, un agradecimiento especial a mis colegas del Equipo de Pronósticos y Climatología, por sus valiosos consejos.*

*Quisiera agradecer también a toda mi familia, que me motivaron constantemente a seguir avanzando con este trabajo y me brindaron bastante apoyo emocional.*

## ÍNDICE GENERAL

RESUMEN.....	viii
ABSTRACT .....	ix
I. INTRODUCCIÓN .....	1
II. REVISIÓN DE LA LITERATURA .....	4
2.1. Antecedentes de la investigación.....	4
2.2. Base teórica .....	5
2.2.1. Análisis exploratorio de los datos.....	5
2.2.2. Climatología .....	6
2.2.3. Pronóstico del tiempo.....	6
2.2.4. Temperatura Máxima y Mínima del Aire .....	7
2.2.5. Visibilidad Horizontal.....	7
2.2.6. Niebla .....	9
2.2.7. Redes neuronales artificiales .....	13
2.2.8. Redes neuronales artificiales de retropropagación .....	17
2.2.9. Evaluación de pronóstico .....	21
III. MATERIALES Y MÉTODOS.....	24
3.1. Materiales y equipos .....	24
3.1.1. Descripción de la zona de estudio.....	24
3.1.2. Información a procesar.....	24
3.1.3. Programas y equipos .....	25
3.2. Desarrollo metodológico .....	26
3.2.1. Análisis e identificación de las condiciones climatológicas y sistemas meteorológicos de influencia que predominan sobre el área de estudio.....	26
3.2.2. Identificación de las variables de entrada para la red neuronal del pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas.....	27

3.2.3. Determinación de los algoritmos apropiados para el entrenamiento y aprendizaje de la red neuronal en la arquitectura seleccionada para el pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas.....	27
3.2.4. Validación del modelo que optimice el pronóstico, con menor grado de error, de las variables meteorológicas Visibilidad Horizontal y Temperaturas Máximas y Mínimas .....	30
IV. RESULTADOS Y DISCUSIÓN .....	31
4.1. Análisis e identificación de las condiciones climatológicas y sistemas meteorológicos de influencia que predominan sobre el área de estudio. ....	31
4.2. Identificación de las variables de entrada para la red neuronal del pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas .....	40
4.3. Determinación de los algoritmos apropiados para el entrenamiento y aprendizaje de la red neuronal en la arquitectura seleccionada para el pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas.....	41
4.3.1. Temperatura Máxima.....	41
4.3.2. Temperatura Mínima.....	43
4.3.3. Visibilidad Horizontal.....	45
4.4. Validación del modelo que optimice el pronóstico, con menor grado de error, de las variables meteorológicas Visibilidad Horizontal y Temperaturas Máximas y Mínimas.....	47
4.4.1. Temperatura Máxima.....	47
4.4.2. Temperatura Mínima.....	48
4.4.3. Visibilidad Horizontal.....	48
V. CONCLUSIONES .....	50
VI. RECOMENDACIONES .....	52
VII. BIBLIOGRAFÍA .....	53
VIII. ANEXOS.....	57

## ÍNDICE DE TABLAS

Tabla 1: Estadísticas descriptivas para el Aeropuerto de Trujillo. ....	32
Tabla 2: Evaluación del pronóstico de la temperatura máxima obtenido por Redes Neuronales Artificiales de Retropropagación. ....	47
Tabla 3: Evaluación del pronóstico de temperatura mínima obtenido por Redes Neuronales Artificiales de Retropropagación. ....	48
Tabla 4: Evaluación del pronóstico de visibilidad horizontal obtenido por Redes Neuronales Artificiales de Retropropagación. ....	49



## ÍNDICE DE FIGURAS

Figura 1. Nieblas de evaporación: a) Esquema de niebla de evaporación por aire frío y b) Esquema de niebla de evaporación por lluvia.....	11
Figura 2. Nieblas por enfriamiento: a) Esquema de niebla de radiación, b) Esquema de niebla orográfica, y c) Esquema de niebla de advección.....	13
Figura 3. Entradas y salidas de una neurona Uj. ....	14
Figura 4. Modelo de arquitectura de una red backpropagation. ....	18
Figura 5. Mapa de ubicación del Aeropuerto “Cap. Carlos Martínez De Pinillos”.....	24
Figura 6. Diagrama de cajas para la temperatura mínima y máxima del Aeropuerto de Trujillo. ....	31
Figura 7. Diagrama de cajas para la visibilidad horizontal del aeropuerto de Trujillo.....	32
Figura 8. Resumen climatológico de la Temperatura del Aire del Aeropuerto de Trujillo.....	33
Figura 9. Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en verano. ....	33
Figura 10. Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en otoño.....	34
Figura 11. Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en invierno.....	34
Figura 12. Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en primavera. ....	35
Figura 13. Resumen climatológico del Viento del Aeropuerto de Trujillo.....	35
Figura 14. Resumen climatológico del Viento Máximo del Aeropuerto de Trujillo.....	36
Figura 15. Resumen climatológico de la presión atmosférica del Aeropuerto de Trujillo.....	36
Figura 16. Resumen climatológico de la precipitación y humedad relativa del Aeropuerto de Trujillo. ....	37

Figura 17. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en verano.....	37
Figura 18. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en otoño. ....	38
Figura 19. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en invierno. ....	38
Figura 20. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en primavera.....	39
Figura 21. Diseño de la red neuronal de cuatro neuronas ocultas para la Temperatura Máxima. ....	41
Figura 22. Diseño de la red neuronal de dos neuronas ocultas para la Temperatura Máxima. ....	42
Figura 23. Diseño de la red neuronal de ocho neuronas ocultas para la Temperatura Máxima. ....	42
Figura 24. Diseño de la red neuronal de cuatro neuronas ocultas para la Temperatura Mínima. ....	43
Figura 25. Diseño de la red neuronal de dos neuronas ocultas para la Temperatura Mínima. ....	44
Figura 26. Diseño de la red neuronal de ocho neuronas ocultas para la Temperatura Mínima. ....	44
Figura 27. Diseño de la red neuronal de cinco neuronas ocultas para la Visibilidad Horizontal.....	45
Figura 28. Diseño de la red neuronal de tres neuronas ocultas para la Visibilidad Horizontal.....	46
Figura 29. Diseño de la red neuronal de diez neuronas ocultas para la Visibilidad Horizontal.....	46
Figura 30. Temperatura Máxima del aeropuerto de Trujillo observada versus data pronosticada con Redes Neuronales Artificiales de Retropropagación.....	47

Figura 31. Temperatura Mnima del aeropuerto de Trujillo observada versus data pronosticada con Redes Neuronales Artificiales de Retropropagacin.....	48
Figura 32. Visibilidad Horizontal del aeropuerto de Trujillo observada versus data pronosticada con Redes Neuronales Artificiales de Retropropagacin.....	49

## ÍNDICE DE ANEXOS

Anexo 1. Ubicación geográfica de los índices utilizados en el análisis de los datos.....	58
Anexo 2. Variables globales: nombre, abreviatura y uso.....	59

## RESUMEN

Es casi imposible hablar de cualquier aspecto de la aviación, sin hacer referencia al entorno meteorológico en el cual una aeronave opera. La no previsión oportuna de los fenómenos meteorológicos adversos para la aviación puede producir retrasos o cancelaciones de los vuelos, generando grandes costos económicos para las compañías aéreas y los pasajeros. Las nuevas tecnologías computacionales que se van desarrollando, permiten una actualización de las técnicas de pronóstico del tiempo, teniendo a las redes neuronales artificiales como un nuevo paradigma de interés. En la presente investigación se realizó el pronóstico de las variables meteorológicas de visibilidad horizontal y temperaturas máximas y mínimas en el aeropuerto internacional de Trujillo utilizando Redes Neuronales Artificiales de Retropropagación (del inglés backpropagation) analizando e identificando las condiciones climatológicas y sistemas meteorológicos de influencia que predominan sobre el lugar. Teniendo como variables de entrada para la Red Neuronal datos diarios de la temperatura máxima y mínima de los años 2008 al 2017 y datos horarios de la visibilidad horizontal de los años 2013 al 2017 del aeropuerto de Trujillo, además de las variables explicativas como índices globales y otras variables meteorológicas locales. Las Redes Neuronales utilizadas para el análisis de la temperatura máxima y mínima del aire constan de ocho neuronas en la capa de entrada y en la capa oculta, y una capa de salida de una neurona; y para la visibilidad horizontal consta de diez neuronas en la capa de entrada y en la capa oculta, y una capa de salida de una neurona. Los resultados mostraron que, para las variables de temperaturas máximas y mínimas, los valores pronosticados y observados son muy cercanos entre sí con una tendencia similar; y para la variable de visibilidad horizontal, se perciben diferencias significativas, sin embargo, sí muestran patrones de tendencia similares.

**Palabras clave:** Redes Neuronales Artificiales Retropropagación, Backpropagation, pronóstico, visibilidad, temperatura, Trujillo, aeronáutica.

## ABSTRACT

It is almost impossible to talk about any aspect of aviation without reference to the meteorological environment in which an aircraft operates. Failure to forecast adverse weather events for aviation in a timely manner can lead to flight delays or cancellations, generating large economic costs for airlines and passengers. The new computational technologies that are being developed, allow an update of weather forecasting techniques, with artificial neural networks as a new paradigm of interest. In the present investigation, the forecast of the meteorological variables of horizontal visibility and maximum and minimum temperatures at the Trujillo airport was carried out using Backpropagation Neural Network (BNN), analyzing and identifying the climatological conditions and meteorological systems of influence that govern the place. Having as input variables for the BNN, daily data of the maximum and minimum temperature from the years 2008 to 2017 and hourly data of the horizontal visibility from the years 2013 to 2017 of the Trujillo airport, in addition to the explanatory variables such as global indices and other local weather variables. The BNNs used for maximum and minimum air temperature analysis consist of eight neurons in the input and hidden layer and an output layer of one neuron; and for horizontal visibility it consists of ten neurons in the input and hidden layer, and an output layer of one neuron. The results showed that, for the maximum and minimum temperature variables, the predicted and observed values are very close to each other with a similar trend; and for the horizontal visibility variable, significant differences are perceived, however, they do show similar trend patterns.

**Keywords:** Backpropagation Neural Network, forecast, visibility, temperature, Trujillo, aeronautics.

## I. INTRODUCCIÓN

La relación entre la aviación y la meteorología es tal, que es casi imposible hablar de cualquier aspecto de la primera sin hacer referencia al entorno meteorológico en el cual la aeronave opera. Las aeronaves son muy vulnerables a ciertas condiciones meteorológicas adversas y, aunque el tiempo es bueno la mayoría de las veces, el piloto debe aprender a reconocer y a respetar los fenómenos adversos. En tales situaciones, un piloto ha de ser capaz de tomar decisiones que minimicen su exposición a los mismos, que optimicen el rendimiento de la aeronave y, por encima de todo, que garanticen la seguridad del vuelo (González, 2005).

Wichter (2017) en *The New York Times* cita a Paul D. Williams, profesor del Departamento de Meteorología en la Universidad de Reading, Inglaterra, que explica que tendemos a ignorar la atmósfera y solo pensamos que el avión vuela a través de un espacio vacío, pero evidentemente no es así. Además, Williams estudia el cambio climático y su efecto en la aviación, y destaca que la atmósfera se está modificando debido al cambio climático por lo cual es necesario una actualización constante de las metodologías del pronóstico del tiempo, para así poder prever de manera oportuna los diferentes problemas que enfrenta la aviación cada vez más frecuentes a medida que el clima mundial cambia.

A mediados de junio del 2017, la aerolínea American Airlines se vio obligada a cancelar más de 40 vuelos en Phoenix, Estados Unidos; la razón fue que la temperatura máxima durante el día alcanzaba los 48 grados Celsius, teniendo demasiado calor para que algunas aeronaves pequeñas pudieran despegar, esto debido a que cuando el aire es muy caliente tiene una menor presión, lo cual dificulta que los aviones puedan elevarse (Wichter, 2017). Nuestro país no ha sido ajeno a los efectos de los fenómenos meteorológicos para la aviación; una ciudad que ha destacado por el gran número de vuelos cancelados por condiciones de mal tiempo ha sido Trujillo, La Libertad; teniendo en el 2014 un vuelo cancelado de la aerolínea LAN procedente de Lima que debió aterrizar a primera hora en el Aeropuerto Internacional Capitán FAP Carlos Martínez de Pinillos, debido a la densa neblina y llovizna

que cubriría la zona por la mañana (Redacción RPP, 2014); esto sumado a más casos presentados en el año 2018, en el mes de abril y mayo. En consecuencia, las cancelaciones o retrasos en los vuelos principalmente generan grandes costos económicos para las compañías aéreas y los pasajeros. Un estudio del 2010 encargado por la Administración Federal de Aviación de los Estados Unidos estima que los retrasos de vuelos cuestan a la industria aérea 8 billones de dólares al año, en gran parte debido al aumento de los gastos en las tripulaciones, combustible y mantenimiento; y los retrasos cuestan a los pasajeros casi 17 billones de dólares (The Associated Press, 2014).

Los centros meteorológicos aeronáuticos velan por proporcionar a los usuarios información meteorológica necesaria para el desempeño de las funciones aeronáuticas en fin del cumplimiento de vuelos planificados. El servicio meteorológico aeronáutico del Perú, dado por Corporación Peruana de Aeropuertos y Aviación Comercial, CORPAC S.A., tiene por finalidad contribuir a la seguridad, regularidad y eficiencia de la navegación aérea nacional e internacional, para ello proporciona la información meteorológica necesaria a los usuarios, miembros de la tripulación de vuelo, a las dependencias de los servicios de tránsito aéreo, de búsqueda y salvamento, a la administración de los aeropuertos y a todas las entidades interesadas en la explotación o desarrollo de la navegación aérea, para el desempeño de sus respectivas funciones (CORPAC S.A., 2018).

Las técnicas de pronóstico del tiempo se encuentran en constante actualización debido a las nuevas tecnologías computacionales que se van desarrollando, que son cada vez más sofisticadas y de fácil acceso. Las redes neuronales dentro del campo de la inteligencia artificial es un nuevo paradigma que ha despertado un gran interés en los últimos años. La novedad de este modelo es una nueva estructura computacional compuesta de un gran número de elementos de cómputo simples, neuronas, interconectados en paralelo. Estos modelos se inspiran en la Biología, en base a la arquitectura de la red y las conexiones que se establecen entre las neuronas como en el cerebro humano. Estos modelos se construyen por medio de algoritmos automáticos de aprendizaje, que ajustan las conexiones entre neuronas para modelizar un conjunto de datos dado. Las redes neuronales han probado su valía para resolver problemas complejos en diversas áreas, incluyendo a la predicción meteorológica y oceánica, siendo un campo muy consolidado en el que continuamente se desarrollan nuevos métodos y aplicaciones (Gutiérrez, Cano, Cofiño & Sordo, 2004).

En el campo de la meteorología aeronáutica se debe ser más precisos, ya que se busca prever



situaciones adversas como las presentadas en el aeropuerto internacional de la ciudad de Trujillo, con cancelaciones de vuelo por mal tiempo, en mucho de los casos por fenómenos meteorológicos de nieblas densas, que provoca malestares en los pasajeros, pérdidas económicas para las aerolíneas y dificultades en el tránsito aéreo. Por lo tanto, en la presente investigación se exploró una nueva metodología para el pronóstico de las variables de visibilidad horizontal y temperaturas máximas y mínimas para el área correspondiente al aeropuerto internacional de Trujillo, teniendo a las redes neuronales artificiales como una nueva tecnología computacional para actualizar las técnicas de pronóstico del tiempo y tener valores cada vez más cercanos a la realidad.

El objetivo general de esta tesis es realizar el pronóstico de las variables meteorológicas de visibilidad horizontal y temperaturas máximas y mínimas en el aeropuerto internacional de Trujillo utilizando las Redes Neuronales Artificiales de Retropropagación; por consiguiente, se plantearon los siguientes objetivos específicos:

- Analizar e identificar las condiciones climatológicas y sistemas meteorológicos de influencia que predominan sobre el área de estudio.
- Identificar las variables de entrada para la red neuronal del pronóstico de visibilidad horizontal y temperaturas máximas y mínimas.
- Determinar los algoritmos apropiados para el entrenamiento y aprendizaje de la red neuronal en la arquitectura seleccionada para el pronóstico de visibilidad horizontal y temperaturas máximas y mínimas.
- Validar el modelo que optimice el pronóstico, con menor grado de error, de las variables meteorológicas visibilidad horizontal y temperaturas máximas y mínimas.

## II. REVISIÓN DE LA LITERATURA

### 2.1. Antecedentes de la investigación

En los últimos años, se han desarrollado diversas investigaciones relacionadas a la predicción de series de tiempo utilizando redes neuronales artificiales, que es un nuevo paradigma que ha despertado gran interés debido al gran avance tecnológico que acontece.

En el campo de la meteorología, tenemos como una de las investigaciones pioneras la realizada por Marzban (2000) patrocinada por el Laboratorio Nacional de Tormentas Severas, el Instituto Cooperativo para Estudios de Meteorología en mesoescala; y el Departamento de Física de la Universidad de Oklahoma, en Norman, Oklahoma, Estados Unidos; para predecir tormentas. Se basan en atributos obtenidos de un radar Doppler, el cual observa diferentes fenómenos que a la larga llegan a producir tornados. Las tormentas eléctricas algunas veces llegan a producirlos, pero no siempre son antecedentes de ello. Se usó una red neuronal alimentada hacia adelante para diagnosticar cuales fenómenos detectados por el radar llegarían a producir un tornado. La red neuronal es diseñada para la identificación del *tornado-yielding*, con ese fin, se desarrollaron procedimientos para determinar el tamaño del conjunto de entrenamiento y el número de nodos ocultos necesarios para el funcionamiento óptimo. Se concluyó que la red neuronal desarrollada de este modo supera a un algoritmo basado en reglas.

Para Sudamérica; Ovando, Bocco y Sayago (2005) de la Universidad Nacional de Córdoba, utilizaron las Redes Neuronales del tipo "backpropagation" para predecir la ocurrencia de heladas a partir de las variables meteorológicas de temperatura, humedad relativa, nubosidad, dirección y velocidad del viento. Para el entrenamiento y la validación de la Red Neuronal se utilizó 24 años de datos meteorológicos obtenidos de la estación de Río Cuarto, Córdoba, Argentina, separados en 10 años correspondientes al conjunto de datos de entrenamiento y 14 años para el conjunto de datos de validación. Se construyeron diferentes modelos para evaluar el comportamiento de las redes utilizando distintos números de variables de entrada y/o neuronas en la capa oculta, teniendo las probabilidades de aciertos en los resultados de predicción para las distintas variables de entrada. Como resultados se

tuvieron que, en los modelos realizados, el porcentaje de días con error de pronóstico fue de 2 por ciento, aproximadamente para 14 años de aplicación; cuando se consideran días de heladas efectivas no pronosticadas los porcentajes oscilan entre un 10 y un 23 por ciento, para el mismo período. Estos muestran el buen desempeño y la pertinencia general de la metodología desarrollada en la estimación de fenómenos de comportamiento no lineal como las heladas.

En el Perú tenemos como ejemplo la investigación de Latínez (2009) titulada “Comparación de los métodos de regresión multivariada adaptativa utilizando Splines (MARS) y Redes Neuronales Artificiales Backpropagation (RNAB) para el pronóstico de lluvias y temperaturas en la cuenca del río Mantaro”. Para el trabajo mencionado se utilizaron datos de las estaciones de Huayao, Santa Ana, Jauja y Viques, los cuales fueron revisados y se eliminaron valores atípicos extremos procediendo al análisis con las técnicas Regresión Multivariada Adaptativa utilizando Splines (MARS) y las Redes Neuronales Artificiales Backpropagation (RNAB). La Red Neuronal utilizada para el análisis fue de 17 nodos en el caso de las precipitaciones y 15 para las temperaturas mínimas como máximas. Las variables explicativas para este estudio fueron variables globales provenientes de información secundaria. Los *inputs* utilizados tienen un desfase de tres meses ( $lag=3$ ). Las conclusiones a las cuales se llegó para este caso fueron que los pronósticos obtenidos por el modelo MARS tienen menor error que los obtenidos con las RNAB, a excepción de la variable Temperatura Máxima de Huayao en donde la RNAB resultó con menos errores que el modelo MARS.

## **2.2. Base teórica**

### **2.2.1. Análisis exploratorio de los datos**

El análisis exploratorio de datos consta de un conjunto de técnicas estadísticas cuyo objetivo es conseguir un entendimiento básico de los datos y de las relaciones existentes entre las variables analizadas. Se realiza mediante métodos sistemáticos sencillos para organizar y preparar los datos, detectar errores en la recolección de los datos, evaluación de datos ausentes, identificación de casos atípicos (*outliers*) y comprobación de los supuestos subyacentes en la mayor parte de las técnicas multivariantes (normalidad, linealidad, homocedasticidad) (Salvador, 2003). Los datos atípicos (*outliers*) son observaciones con características diferentes a los demás, no pueden ser caracterizados categóricamente como benéficos o problemáticos, sino que deben ser revisados en el contexto del análisis y a su vez se debe evaluar el tipo de información que puedan proporcionar, ya que pueden producir

aumento en la variabilidad. Los datos ausentes (*missing*) son habituales en el campo del análisis climático, se deben determinar las razones de las ausencias buscando entender el proceso para seleccionar la acción más apropiado para su tratamiento.

El análisis exploratorio de datos presenta las siguientes etapas:

1. Preparar los datos para su aplicación en alguna técnica estadística.
2. Realizar un análisis estadístico gráfico y numérico de las variables del estudio, para así tener una idea inicial de la información contenida en el conjunto de datos, pudiendo detectar oportunamente la existencia de posibles errores. Se pueden utilizar histogramas o *boxplot*.
3. Realizar un examen gráfico de las relaciones entre las variables analizadas y un análisis descriptivo numérico que cuantifique el grado de interrelación existente entre ellas. Posteriormente, evaluar, si fuera necesario, algunos supuestos básicos subyacentes a muchas técnicas estadísticas como, por ejemplo, la normalidad, linealidad y homocedasticidad.
4. Identificar los posibles casos atípicos (*outliers*) y evaluar el impacto potencial que puedan ejercer en análisis estadísticos posteriores. Los valores atípicos tienen gran influencia en el cálculo de la media, variancia y otros parámetros estadísticos.
5. Finalmente, evaluar, si fuera necesario, el impacto potencial que pueden tener los datos ausentes (*missing*) sobre la representatividad de los datos analizados.

#### 2.2.2. Climatología

La climatología consiste en el estudio del clima, el funcionamiento del sistema climático, sus variaciones y extremos. En la guía de prácticas climatológicas, OMM - N°100, se afirma que “en sentido estricto, se entiende por clima las condiciones meteorológicas normales correspondientes a un lugar y período de tiempo determinados” (OMM, 2018 p. 1). La climatología parte de descripciones estadísticas de diferentes variables meteorológicas como la temperatura, la precipitación, la presión atmosférica, la humedad y los vientos, o de la interacción de estos, como tipos y fenómenos meteorológicos, que son característicos de un lugar o región, durante cualquier período de tiempo.

#### 2.2.3. Pronóstico del tiempo

Los pronósticos son premisas o suposiciones básicas de que lo que puede suceder, a partir de los cuales se basan la planeación y la toma de decisiones. Las diferentes técnicas de

pronóstico nos ayudan a disminuir la incertidumbre sobre el futuro. Lo fundamental en los pronósticos es ser consistente y lograr el mayor porcentaje de acierto. En el campo de la meteorología, pronóstico se refiere a una declaración del estado anticipado de la atmósfera en un lugar o área en particular, en un período específico de tiempo. La palabra fue introducida por Robert FitzRoy, primer director del Departamento Meteorológico de la Junta de Comercio en el Reino Unido. Las previsiones se definen de acuerdo a los periodos que cubren: pronóstico del tiempo a largo plazo, previsión meteorológica a medio plazo y predicción meteorológica a corto plazo (Smith, 2006).

En el campo de la aeronáutica, la RAP (Regulaciones Aeronáuticas del Perú) 303: Servicio meteorológico para la navegación aérea, define como pronóstico la declaración de las condiciones meteorológicas previstas para una hora o período en específico y respecto a una cierta área o porción del espacio aéreo (Resolución Directoral N° 302-2010-MTC/12, 2010).

#### 2.2.4. Temperatura Máxima y Mínima del Aire

El aire está constituido por moléculas que se hallan en continuo movimiento y chocan entre ellas. La medida promediada de la energía cinética que generan estas moléculas, se le denomina temperatura del aire. Cuanto mayor es la actividad de estas moléculas más alta es la temperatura y, si el movimiento molecular fuera nulo, se alcanzaría la temperatura más baja posible. Se utilizan varias escalas de temperatura: Celsius, Fahrenheit y Kelvin (González, 2005).

La temperatura del aire está sujeta a factores de muy diversos, sin embargo, esta principalmente influenciada por la radiación solar. La temperatura media global de la atmósfera presenta desigualdades espaciales y un gran gradiente desde el Ecuador hasta los polos debido fundamentalmente a las variaciones latitudinales en el balance de radiación. El calor, procedente de la radiación, que se acumula en la superficie se cede a la atmósfera por cuatro mecanismos: conducción, convección, radiación infrarroja y evaporación, variando la temperatura del aire. La variación térmica diaria depende mucho de la radiación y se caracteriza por un máximo situado poco después de mediodía y un mínimo situado poco después de la salida del sol. Este esquema básico puede ser modificado por la nubosidad o por advecciones de aire frío o cálido (Alonso, 2000).

#### 2.2.5. Visibilidad Horizontal

En meteorología se define a la visibilidad horizontal como la máxima distancia horizontal a la que pueda verse y reconocerse un objeto de color negro de dimensiones convenientes,

situado cerca del suelo, al ser observado ante un fondo brillante; o la distancia máxima a la que puedan verse e identificarse las luces de aproximadamente 1000 candelas ante un fondo no iluminado, con ayuda del ojo exclusivamente. Se mide en metros o kilómetros (González, 2005).

#### 2.2.5.1. Fenómenos que reducen la visibilidad

Gonzales (2005) nos menciona que la transparencia del aire, que es a lo que se hace alusión cuando se habla de visibilidad horizontal, puede verse reducida por la presencia de partículas, sólidas o líquidas, en suspensión en la atmósfera, que absorben, dispersan y reflejan la luz. El grado de reducción de la visibilidad depende de la concentración de dichas partículas. A continuación, se describen los fenómenos que reducen la visibilidad:

- a) Niebla: es una nube en contacto con el suelo o a poca altura (hasta 15 metros sobre el suelo), que reduce la visibilidad horizontal a valores inferiores a 1000 metros, debido a la presencia de numerosas gotitas de agua suspendidas en el aire.
- b) Bruma: es una nube en contacto con el suelo o a poca altura (hasta 15 metros sobre el suelo), que reduce la visibilidad horizontal a valores comprendidos entre 1000 y 5000 metros. En el caso de la bruma, el número de gotitas suspendidas en el aire por unidad de volumen es menor a comparación de la niebla.
- c) Calima: es el enturbiamiento de la atmósfera por pequeñas partículas sólidas y secas, que se encuentran en suspensión. Son invisibles a simple vista, dando al cielo un aspecto opalescente. En este caso la visibilidad se reduce a valores inferiores o iguales a 5000 metros, con una humedad relativa de hasta 60 por ciento.
- d) Humo: es la suspensión en el aire de partículas producidas por combustión, que reducen la visibilidad a valores inferiores o iguales a 5000 metros, en función de su extensión o concentración.
- e) Smog: es el término utilizado para expresar la presencia de humo y niebla. El enturbiamiento que se produce en la atmósfera, se debe a la formación de niebla sobre núcleos de condensación que proceden de la combustión.
- f) Arena y polvo: son partículas en suspensión levantadas por el viento y en movimiento por turbulencia. Su presencia es favorecida por los vientos fuertes y una atmósfera inestable. Las tormentas de arena o de polvo pueden reducir la visibilidad hasta valores inferiores a 1000 metros.

g) Precipitaciones: reducen la visibilidad en función del tamaño de las gotas y de su concentración a los valores siguientes: lluvia moderada (de 3 a 10 kilómetros), lluvia fuerte (de 50 a 500 metros), llovizna moderada (de 500 metros a 3 kilómetros), llovizna densa (de 50 a 500 metros), nieve moderada (menos de 1000 metros), nieve fuerte (de 50 a 200 metros), granizo moderado (menos de 1000 metros) y granizo fuerte (de 50 a 200 metros)

De todos los fenómenos meteorológicos que reducen la visibilidad, las nieblas son particularmente peligrosas en las operaciones de aterrizaje, despegue y rodaje.

#### 2.2.6. Niebla

La niebla se define como la suspensión en el aire de gotitas muy pequeñas de agua o de cristales de hielo sobre la superficie, normalmente microscópicas, que reducen la visibilidad horizontal en las capas bajas de la atmósfera, a menos de 1000 metros de distancia (Gonzales, 2005). Se origina principalmente por la condensación o sublimación de la masa de aire húmedo, esta saturación de la masa de aire se produce por el enfriamiento de la masa de aire o la transferencia del vapor de agua de algún cuerpo libre de agua (como lagunas, lagos, ríos), la vegetación o el océano hacia la atmósfera. Vistin (2014) nos menciona que la niebla varía en su composición según a la temperatura en la que esta se desarrolle. Si la temperatura se encuentra superior a 0°C, estará constituida por gotitas en suspensión, pero si la temperatura es menor a 0°C estará constituida por cristales de hielo muy pequeños.

##### 2.2.6.1. Factores que influyen en la formación de niebla en la costa peruana

En la costa occidental de América del Sur, el clima es principalmente determinado por el Anticiclón del Pacífico Sur y la corriente peruana de aguas frías, especialmente para la presencia de nubes y nieblas en la costa peruana. El Anticiclón del Pacífico Sur consiste en masas de aire de alta presión, generando bajas precipitaciones y causando un área de gran aridez, que a su vez produce vientos suaves provenientes del sur y del sureste cargando a la atmósfera de la humedad que proviene del mar (Pérez, 2019). La corriente peruana, también llamada corriente de Humboldt, viaja de sur a norte, desde el tercio superior de Chile hasta los límites entre Perú y Ecuador. El movimiento de la Tierra empuja las aguas profundas, y por lo tanto frías, hacia el norte y hacia la superficie (Redacción BBC News Mundo, 2019). Según Gilford et. al (1992), la corriente fría de Humboldt regula la temperatura del aire debajo de los 610 m.s.n.m., la cual produce una capa límite marítima que no supera las 20 millas marinas tierra adentro, teniendo una influencia en las zonas costeras del continente.

La intensificación del Anticiclón del Pacífico Sur incrementa la intensidad de los vientos alisios que a su vez intensifica el afloramiento costero y si éstas son más frías que la región donde se traslada establecerá inversión térmica. En las zonas costeras es conocida como la inversión de los alisios, estableciendo así también un gradiente térmico entre el mar y el continente. Cuando el flujo marino del anticiclón atraviesa la superficie rugosa de la costa desértica, se produce una turbulencia atmosférica mecánica que tiende a igualar el contenido de vapor de agua en la capa, llegando el aire a saturarse a cierta distancia por debajo de la capa de fricción, dando lugar a la condensación en un nivel por encima del suelo, el cual se conoce como Nivel de condensación, este nivel se incrementa entre la base de la inversión térmica o inversión de subsidencia y la base de la nubosidad tipo estratiforme predominante (Quispe & Chira, 2015)

La formación de niebla principalmente se da cuando la temperatura del punto de rocío se aproxima a la temperatura de la superficie del mar. Pinche (1986) indica que el flujo del aire proveniente del Anticiclón del Pacífico Sur se desliza sobre aguas cada vez más frías a medida que se aproxima a la costa, enfriándose y, en consecuencia, la humedad relativa aumenta, creando las condiciones propicias para la formación de nubosidad estratiforme y nieblas típicas a lo largo de la costa.

Según indica el Meteorólogo Fernando Allasi (comunicación personal, 28 de marzo de 2022), la niebla en la costa peruana también está influenciada por la masa de agua subtropical superficial, que es una masa de agua cálida con temperatura superior de los 17°C localizada hacia el oeste de la corriente marina peruana. La incursión de la masa de agua subtropical superficial oceánica hacia la costa ocasiona la transferencia del vapor de agua hacia la atmósfera. La concentración del vapor de agua en la atmósfera asociada con el descenso de la temperatura del aire, genera en la madrugada y primeras horas de la mañana la saturación de la masa de aire, esta saturación ocasiona la transformación física del vapor de agua de estado gaseoso a líquido que origina la niebla.

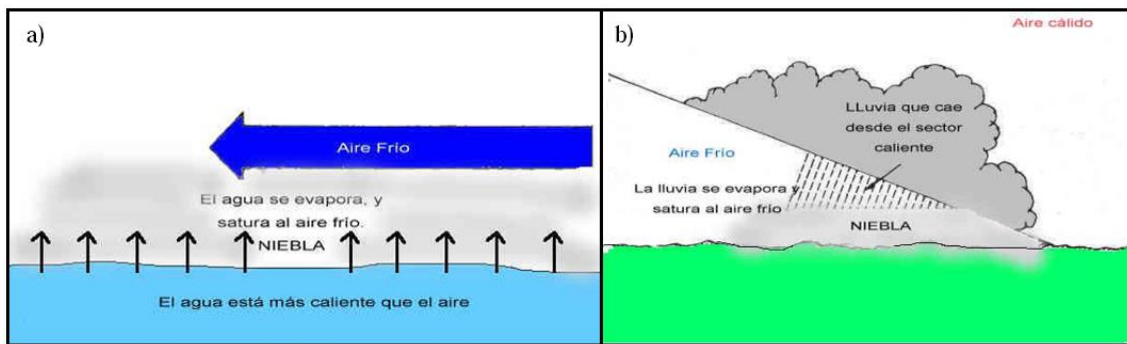
#### 2.2.6.2. Clasificación de la niebla de acuerdo a su génesis

Existen algunos procesos físicos para la generación de la niebla en la costa peruana, en las interfaces océano - atmósfera, océano - costa o la costa - atmósfera, de los cuales se destacan los siguientes:

A. Niebla de evaporación: Se producen cuando se evapora agua en el aire frío. Este cambio de estado del agua puede ocurrir de dos maneras: cuando una corriente de aire frío y



relativamente seco fluye o permanece en reposo sobre una superficie de agua de mayor temperatura, siendo común en las zonas polares y sobre los lagos y lagunas en invierno; o cuando llueve, si el agua que cae tiene mayor temperatura que el aire del entorno provoca que las gotas de lluvia se evaporan saturando el aire, principalmente se originan dentro del aire frío de los frentes de lento movimiento como los estacionarios, calientes o los frentes fríos lentos (Aránguiz, Morales, Nieto, & Silva, 2009).



**Figura 1.** Nieblas de evaporación: **a)** Esquema de niebla de evaporación por aire frío y **b)** Esquema de niebla de evaporación por lluvia.

**FUENTE:** Aránguiz, Morales, Nieto, & Silva (2009)

B. Niebla por enfriamiento: Es ocasionada mediante la disminución que experimenta la capacidad del aire para contener vapor de agua cuando disminuye la temperatura. Se clasifica en tres tipos:

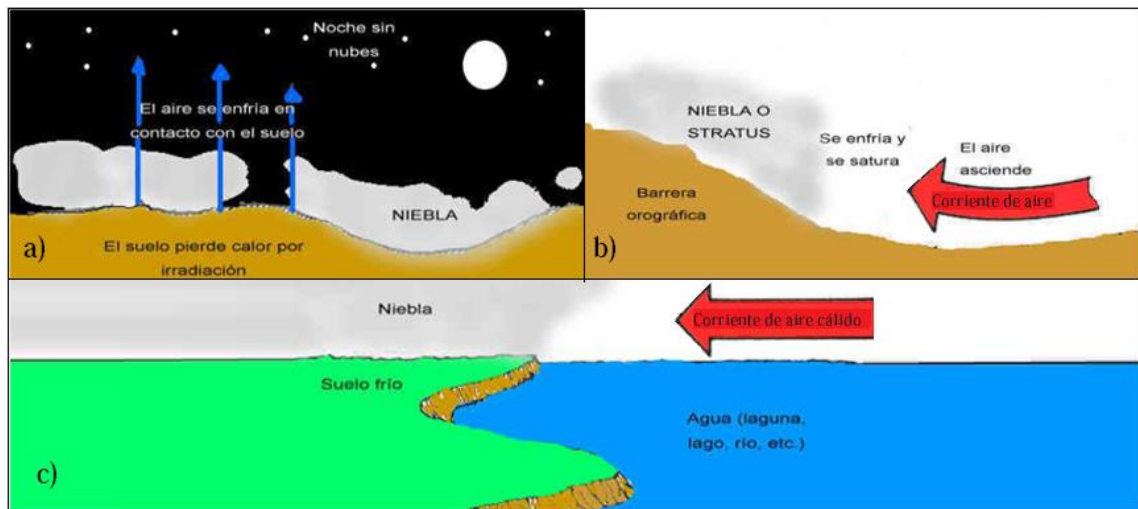
- Niebla advectiva: Según Pettersen (1956), es producida por el transporte de masas de aire húmedo sobre una superficie fría, provocando el enfriamiento de los niveles superficiales hasta el punto de rocío, es relativamente superficial y estable caracterizada por una inversión superficial. Para que este tipo de niebla se forme es necesario que el viento sople con una intensidad entre 4 y 12 nudos para que se pueda mantener constante el flujo de aire cálido y húmedo, si se sobrepasan estos valores la niebla se desprenderá del suelo generando una nube baja llamada estrato turbulento; y si el aire, por el contrario, está calmo, el vapor de agua se depositará sobre el suelo formando rocío. Se producen principalmente en zonas costeras, especialmente en invierno, cuando el aire relativamente más cálido y húmedo procedente del mar fluye hacia la tierra más fría; por el contrario, en verano se producen de forma inversa, es decir sobre el mar, cuando el aire más cálido de la tierra se desplaza sobre el agua relativamente más fría (Aránguiz, Morales, Nieto, & Silva, 2009).

Según indica el Meteorólogo Fernando Allasi (comunicación personal, 28 de marzo de 2022), la circulación del vapor de agua en la costa peruana se origina de la dinámica del Anticiclón del Pacífico sur o la interacción simultánea de la Baja presión Subpolar y el

Anticiclón del Pacífico sur en el océano o la costa, donde se identifican los siguientes patrones generadores de la niebla en la costa: la deformación ondulatoria del Anticiclón Pacífico, sur sobre la costa del Perú, la configuración zonal de oeste a este del Anticiclón Pacífico sur sobre el océano, la configuración meridional de sur a norte del Anticiclón Pacífico sur sobre la costa del Perú, la incursión de la cuña del Anticiclón Pacífico sur sobre la costa, el cizalle (fluctuación) del viento del Anticiclón Pacífico sur sobre la costa meridional del Perú, la incursión zonal del Anticiclón Pacífico sur sobre Sudamérica, la interacción simultánea de la Baja presión Subpolar y el Anticiclón del Pacífico sur sobre la costa o la contracción espacial y la deformación meridional de sur a norte del Anticiclón Pacífico sur sobre la costa, ocasionado por la interacción másica de la Baja presión Subpolar sobre la zona occidental del Anticiclón del Pacífico sur, la intensificación del Anticiclón Pacífico sur sobre la costa del Perú o la división del Anticiclón Pacífico sur y la localización de un núcleo secundario de Alta presión sobre la costa oeste Sudamérica.

- La niebla de irradiación: ocurre cuando el suelo pierde calor a través de la emisión de radiación infrarroja, por lo que el suelo enfriado produce la condensación del aire próximo a éste, es una niebla de corta duración y poca altura (Aránguiz, Morales, Nieto, & Silva, 2009). Se producen principalmente en inviernos y sobre el relieve de la costa (desierto, pampas, lomas, valles costeros). Según indica el Meteorólogo Fernando Allasi (comunicación personal, 28 de marzo de 2022), la disipación de la nubosidad estratiforme generadora de la irradiación de la superficie está asociada con las siguientes configuraciones sinópticas de los sistemas meteorológicos en la costa: el Anticiclón Pacífico sur mediante la incursión de la cuña del Anticiclón Pacífico sur sobre la costa del Perú; la circulación Anticiclónica del océano Pacífico sur a 500 mb mediante la cuña de la circulación Anticiclónica del océano Pacífico sur a 500 mb sobre la atmósfera la costa del Perú, la circulación Anticiclónica del océano Pacífico sur a 500 mb sobre la atmósfera de la costa del Perú, la incursión y el posicionamiento de la circulación Anticiclónica del océano Pacífico sur a 500 mb sobre la atmósfera la costa del Perú y la elongación meridional de la cuña de la circulación Anticiclónica del océano Pacífico sur a 500 mb a lo largo de la costa; y la circulación Anticiclónica de Sudamérica a 500 mb mediante la incursión de la circulación Anticiclónica de Sudamérica a 500 mb sobre la atmósfera de la costa del Perú, la elongación y la disgregación de la circulación Anticiclónica de Sudamérica a 500 mb sobre la atmósfera de la costa y la cuña de la circulación Anticiclónica de Sudamérica a 500 mb sobre la atmósfera la costa del Perú.

- Niebla orográfica: se genera por las corrientes de aire que ascienden sobre las laderas montañosas o elevaciones del terreno. Esto se debe a que cuando el aire asciende, se expande y se enfría. Este enfriamiento conlleva un aumento de la humedad relativa pudiendo alcanzarse la saturación, siendo necesario que la humedad relativa inicial sea elevada y que el viento sea persistente y no muy intenso (Aránguiz, Morales, Nieto, & Silva, 2009).



**Figura 2.** Nieblas por enfriamiento: a) Esquema de niebla de radiación, b) Esquema de niebla orográfica, y c) Esquema de niebla de advección.

**FUENTE:** Aránguiz, Morales, Nieto, & Silva (2009)

## 2.2.7. Redes neuronales artificiales

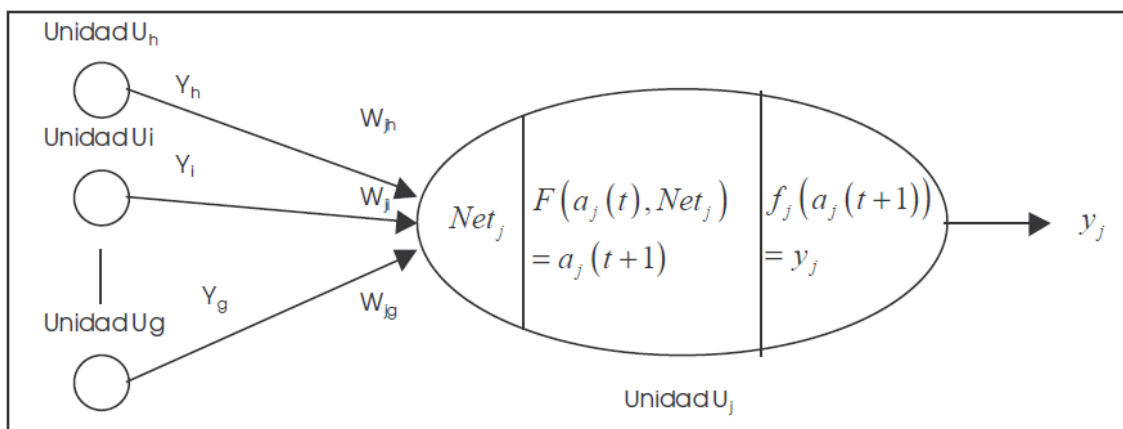
### A. Definición

Su concepto es definido por varios autores, pero el más aceptado es el mencionado por Hetch – Nielsen (1988), donde define como un Sistema de computación constituido por un gran número de elementos simples de procesamiento muy interconectados, que procesan información por medio de su estado dinámico como respuesta a entradas externas. Las Redes Neuronales poseen las siguientes ventajas: A) Poseen un aprendizaje adaptativo, donde los sistemas son dinámicos y auto adaptativos a nuevas condiciones. B) Autoorganización, consiste en una modificación de la red neuronal para llevar a cabo un objetivo específico, logrando resolver problemas donde la información de entrada no es muy clara. C) Tolerancia a fallos, la información que se maneja se distribuye por todas las neuronas generando cierto grado de redundancia. D) Las redes neuronales pueden ser entrenadas y trasladadas a cualquier hardware de bajo costo gracias a las diferentes herramientas computacionales que existen hoy en día.

## B. Elementos de la Red Neuronal

Como toda neurona biológica posee sus elementos, también los tiene la red neuronal, donde Haykin (1999) lo distribuye así: A) Entradas o nodos de entrada: son los valores que se ingresan a la red, de acuerdo al tipo de problema que se está estudiando. B) Salidas o nodos de salida: valores que arroja el modelo de red como un resultado del aprendizaje. C) Peso sináptico: son los valores numéricos que expresan la importancia que corresponde a la entrada de la red. D) Un punto de suma de entradas ponderadas: se realiza una suma de todas las entradas multiplicadas por sus respectivos pesos sinápticos. E) Una función de activación: es aquel que define el rango de salida de toda neurona. F) Sesgo: es aquel valor de entrada fija e igual a 1.

Para describir a mejor detalle los conceptos básicos de cada uno de los elementos de la red neuronal, Hilera y Martínez (2000) nos detallan que, si se tienen  $N$  neuronas, estas se pueden ordenar arbitrariamente y designarlas como una  $j$ -ésima unidad ( $U_j$ ). Su función de estas es recibir las entradas de las células vecinas y calcular el valor de salida, el cual es enviado a todas las células restantes (Figura 3).



**Figura 3.** Entradas y salidas de una neurona  $U_j$ .  
**FUENTE:** Hilera y Martínez (2000).

Todas las neuronas que componen la red neuronal se hallan en cierto estado. Se puede decir que hay dos posibles estados, en reposo o excitado; estos se denominan estados de activación, y a estos se les asigna un valor. Los valores de activación pueden ser continuos o discretos, pudiendo ser limitados o ilimitados. Estos estados de activación del conjunto de neuronas en el tiempo  $(t)$ , son especificados en un vector de  $N$  números reales  $A(t)$ . La activación de una unidad  $U_i$  en el tiempo  $t$  se designa por  $a_i(t)$ :

$$A(t) = (a_1(t), a_2(t), \dots; a_i(t), \dots, a_N(t)) \quad (1)$$

Entre las neuronas que forman la red neuronal existe un conjunto de conexiones que unen unas con otras. Cada neurona transmite señales a las que están conectadas a su salida. Cada neurona  $U_i$  tiene una función de salida  $f_i(a_i(t))$ , que transforma el estado de activación  $a_i(t)$  en una señal de salida  $y_i(t)$  de la forma:

$$y_i(t) = f_i(a_i(t)) \quad (2)$$

El vector que contiene las salidas de todas las neuronas en un instante  $t$  es:

$$Y(t) = (f_1(a_1(t)), f_2(a_2(t)), \dots, f_i(a_i(t)), \dots, f_N(a_N(t))) \quad (3)$$

Hilera y Martínez (2000) mencionan que las conexiones que unen a las neuronas que forman la red neuronal tienen un peso asociado, que es el que hace que la red adquiera conocimiento. Consideramos  $y_i$  como el valor de salida de una neurona  $i$  en un instante dado. Una neurona recibe un conjunto de señales que le dan información del estado de activación de todas las neuronas con las que se encuentra conectada. Cada conexión entre la neurona  $i$  y la neurona  $j$  está ponderada por un peso  $w_{ji}$ . Normalmente, se considera que el efecto de cada señal es aditivo, de tal forma que la entrada neta que recibe una neurona net  $j$  es la suma del producto de cada señal individual por el valor de la sinapsis que conecta ambas neuronas.

$$net_j = \sum_i^N w_{ji} \cdot y_i \quad \dots \quad (4)$$

La ecuación 4 muestra el procedimiento a seguir para combinar los valores de entrada a una unidad con los pesos de las conexiones que llegan a esa unidad y es conocida como función de propagación. Suele utilizarse una matriz  $W$  con todos los pesos  $w_{ji}$  que reflejan la influencia que tiene la neurona  $i$  sobre la neurona  $j$ .  $W$  es un conjunto de elementos positivos, negativos o nulos. Si  $w_{ji}$  es positivo, indica que la interacción entre las neuronas  $i$  y  $j$  es excitadora; es decir, siempre que la neurona  $i$  este activada, la neurona  $j$  recibirá una señal de  $i$  que tenderá a activarla. Si  $w_{ji}$  es negativo, la sinapsis será inhibitoria. En este caso, si  $i$  esta activada, enviará una señal a  $j$  que tenderá a desactivar a esta. Finalmente, si  $w_{ji} = 0$ , se supone que no hay conexión entre ambas.

También hay la necesidad de una función que combine las entradas con el estado actual de la neurona para producir un nuevo estado de activación. Esta función  $F$  produce un nuevo estado de activación en una neurona a partir del estado ( $a_i$ ) que existía y la combinación de las entradas con los pesos de las conexiones ( $net_i$ ). Dado el estado de activación  $a_i(t)$  de la

unidad  $U_i$  y la entrada total que llega a ella,  $Net_i$ , el estado de activación siguiente,  $a_i(t+1)$ , se obtiene aplicando una función  $F$ , llamada función de activación reflejada en la siguiente ecuación:

$$a_i(t + 1) = F(a_i(t), Net_i) \quad (5)$$

En la mayoría de los casos,  $F$  es la *función identidad*, por lo que el estado de activación de una neurona  $t + 1$  coincidirá con el  $Net$  de la misma en  $t$ . En este caso, el parámetro que se le pasa a la función de salida,  $f$ , de la neurona será directamente el  $Net$ . El estado de activación anterior no se tiene en cuenta. Según esto, la salida de una neurona  $i$  ( $y_i$ ) quedará así:

$$y_i(t + 1) = f(Net_i) = f\left(\sum_{j=1}^N w_{ij} y_j(t)\right) \quad (6)$$

En lo sucesivo, se considera únicamente la función  $f$ , a la que se le denominará indistintamente función de activación. Además, esta normalmente no está centrada en el origen del eje que representa el valor de la entrada neta, sino que existe cierto desplazamiento debido a las características internas de la propia neurona y que no es igual en todas ellas. Este valor se denota como  $\theta_i$  y representa el umbral de activación de la neurona  $i$ .

$$y_i(t + 1) = f(Net_i - \theta_i) = f\left(\sum_{j=1}^N w_{ij} y_j(t) - \theta_i\right) \quad (7)$$

Finalmente, en el caso de las redes neuronales, se puede considerar que el conocimiento se encuentra representado en los pesos de las conexiones entre neuronas. Todo proceso de aprendizaje implica cierto número de cambios en estas conexiones. En realidad, puede decirse que se aprende modificando los valores de los pesos de la red.

### C. Capas de la Red Neuronal

Toda red neuronal posee capas o niveles, siendo las siguientes: a) Capa de entrada: es aquella que recibe directamente la información de las fuentes externas de la red. b) Capa oculta: son internas a la red y no tienen contacto directo con el exterior. El número de niveles de capas ocultas que puede tener una red neuronal se encuentra entre 0 y un número elevado. Cada neurona en su capa oculta puede estar interconectada de diferentes formas, de acuerdo a la topología de red que se escoja. c) Capa de salida: transportan la información de la red hacia el exterior.

#### D. Fases de la Red Neuronal

Toda red neuronal posee fases, siendo una de ellas la de aprendizaje, el cual consiste en la adaptación de los pesos sinápticos a la información que extrae de los patrones o procesos de entrenamientos, logrando que dichos pesos se minimicen o maximicen de acuerdo a alguna función de energía. Y la fase de pruebas, en donde los resultados que brinda la fase anterior pueden ser calculados de una vez o adaptarlos de una forma iterativa, dependiendo del tipo de red neuronal.

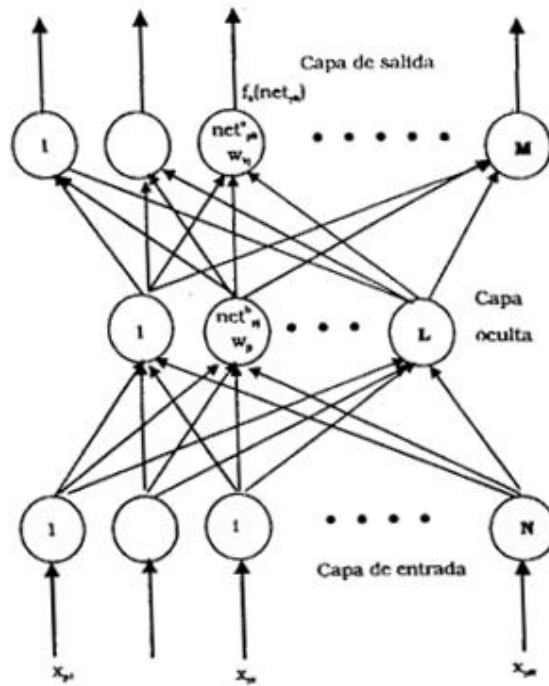
#### 2.2.8. Redes neuronales artificiales de retropropagación

##### A. Definición.

Es un método propuesto por Werbos (1974) y reinventado por Parker (1987) para el entrenamiento de las capas ocultas llamado Red *Backpropagation* o de retropropagación en español, cuyas características son las siguientes: a) derivación de las funciones de activación, consiste en las funciones sigmoideal y tangente hiperbólica, b) modos de entrenamiento que puedan poseer para la actualización de los pesos y c) batch, donde la actualización de los pesos se realiza gracias a la presentación de los ejemplos de entrenamiento que ello constituye. Una de sus desventajas, es la paralización de la red provocada por el nivel máximo que pueda alcanzar la capa de entrada y la búsqueda anormal del mínimo valor en la superficie del error.

##### B. Mecanismo de funcionamiento.

El funcionamiento de una red de retropropagación consiste en un aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo, empleando un ciclo propagación-adaptación de dos fases: primero se utiliza un patrón de entrada como estímulo para la primera capa de neuronas de la red, esta se va propagando a través de todas las capas superiores hasta generar una salida, finalmente se compara el resultado obtenido en las neuronas de salida con la salida. A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su relación relativa al error total. Basándose en el valor del error obtenido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez que se muestre el mismo patrón, la salida esté más cercana a la deseada teniendo un menor error.



**Figura 4.** Modelo de arquitectura de una red backpropagation.  
**FUENTE:** Hilera y Martínez (2000).

### C. Aplicación del algoritmo de retropropagación

Hilera y Martínez (2000) presentan los pasos y ecuaciones que se utilizan para aplicar el algoritmo de entrenamiento:

1. Empezar con valores pequeños aleatorios en los pesos de la red.
2. Presentar un patrón de entrada  $X_p: x_{p1}, x_{p2}, \dots, x_{pN}$ , y especificar la salida deseada que debe generar la red  $d_1, d_2, \dots, d_M$  (si la red se utiliza como un clasificador, todas las salidas deseadas serán cero, excepto una, que será la de la clase a la que pertenece el patrón de entrada).
3. Obtener la salida actual de la red, para ello se presentan las entradas a la red y se va calculando la salida que presenta cada capa hasta obtener la capa de salida, esta será la salida de la red  $y_1, y_2, \dots, y_M$ . Los pasos son los siguientes:

- Calcular las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada. Para una neurona  $j$  oculta:

$$\text{net}_{pj}^h = \sum_{i=1}^N w_{ji}^h x_{pi} + \theta_j^h \quad (8)$$

donde el índice  $h$  hace referencia a las magnitudes de la capa oculta; el subíndice  $p$ , al  $p$ -ésimo vector de entrenamiento, y  $j$  a la  $j$ -ésima neurona oculta. El termino  $\theta$  es opcional ya que actúa como una entrada más.



- Calcular las salidas de las neuronas ocultas:

$$y_{pj} = f_j^h(\text{net}_{pj}^h) \quad (9)$$

- Realizar los mismos cálculos para obtener las salidas de las neuronas de salida (capa *o*: output)

$$\text{net}_k^o = \sum_{j=1}^L w_{kj}^o y_{pj} + \theta_k^o \quad (10)$$

$$y_{pk} = f_k^o(\text{net}_{pk}^o) \quad \dots \quad (11)$$

4. Obtener los términos de error para todas las neuronas. Si la neurona *k* es una neurona de la capa de salida, el valor de delta es:

$$\delta_k^o = (d_{pk} - y_{pk}) f_k^o(\text{net}_{pk}^o) \quad (12)$$

La función *f*, que se mencionó anteriormente, debe cumplir con el requisito de ser derivable, lo que implica la imposibilidad de usar una función escalar. En general, disponemos de dos formas de función de salida que nos pueden servir: la función lineal de salida ( $f|k(\text{net}_{jk}) = \text{net}_{jk}$ ) y la función sigmoideal definida por la expresión:

$$f_k(\text{net}_{jk}) = \frac{1}{1+e^{-\text{net}_{jk}}} \quad (13)$$

La selección de la función de salida depende de la forma en la que se decida representar los datos de salida: si se desea que las neuronas de salida sean binarias, se utiliza la función sigmoideal, puesto que esta función es casi estable y, además es derivable. En otros casos es tan aplicable una función como otra.

Para la función lineal, tenemos:  $f_k^{o'} = 1$ , mientras que la derivada de una función *f* sigmoideal es:

$$f_k^{o'} = f_k^o(1 - f_k^o) = y_{pk}(1 - y_{pk}) f_k^{o'} = f_k^o(1 - f_k^o) = y_{pk}(1 - y_{pk}) \quad (14)$$

por lo que los términos de error para las neuronas de salida nos quedan:

$$\delta_{pk}^o = (d_{pk} - y_{pk}) \quad (15)$$

para la salida lineal, y

$$\delta_{pk}^o = (d_{pk} - y_{pk}) y_{pk}(1 - y_{pk}) \quad (16)$$

para la salida sigmoïdal.

Si la neurona  $j$  no es de salida, entonces la derivada parcial del error no debe ser evaluada directamente. Por tanto, se obtiene el desarrollo a partir de valores que son conocidos y otros son evaluados. La expresi3n obtenida en este caso es:

$$\delta_{pj}^h = f_j^h(\text{net}_{pj}^h) \sum_k \delta_{pk}^o w_{kj}^o \quad (17)$$

donde  $k$  hace referencia a todas las neuronas de la capa superior a la de la neurona  $j$ . Entonces, el error que se obtiene en una neurona oculta es proporcional a la suma de los errores conocidos que resultan en las neuronas a las que est1 conectada la salida de esta, multiplicado cada uno de ellos por el peso de la conexi3n. Los umbrales internos de las neuronas se adaptan de forma similar, considerando que est1n conectados con los pesos desde entradas auxiliares de valor constante.

#### 5. Actualizaci3n de los pesos.

Para ello, se utiliza el algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atr1s hasta llegar a la capa de entrada, ajustando los pesos de la forma siguiente:

para los pesos de las neuronas de la capa de salida se tiene

$$\begin{aligned} w_{kj}^o(t+1) &= w_{kj}^o(t) + \Delta w_{kj}^o(t+1); \\ \Delta w_{kj}^o(t+1) &= \alpha \delta_{pk}^o y_{pj} \end{aligned} \quad (18)$$

y para los pesos de las neuronas de la capa oculta

$$\begin{aligned} w_{ji}^h(t+1) &= w_{ji}^h(t) + \Delta w_{ji}^h(t+1); \\ \Delta w_{ji}^h(t+1) &= \alpha \delta_{pj}^h x_{pi} \end{aligned} \quad (19)$$

En ambos casos, para acelerar el proceso de aprendizaje, se puede a1adir un t1rmino momento de valor:  $\beta (w_{ji}^h(t) - w_{ji}^h(t-1))$  cuando se trata de una neurona oculta.

6. El proceso se repite hasta que el termino de error

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2 \quad (20)$$

resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

### 2.2.9. Evaluación de pronóstico

Existen ciertos métodos que evalúan el cumplimiento del modelo de pronóstico realizado, siendo los siguientes definidos por Pielke (1984) y Stauffer y Seaman (1990):

#### A. Error Cuadrático Medio (ECM)

Consiste en la suma de las diferencias entre el valor pronosticado y el observado (aquellos datos obtenidos en la estación meteorológica) por cada periodo, elevadas al cuadrado, dividiendo el resultado entre el número total de periodos. Se tiene que la precisión del modelo es mayor cuando el valor del ECM es menor. Se puede expresar en la ecuación 21:

$$ECM = \sum_{i=1}^N \frac{(x_i - x_{iobs})^2}{N} \quad (21)$$

Donde  $x_i$  es el valor pronóstico para la fila  $i$ ,  $x_{iobs}$  es el valor observado para la fila  $i$ , y  $N$  es el número de valores analizados.

#### B. Raíz del Error Cuadrático Medio (RECM)

La RECM calcula la raíz cuadrada de la medida de las diferencias en promedio entre los valores pronósticos y los observados. Esta es mejor opción a usar que el ECM, ya que este último da como resultado unidades elevadas al cuadrado. Se puede expresar en la ecuación 22:

$$RECM = \sqrt{\sum_{i=1}^N \frac{(x_i - x_{iobs})^2}{N}} \quad (22)$$

Donde  $x_i$  es el valor pronóstico para la fila  $i$ ,  $x_{iobs}$  es el valor observado para la fila  $i$ , y  $N$  es el número de valores analizados.

#### C. Error Absoluto Medio (EAM)

Mide el error global de un modelo de pronóstico calculándose a partir de la suma de todos los errores absolutos de todos los periodos, entre el número total de aquellos periodos. En términos generales, es un promedio de los errores donde un valor alto representa un error alto del modelo. Se puede expresar en la ecuación 23:

$$EAM = \sum_{i=1}^N \frac{|x_i - x_{iobs}|}{N} \quad (23)$$

Donde  $x_i$  es el valor pronóstico para la fila  $i$ ,  $x_{iobs}$  es el valor observado para la fila  $i$ , y  $N$  es el número de valores analizados.

#### D. Error Absoluto Medio Normalizado (EAMN)

Para tener en cuenta el peso del error respecto al valor de la variable medida, se debe normalizar el error absoluto. Se expresa en la ecuación 24:

$$EAMN = \sum_{i=1}^N \frac{|x_i - x_{iobs}| x_{iobs}}{N} \quad (24)$$

Donde  $x_i$  es el valor pronóstico para la fila  $i$ ,  $x_{iobs}$  es el valor observado para la fila  $i$ , y  $N$  es el número de valores analizados.

#### E. BIAS

También conocido como sesgo. Proporciona información sobre la tendencia del modelo a sobrestimar o subestimar una variable, cuantificando el error sistemático. Un valor positivo indica una sobrestimación del modelo y un valor negativo indica una subestimación. Se expresa en la ecuación 25:

$$BIAS = \sum_{i=1}^N \frac{(x_i - x_{iobs})}{N} \quad (25)$$

Donde  $x_i$  es el valor pronóstico para la fila  $i$ ,  $x_{iobs}$  es el valor observado para la fila  $i$ , y  $N$  es el número de valores analizados.

#### F. Correlación (COR)

La correlación implica la fuerza y dirección de una relación lineal entre dos variables aleatorias. Dos variables cuantitativas están correlacionadas cuando los valores de una de ellas varían sistemáticamente con respecto a los valores homónimos de la otra. Por ejemplo, si tenemos dos variables (X e Y) existe correlación si al aumentar los valores de X lo hacen también los de Y, y viceversa. Para la evaluación del pronóstico, la correlación es utilizada como un indicador de que los valores pronosticados tienen el sentido y la fuerza de una relación lineal con los valores observados. Se tiene un buen modelo de pronóstico cuando la correlación entre los valores pronosticados y observados tenga un valor cercano a uno y que sea positiva. Freedman, Pisani, Purves y Adhikari (1993) lo definen a partir de la siguiente relación, expresada en la ecuación 26:

$$COR = r_{xy} = \frac{Cov_{xy}}{S_x S_y} \quad (26)$$

donde:

$$Cov_{xy} = \sum_{i=1}^N \frac{(x_i - \bar{x})(y_i - \bar{y})}{N-1}, S_x = \sqrt{\sum_{i=1}^N \frac{(x_i - \bar{x})^2}{N-1}} \text{ y } S_y = \sqrt{\sum_{i=1}^N \frac{(y_i - \bar{y})^2}{N-1}}.$$

### III. MATERIALES Y MÉTODOS

#### 3.1. Materiales y equipos

##### 3.1.1. Descripción de la zona de estudio

El Aeropuerto “Cap. Carlos Martínez De Pinillos” se encuentra ubicado en Perú, la región La Libertad, provincia de Trujillo y distrito de Huanchaco a 11 Km. de la ciudad. Posee una temperatura de referencia de 25.8 ° C, una elevación de 39 m.s.n.m. y sus coordenadas geográficas son: 08° 04’54.15’’Sur - 79°06’31.13’’ Oeste (CORPAC S.A.).



**Figura 5.** Mapa de ubicación del Aeropuerto “Cap. Carlos Martínez De Pinillos”.

**FUENTE:** Google Earth (2022)

##### 3.1.2. Información a procesar

Los datos de visibilidad horizontal y temperaturas máximas y mínimas fueron tomados de la estación meteorológica aeronáutica del aeropuerto de Trujillo, los cuales fueron proporcionados por CORPAC S.A. Esta estación se encuentra ubicada aproximadamente en la latitud 08°06'S y longitud 79°02', con una altitud de 26 m. Se utilizaron cinco años de datos horarios de visibilidad horizontal, de los años 2013 al 2017, y diez años de datos diarios de temperaturas máximas y mínimas, de los años 2008 al 2017.

Se utilizaron como variables explicativas otras variables meteorológicas tomadas de la estación meteorológica aeronáutica del aeropuerto de Trujillo como la humedad relativa y el viento (velocidad y dirección); y un conjunto de índices globales como: índices globales de la temperatura superficial del mar en el Pacífico ecuatorial, conocidos como las regiones Niño 1+2, Niño 3, Niño 4 y Niño 3.4 que provienen del *Earth System Research Laboratory*<sup>1</sup> (ESRL); la presión atmosférica a nivel del mar en Darwin, Tahití y el Índice de Oscilación del Sur (SOI), provienen del *Climate Prediction Center*<sup>2</sup> (CPC) de la NOAA<sup>3</sup>. Estas últimas variables explicativas se recolectaron a través de internet, accediendo desde cada portal antes mencionado. En total, para las variables temperatura mínima y máxima del aire se utilizaron 08 variables explicativas correspondientes a un análisis diario; y para la variable visibilidad horizontal se utilizaron 10 variables explicativas correspondientes a un análisis horario.

### 3.1.3. Programas y equipos

Se utilizaron los siguientes programas informáticos para el desarrollo de la presente investigación:

- Análisis exploratorio de datos y desarrollo de climatología / Microsoft Office Excel 2019:

Excel es un programa desarrollado por *Microsoft* que pertenece a la línea Office, para *Windows*, *macOS*, *Android* e *iOS*. Se compone de una hoja de cálculo que permite manipular datos numéricos y de texto en tablas formadas por la unión de filas y columnas. Cuenta con cálculo, gráficas, tablas y un lenguaje de programación macro llamado *Visual Basic* para aplicaciones.

- Metodología de Redes Neuronales Artificiales / Weka 3.8<sup>4</sup>:

Weka es un conjunto de librerías en JAVA que permite la extracción de conocimientos a partir de bases de datos. Es un software que ha sido desarrollado en la universidad de Waikato (Nueva Zelanda) bajo licencia GPL, siendo una de las suites más utilizadas en los últimos años. Contiene las herramientas necesarias para realizar transformaciones, tareas de clasificación, regresión, *clustering*, asociación y visualización. Está diseñada como una herramienta a la cual se le pueden agregar extensiones para así añadir nuevas funcionalidades de manera sencilla. Sin embargo, se tiene escasa documentación orientada al usuario, lo que la hace una herramienta difícil de comprender y manejar sin

---

<sup>1</sup> Se pueden bajar desde la siguiente dirección: <http://www.esrl.noaa.gov/>

<sup>2</sup> <http://www.cpc.ncep.noaa.gov/>

<sup>3</sup> NOAA – National Oceanic and Atmospheric Administration

<sup>4</sup> <http://www.cs.waikato.ac.nz/ml/weka/>

información adicional; por lo cual se recomienda revisar manuales para el mejor entendimiento<sup>5</sup>.

Se utilizaron los siguientes equipos para el desarrollo de la presente investigación:

- *Laptop* (AMD Ryzen, 8 GB de RAM).
- USB 16Gb.
- Materiales de escritorio.

### 3.2. Desarrollo metodológico

#### 3.2.1. Análisis e identificación de las condiciones climatológicas y sistemas meteorológicos de influencia que predominan sobre el área de estudio.

Antes procesar los datos se procedió con el análisis exploratorio mediante la metodología planteada por la investigación de Latínez (2009), mediante técnicas como diagramas de cajas para detectar errores de medición, como datos faltantes o datos atípicos. Para la detección de valores atípicos se utilizó la diferencia entre el primer cuartil  $Q_1$  y el tercer cuartil  $Q_3$ , es decir el rango Intercuantil (IQR). En un diagrama de caja se consideran dos tipos de valores atípicos: valor atípico leve y valor atípico extremo.

Para calcular el valor atípico leve tenemos que:

$$LI: < Q_1 - 1.5 \times IQR, \text{ ó } LS: > Q_3 + 1.5 \times IQR. \quad (27)$$

Los resultados determinan los límites interiores  $\langle LI_i, LS_i \rangle$ , a partir de los cuales la observación se considera un atípico leve.

Los valores atípicos extremos son observaciones más allá de los siguientes límites externos  $\langle LI_e, LS_e \rangle$ :

$$LI: < Q_1 - 3 \times IQR, \text{ ó } LS: > Q_3 + 3 \times IQR. \quad (28)$$

Para el caso de datos meteorológicos, los valores atípicos leves no serán removidos debido a que las variaciones de las variables respuesta pueden presentar algún valor fuera de rango establecido, ocasionado por algún fenómeno físico particular; también se debe considerar que algunas variables meteorológicas, en ciertos casos, presentan grandes fluctuaciones, por

---

<sup>5</sup> Manual sugerido en <http://metaemotion.com/diego.garcia.morate/download/weka.pdf>



lo cual se consideró sólo los valores atípicos leves mas no los valores atípicos extremos.

Se desarrolló y analizó la climatología de la estación meteorológica aeronáutica del aeropuerto de Trujillo teniendo en cuenta lo establecido en la RAP 303, capítulo 8: Información Climatológica Aeronáutica, donde indica que: “la información climatológica aeronáutica se debe basar en observaciones efectuadas a lo largo de un período de cinco años como mínimo, y dicho período se debe indicar en la información proporcionada”. En este caso se utilizó un periodo diez años (2008 – 2017) para analizar la climatología de las variables de temperaturas extremas y de un periodo de cinco años (2013 – 2017) para analizar la climatología de la variable de visibilidad horizontal; y así también pudiendo determinar un contexto climatológico del lugar al analizar otras variables en un periodo de diez años de la temperatura media del aire, y un periodo de cinco años para otras variables como la presión atmosférica, techo de nubes, dirección y velocidad del viento prevaleciente y máximo, humedad relativa y precipitación. También se estudiaron los sistemas meteorológicos que influyen en el área de estudio para específicamente el caso de la reducción de visibilidad horizontal por niebla, identificando los patrones sinópticos más característicos mediante revisión bibliográfica.

### 3.2.2. Identificación de las variables de entrada para la red neuronal del pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas

Se utilizaron cinco años de datos horarios de visibilidad horizontal (2013-2017) y diez años de datos diarios de temperaturas máximas y mínimas (2008-2017) de la estación meteorológica aeronáutica del aeropuerto de Trujillo. Como variables explicativas se tomaron en cuenta un conjunto de índices globales y otras variables meteorológicas tomadas de la estación meteorológica aeronáutica del aeropuerto de Trujillo, analizando cuales tendrán mejor respuesta en nuestro modelo de pronóstico y así pudiendo identificar cuales tienen una mejor relación con nuestras variables meteorológicas en la zona de interés.

### 3.2.3. Determinación de los algoritmos apropiados para el entrenamiento y aprendizaje de la red neuronal en la arquitectura seleccionada para el pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas

Mediante análisis documentario de otros estudios anteriores de aplicaciones de redes neuronales para distintas variables meteorológicas en otras zonas de estudio se determinó la arquitectura más adecuada para el tipo de red neuronal enfocada en pronóstico meteorológico. Primero se identificaron las capas o niveles de la Red Neuronal: Capa de

entrada, aquella que recibe directamente la información de las fuentes externas de la red; Capa oculta, internas en la red y no tienen contacto directo con el exterior, éstas se encuentran entre cero y un número elevado por lo cual se evaluó el valor más óptimo; y Capa de salida, que transportan la información de la red hacia el exterior.

Esto se desarrolló en fases, siendo una de ellas la de aprendizaje, el cual consiste en la adaptación de los pesos sinápticos a la información que extrae de los patrones o procesos de entrenamientos, logrando que dichos pesos se minimicen o maximicen de acuerdo a alguna función de energía. Y la fase de pruebas, en donde los resultados que brinda la fase anterior pueden ser calculados de una vez o adaptarlos de una forma iterativa, dependiendo del tipo de red neuronal. Esto ayudó a tener la mejor red neuronal que explique los datos de visibilidad horizontal y temperaturas máximas y mínimas para la estación meteorológica aeronáutica del aeropuerto de Trujillo mediante el programa estadístico WEKA que trabaja con la aplicación del algoritmo de retropropagación de la siguiente manera:

- I. Empezar con valores pequeños aleatorios en los pesos de la red.
- II. Presenta un patrón de entrada  $X_p: x_{p1}, x_{p2}, \dots, x_{pN}$ , y especifica la salida deseada que debe generar la red  $d_1, d_2, \dots, d_M$ .
- III. Calcula la salida actual de la red, para ello se presenta las entradas a la red y se calcula la salida que presenta cada capa hasta llegar a la capa de salida, esta será la salida de la red  $y_1, y_2, \dots, y_M$ . Los pasos son los siguientes:
  - Calcular las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada. Para una neurona  $j$  oculta se utiliza la ecuación 8 mencionada en la revisión de literatura, en donde el índice  $h$  se refiere a magnitudes de la capa oculta; el subíndice  $p$ , al  $p$ -ésimo vector de entrenamiento, y  $j$  a la  $j$ -ésima neurona oculta. El término  $\theta$  es opcional ya que actúa como una entrada más.
  - Se calculan las salidas de las neuronas ocultas mediante la ecuación 9 mencionada en la revisión de literatura
  - Se realiza los mismos cálculos para obtener las salidas de las neuronas de salida (capa  $o$ : output) con las ecuaciones 10 y 11 mencionadas en la revisión de literatura
- IV. Obtener los términos de error para todas las neuronas. Si la neurona  $k$  es una neurona de la capa de salida, el valor de la delta es calculado mediante la ecuación 12 mencionada en la revisión de literatura.

La función  $f$  debe ser derivable, lo que implica la imposibilidad de usar una función escalar. En general se disponen de dos formas de función de salida que nos pueden servir: la función lineal de salida ( $f|k(net_{jk}) = net_{jk}$ ) y la función sigmoideal definida por la ecuación 13 mencionada en la revisión de literatura.

La selección de la función de salida depende de la forma en la que se decida representar los datos de salida: si se desea que las neuronas de salida sean binarias, se utiliza la función sigmoideal, puesto que esta función es casi estable y, además es derivable. En otros casos es tan aplicable una función como otra.

Para la función lineal, se tiene:  $f'_k = 1$ , mientras que la derivada de una función  $f$  sigmoideal es representada en la ecuación 14. Por lo que los términos de error para las neuronas de salida nos quedan, para el caso de la salida lineal expresados en la ecuación 15, y para la salida sigmoideal son expresados en la ecuación 16 mencionada en la revisión de literatura

Si la neurona  $j$  no es de salida, entonces la derivada parcial del error no debe ser evaluada directamente. Entonces, se obtiene el desarrollo a partir de valores conocidos y otros que pueden ser evaluados. La expresión obtenida en este caso es presentada en la ecuación 17 mencionada en la revisión de literatura, donde  $k$  se refiere a todas las neuronas de la capa superior a la de la neurona  $j$ . Así, el error que se obtiene en una neurona oculta es proporcional a la suma de los errores conocidos que resultan de las neuronas a las que está conectada la salida de esta, multiplicado cada uno de ellos por el peso de la conexión. Los umbrales internos de las neuronas se adaptan de forma similar, considerando que están conectados con los pesos desde entradas auxiliares de valor constante.

- V. Actualización de los pesos: para ello, se utiliza el algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la forma siguiente: para las neuronas de la capa de salida se tiene sus pesos expresados en la ecuación 18 y para los pesos de las neuronas de la capa oculta se tiene la ecuación 19 mencionada en la revisión de literatura. En ambos casos, para acelerar el proceso de aprendizaje, se puede añadir un término momento de valor:  $\beta (w_{ji}^h(t) - w_{ji}^h(t - 1))$  cuando se trata de una neurona oculta.

VI. El proceso se repite hasta que el termino de error resulta aceptablemente pequeño para cada uno de los patrones aprendidos. Esto se expresa en la ecuación 20 mencionada en la revisión de literatura.

#### 3.2.4. Validación del modelo que optimice el pronóstico, con menor grado de error, de las variables meteorológicas Visibilidad Horizontal y Temperaturas Máximas y Mínimas

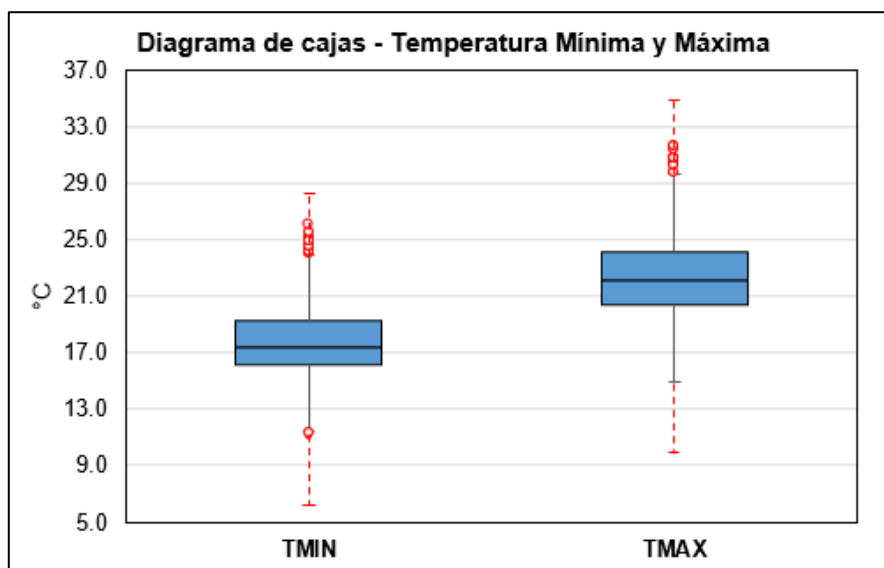
Se evaluó el rendimiento del modelo seleccionado preliminarmente, realizado mediante redes neuronales artificiales de retropropagación para cada una de las variables de visibilidad horizontal y temperaturas máximas y mínimas en el aeropuerto de Trujillo, hallando los indicadores de Error Cuadrático Medio (ECM) expresado en la ecuación 21, Raíz del Error Cuadrático Medio (RECM) expresada en la ecuación 22, Error Absoluto Medio (EAM) expresado en la ecuación 23, Error Absoluto Medio Normalizado (EAMN) expresado en la ecuación 24, BIAS expresado en la ecuación 25, y la Correlación expresada en la ecuación 26 y 27 mencionadas en la revisión de literatura. El modelo se reconoció como óptimo cuando el valor del ECM, RECM, EAM y EAMN fue bajo, cuando el valor del BIAS fue cercano a cero y cuando la Correlación tuvo un valor cercano a uno y fue positiva.

Finalmente, se presentaron los resultados del modelo de pronóstico de la visibilidad horizontal y temperaturas máximas y mínimas en el aeropuerto de Trujillo, mediante una comparación entre los datos reales obtenidos en la estación aeronáutica y los valores resultantes de la red neuronal, haciendo una descripción de estos.

## IV. RESULTADOS Y DISCUSIÓN

### 4.1. Análisis e identificación de las condiciones climatológicas y sistemas meteorológicos de influencia que predominan sobre el área de estudio.

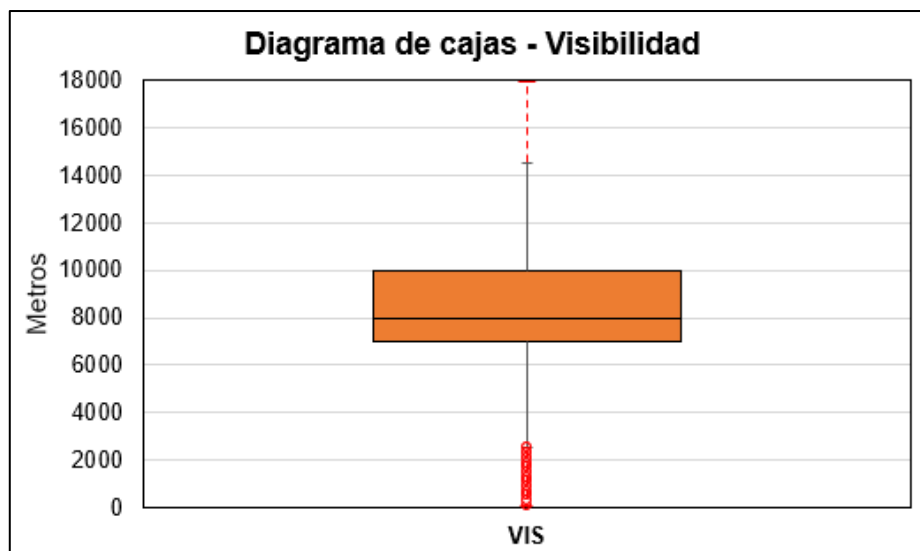
A partir del desarrollo del análisis exploratorio de los datos de la temperatura mínima, podemos ver en su diagrama de cajas (Figura 6) que tienen una asimetría positiva o segada a la derecha, teniendo que los datos se concentran en la parte inferior de la distribución, entonces la media es mayor que la mediana; no se encontraron valores atípicos extremos, debido a esto no se eliminó ninguna observación a pesar de que hay 14 valores atípicos leves. Respecto a los datos de la temperatura máxima (Figura 6), podemos ver que tienen una asimetría positiva o segada a la derecha muy pequeña, teniendo que los datos se concentran en la parte inferior de la distribución, entonces la media es mayor que la mediana; tampoco se encontraron valores atípicos extremos, debido a esto no se eliminó ninguna observación, pero se presentaron nueve valores atípicos leves.



**Figura 6.** Diagrama de cajas para la temperatura mínima y máxima del Aeropuerto de Trujillo.

Para la variable de visibilidad horizontal, podemos ver que tiene una asimetría negativa, teniendo que la media es menor que la mediana; sin embargo, se tienen que los datos se

concentran en la parte inferior de la distribución debido a la distorsión provocada por valores muy grandes con un límite de medición de 9999 metros por la codificación METAR de los datos proporcionados. Tampoco se encontraron valores atípicos extremos, debido a esto no se eliminó ninguna observación, pero se presentaron 13 valores atípicos leves.



**Figura 7.** Diagrama de cajas para la visibilidad horizontal del aeropuerto de Trujillo.

A continuación, se presentan las estadísticas descriptivas de la temperatura mínima, temperatura máxima y visibilidad, en la Tabla 1.

**Tabla 1:** Estadísticas descriptivas para el Aeropuerto de Trujillo.

	Mínimo	Máximo	Mediana	Media	Error Estándar de la Media	Desviación Estándar	Asimetría
Temp. Mínima (°C)	11.10	26.00	17.40	17.76	0.36	2.15	0.58
Temp. Máxima (°C)	15.90	31.60	22.10	22.30	0.04	2.59	0.35
Visibilidad (m)	100.00	9999.00	8000.00	7981.37	9.59	2007.63	-0.86

A continuación, se muestra la climatología desarrollada para el aeropuerto de Trujillo con los datos proporcionados de la estación meteorológica aeronáutica de CORPAC S.A, para los años del 2008 al 2017 para las variables de temperaturas medias del aire y para los años del 2013 al 2017 para las variables visibilidad, viento prevaleciente y máximo, presión atmosférica, humedad relativa, precipitación y techo de nubes.

A partir de la Figura 8 podemos ver que la temperatura del aire media mensual está entre los valores de 22.6 y 17.4°C, la temperatura máxima del aire media mensual está entre los valores de 25.5 y 20.1°C y la temperatura mínima del aire media mensual está entre los

valores de 15.9 y 20.4°C. En conjunto vemos que los valores más altos se dan en la estación de verano, durante los meses de febrero y marzo, y los valores más bajos se dan durante la transición de la estación de invierno a primavera, durante los meses de setiembre a octubre.

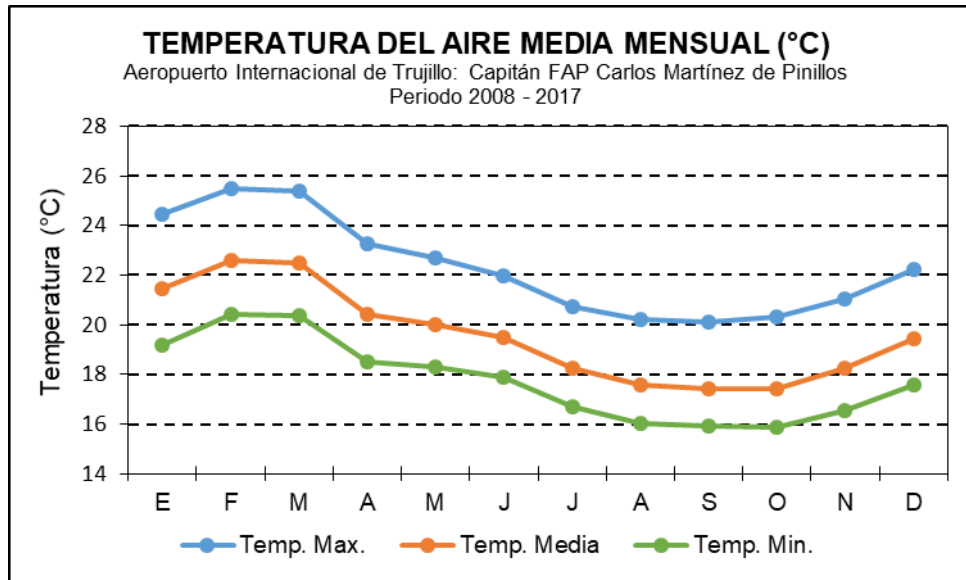


Figura 8. Resumen climatológico de la Temperatura del Aire del Aeropuerto de Trujillo.

En la Figura 9 podemos ver que la visibilidad durante la estación de verano presenta una mayor frecuencia de valores entre 5000 a 1000 metros en horas de la madrugada y mañana, y una mayor frecuencia de valores menores a 1000 metros en horas del amanecer.

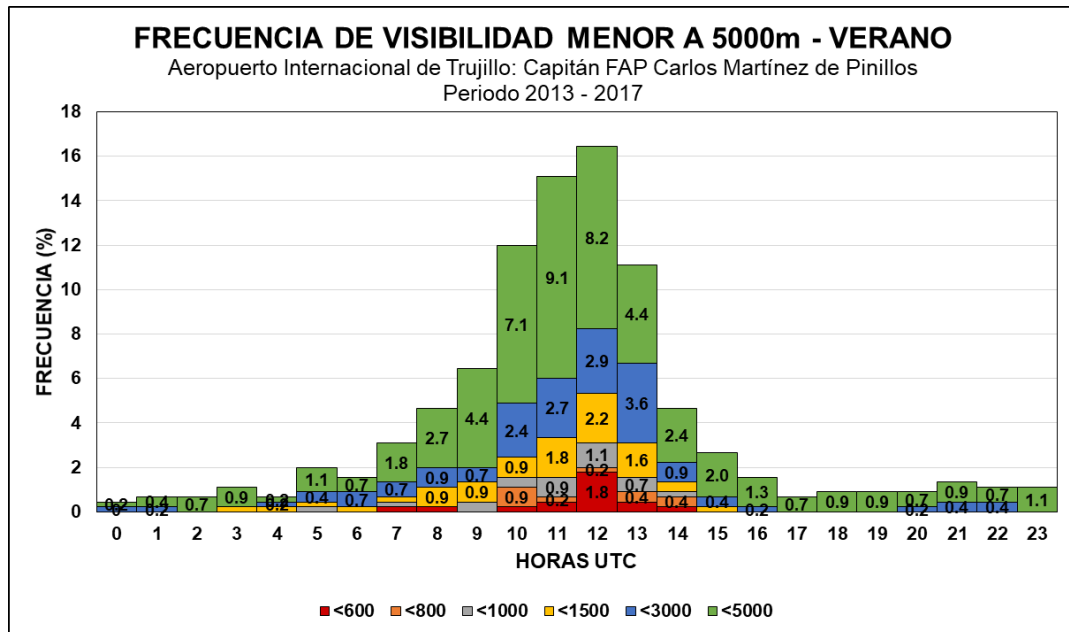
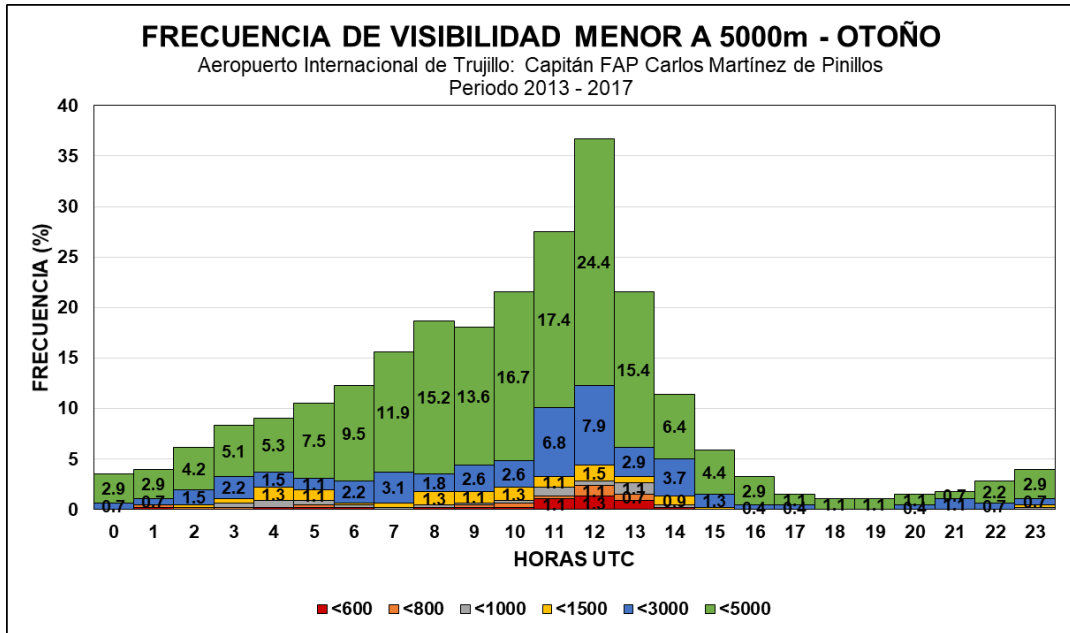


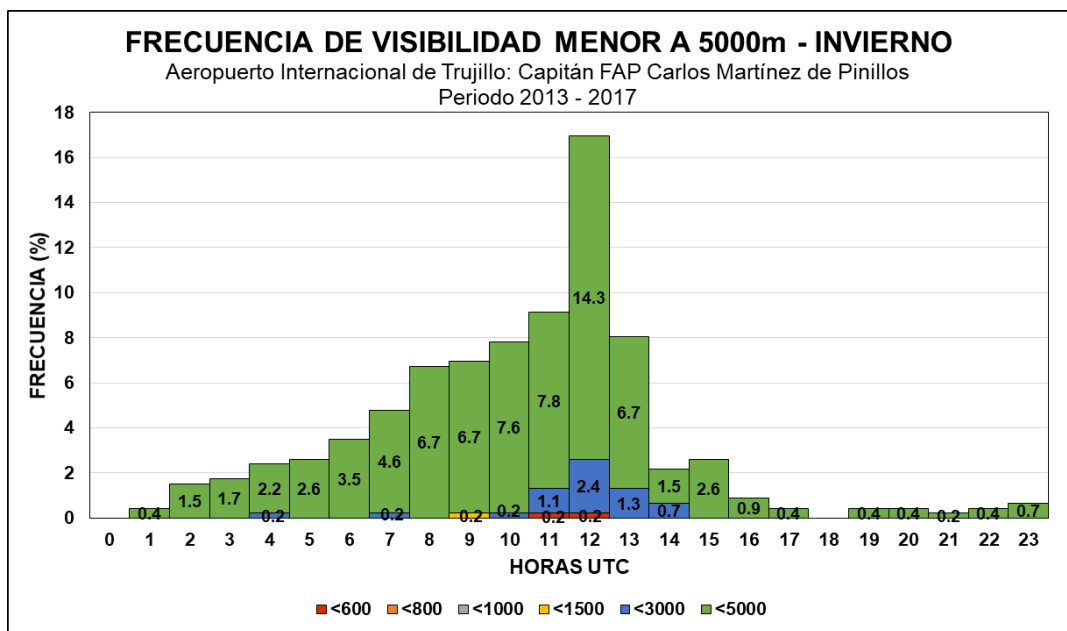
Figura 9. Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en verano.



**Figura 10.** Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en otoño.

En la Figura 10 podemos ver que la visibilidad durante la estación de otoño presenta una mayor frecuencia de valores entre 5000 a 1000 metros en últimas horas de la noche, madrugada y primeras horas de la mañana, y una mayor frecuencia de valores menores a 1000 metros en horas del amanecer.

En la Figura 11 podemos ver que durante la estación de invierno se presenta una mayor frecuencia de valores entre 5000 a 1000 metros de visibilidad en últimas horas de la noche, madrugada y al amanecer. En esta estación se presentan escasos días con niebla.



**Figura 11.** Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en invierno.



En la Figura 12 podemos ver que la visibilidad durante la estación de primavera presenta una mayor frecuencia de valores entre 5000 a 1000 metros en horas de la madrugada y primeras horas de la mañana, y una mayor frecuencia de valores menores a 1000 metros en horas del amanecer.

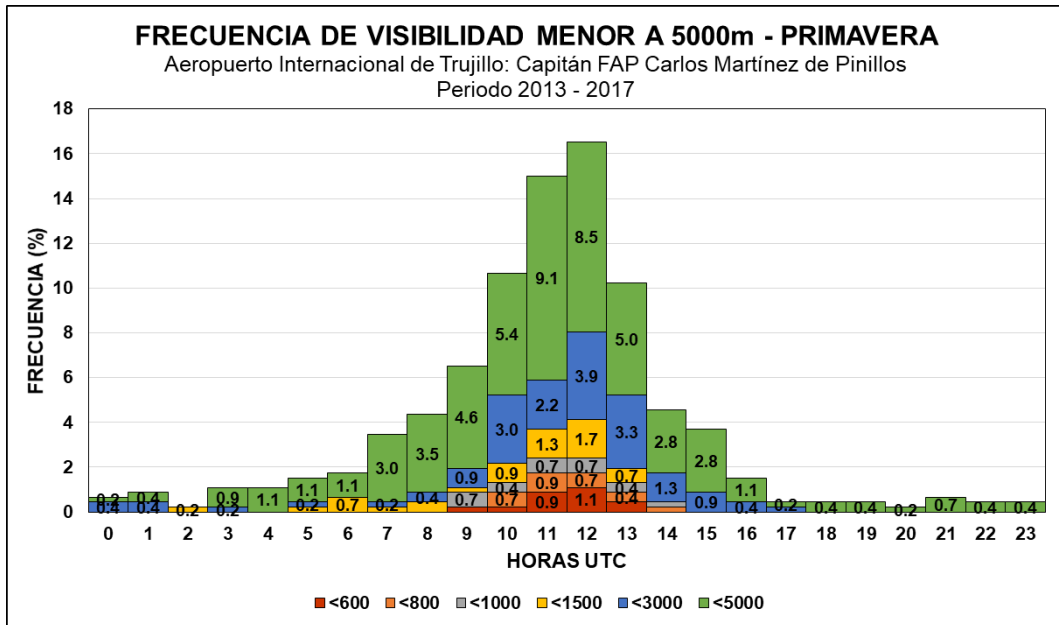


Figura 12. Resumen climatológico de la Visibilidad del Aeropuerto de Trujillo en primavera.

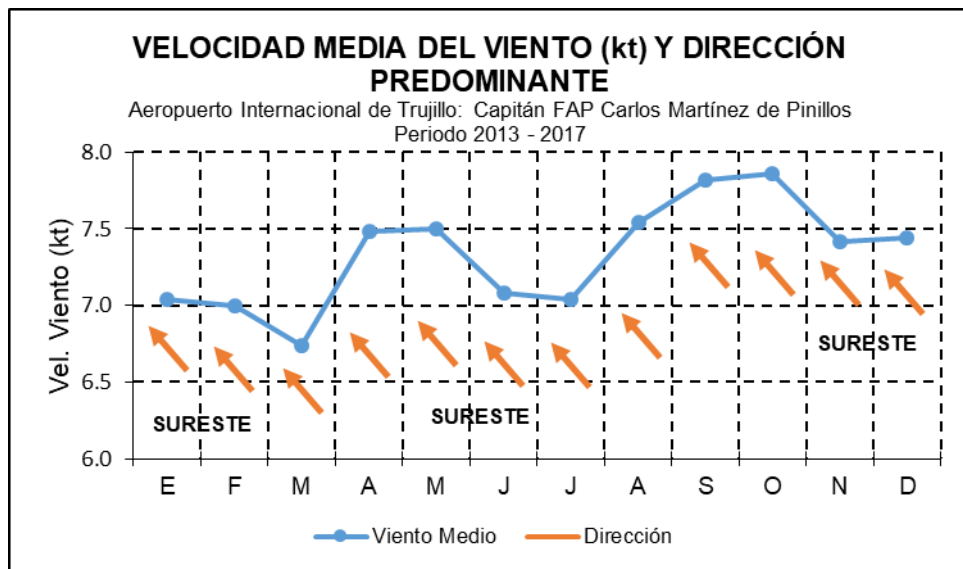
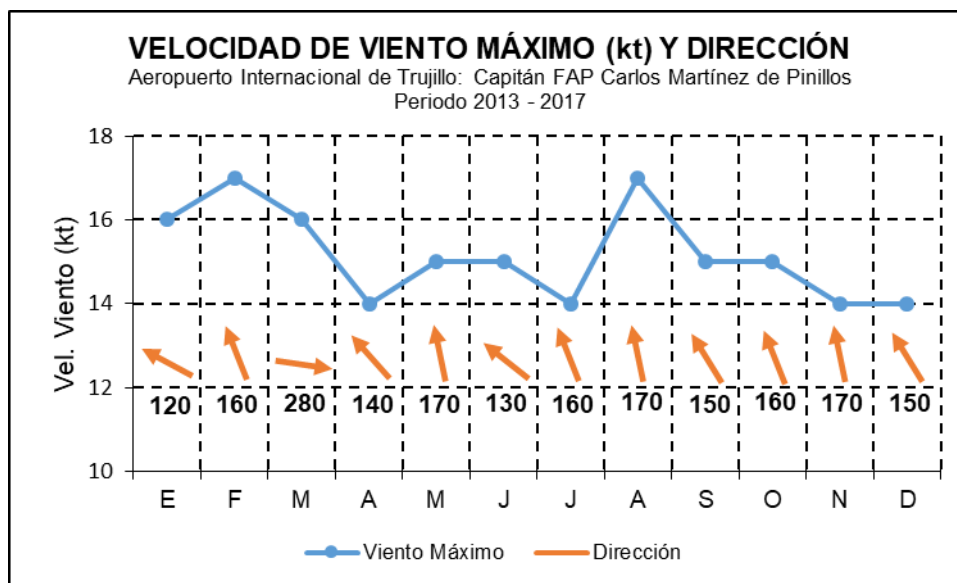


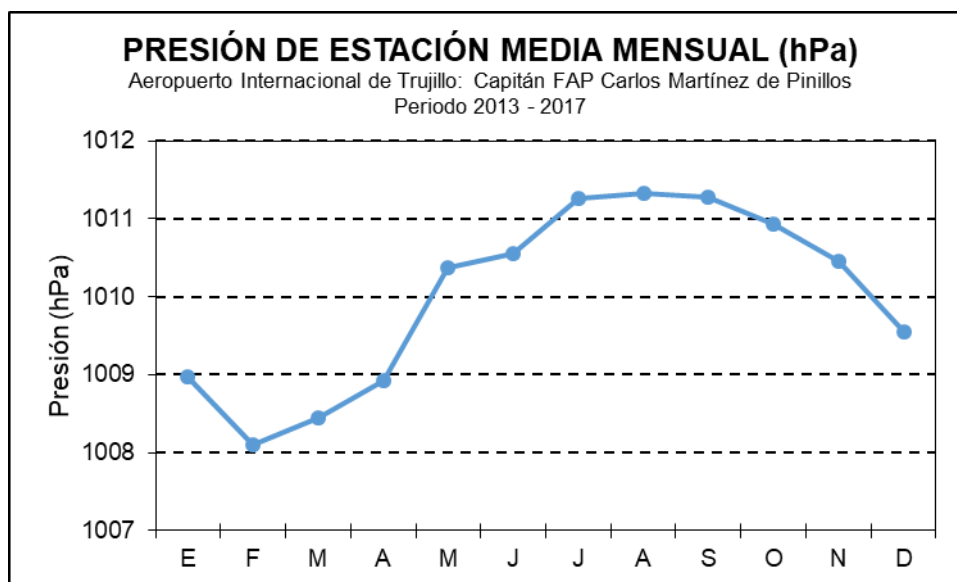
Figura 13. Resumen climatológico del Viento del Aeropuerto de Trujillo.

A partir de la Figura 13 podemos ver que la velocidad media del viento mensual está entre los valores de 6.7 y 7.9 nudos, y la dirección del viento predominante es del Sureste. En conjunto vemos que los valores más altos se dan durante la transición de la estación de invierno a primavera, en los meses de setiembre a octubre; y los valores más bajos se dan en la estación de verano, durante los meses de febrero y marzo.



**Figura 14.** Resumen climatológico del Viento Máximo del Aeropuerto de Trujillo.

A partir de la Figura 14 podemos ver que la velocidad máxima del viento media mensual está entre los valores de 14 y 17 nudos, y la dirección del viento es variable teniendo cierta predominancia del Sureste y Sur. En conjunto vemos que los valores más altos se dan durante la estación de verano e invierno, y los valores más bajos se dan en la estación de otoño y primavera.



**Figura 15.** Resumen climatológico de la presión atmosférica del Aeropuerto de Trujillo.

A partir de la Figura 15 podemos ver que la presión atmosférica media mensual está entre los valores de 1008.1 y 1011.3 hPa. Los valores más altos se dan durante la estación de invierno, y los valores más bajos se dan en la estación de verano.

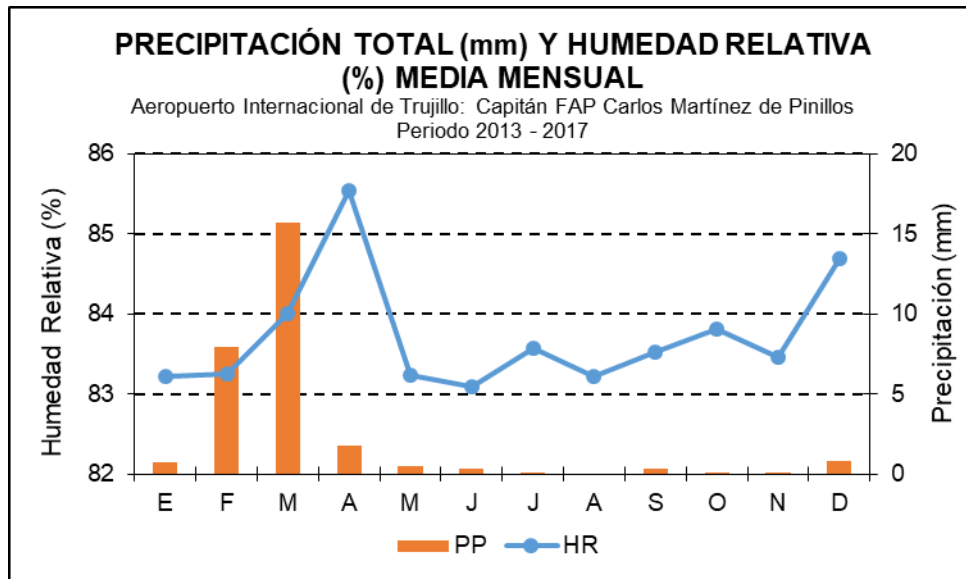


Figura 16. Resumen climatológico de la precipitación y humedad relativa del Aeropuerto de Trujillo.

A partir de la Figura 16 podemos ver que la precipitación total media mensual está entre los valores de 0.0 y 15.7 mm, teniendo los valores más altos en los meses de verano y siendo menor el resto del año. La humedad relativa media mensual está entre los valores de 83 y 86%, teniendo los valores más altos en otoño, en el mes de abril, y siendo menor el resto del año.

En la Figura 17 podemos ver que la nubosidad durante la estación de verano presenta una mayor frecuencia de altura de techo de nubes inferior a 450 metros con presencia de casos menores a 60 metros, durante la madrugada y primeras horas de la mañana.

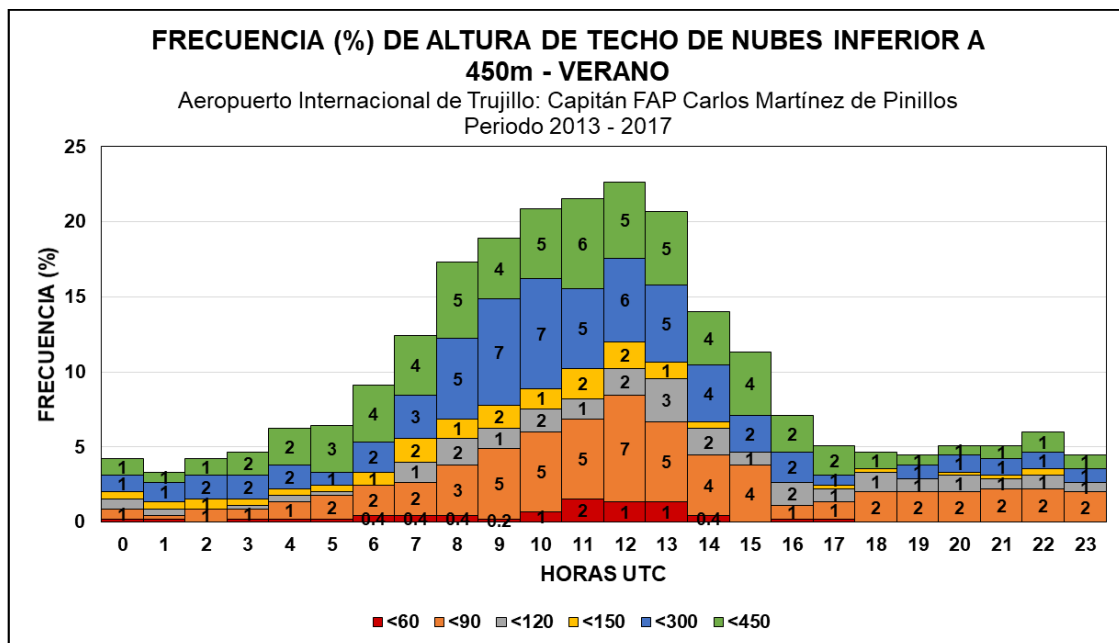


Figura 17. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en verano.

En la Figura 18 podemos ver que la nubosidad durante la estación de otoño presenta una mayor frecuencia de altura de techo de nubes inferior a 450 metros con presencia de casos menores a 60 metros, durante las últimas horas de la noche, madrugada y primeras horas de la mañana.

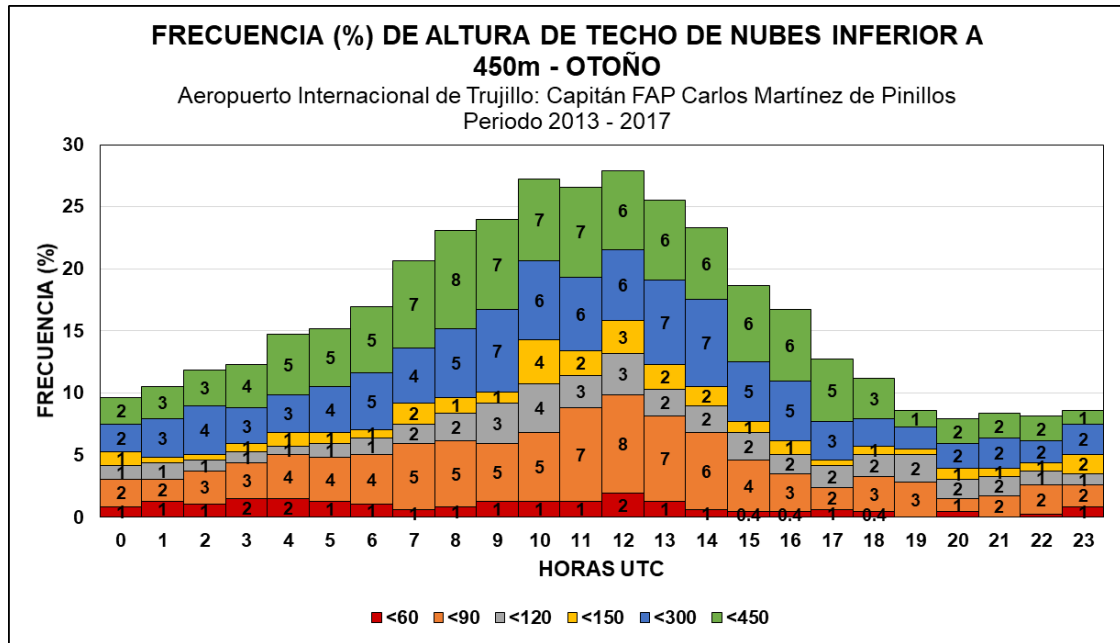


Figura 18. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en otoño.

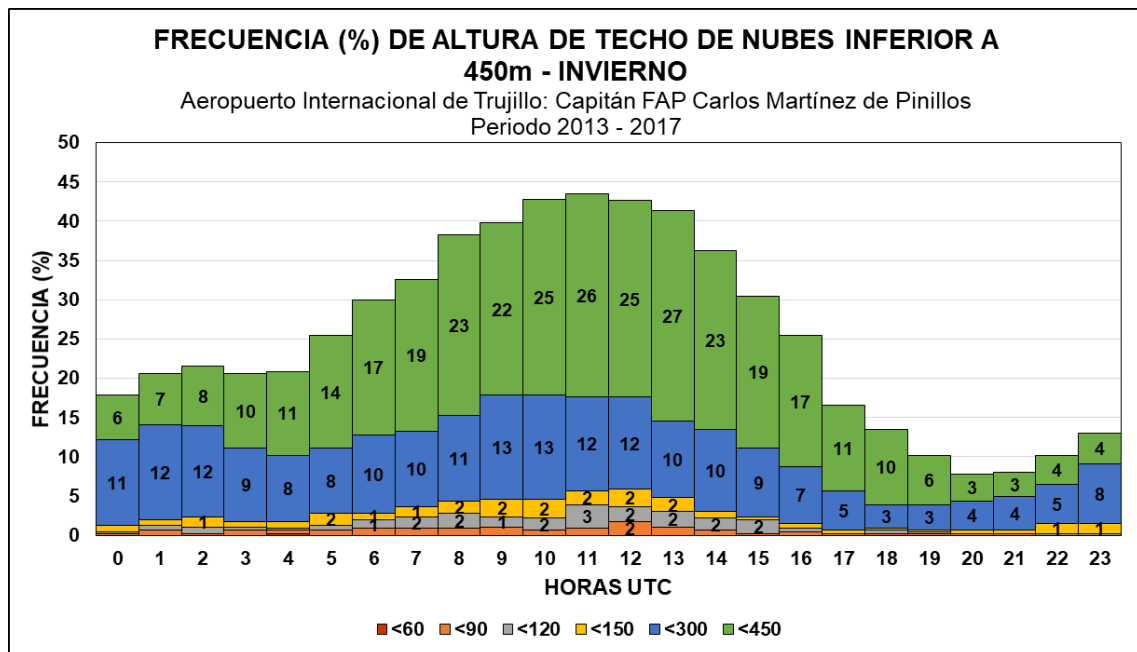
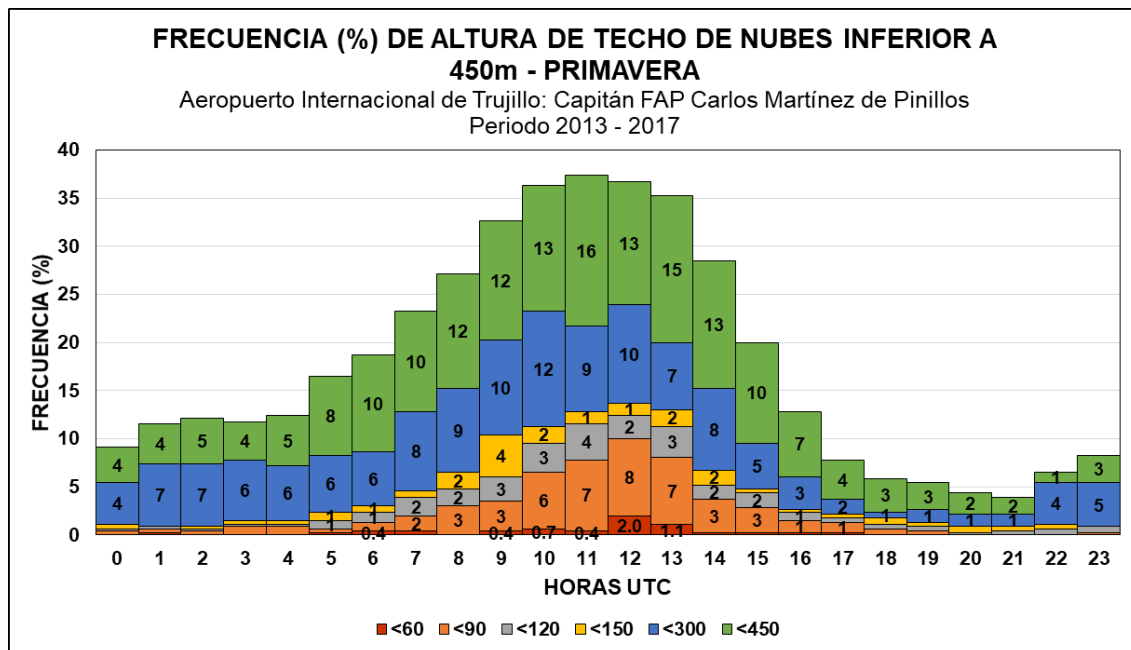


Figura 19. Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en invierno.

En la Figura 19 podemos ver que la nubosidad durante la estación de invierno presenta una mayor frecuencia de altura de techo de nubes inferior a 450 metros hasta los 60 metros, durante la noche, madrugada y primeras horas de la mañana.



**Figura 20.** Resumen climatológico de la Nubosidad del Aeropuerto de Trujillo en primavera.

En la Figura 20 podemos ver que la nubosidad durante la estación de primavera presenta una mayor frecuencia de altura de techo de nubes inferior a 450 metros con presencia de casos menores a 60 metros, durante la madrugada y primeras horas de la mañana.

Mediante revisión de literatura, podemos analizar que, para el aeropuerto de Trujillo, por ubicarse en una ciudad costera, se debe tomar en cuenta al componente oceánico como principal fuente de humedad. Por lo tanto, los sistemas que influyen en la zona serán el mar peruano y la masa de agua subtropical superficial evidenciadas en el estado térmico del agua superficial del mar, y el Anticiclón del Pacífico sur cuya principal influencia se evidencia en los gradientes de presión que se generan por su fluctuación, originando el viento que interactúa cinemáticamente con las masas de aire desarrollando diversas condiciones del tiempo en la zona.

Para la visibilidad horizontal es importante señalar los factores que causan su disminución para así poder tomarlo en cuenta en el desarrollo de la red neuronal artificial. Una de las causas principales de su disminución es la niebla, donde para la zona de Trujillo es importante tener en cuenta los procesos físicos de la generación de la niebla en la costa peruana en las interfaces océano - atmósfera, océano - costa o la costa - atmósfera, de los cuales se destacan la niebla advectiva y la niebla de irradiación mencionadas en la revisión de literatura. Al analizar los factores de formación de nieblas en la costa peruana se tiene un panorama más claro de las variables a tomar en cuenta en el desarrollo del modelo de redes neuronales en el punto a continuación.

#### **4.2. Identificación de las variables de entrada para la red neuronal del pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas**

Para empezar con el desarrollo de la red neuronal, se tomaron en cuenta todos los valores de la serie de diez años de datos diarios de la temperatura máxima y mínima de los años 2008 al 2017, y todos los valores de la serie de cinco años de datos horarios de la visibilidad horizontal de los años 2013 al 2017, del aeropuerto de Trujillo, habiendo realizado previamente el AED. Posteriormente se definieron sus variables explicativas teniendo las siguientes: los índices globales de la temperatura superficial del mar (TSM) en el pacífico ecuatorial en las regiones Niño 1+2, Niño 3, Niño 4 y Niño 3.4, la presión al nivel del mar en Darwin y Tahití, y el Índice de Oscilación del Sur (SOI), además para las variables temperatura máxima y mínima del aire se consideró a la humedad relativa diaria de los años 2008 al 2017 y para la variable visibilidad horizontal se consideró la humedad relativa y viento horario de los años 2013 al 2017, de la estación meteorológica del aeropuerto de Trujillo.

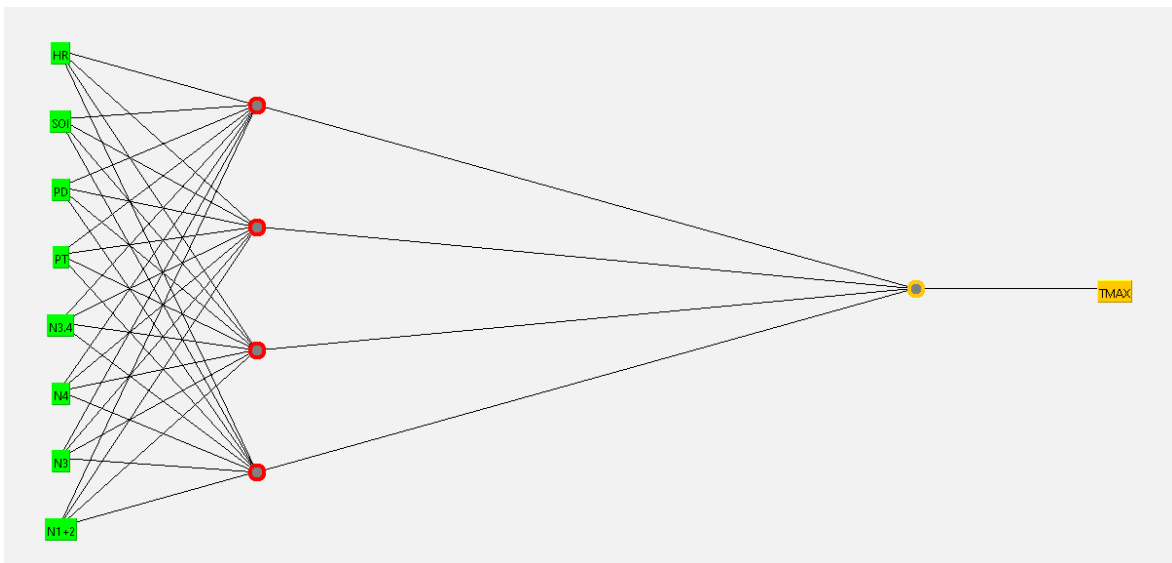
Se tomaron en cuenta estos índices globales mencionados líneas arriba debido a que se encuentran ubicados en la región geográfica del océano pacífico que es de influencia directa con nuestra área de interés del aeropuerto de Trujillo que está ubicada dentro del litoral costero peruano (Anexo 3). Los índices globales de la TSM en el pacífico ecuatorial en las regiones Niño 3, Niño 3.4, Niño 4 y Niño 1+2, nos dan una perspectiva de las condiciones de las temperaturas superficiales del mar, donde según los valores que presenten, reflejarán el intercambio de vapor de agua a la atmósfera desde la superficie del océano; por otro lado la presión al nivel del mar en Darwin y Tahití, y el Índice de Oscilación del Sur (SOI), nos dan una perspectiva de las fluctuaciones de los sistemas de presión en el océano pacífico asociados a patrones de vientos, donde los valores del SOI están relacionados a los fenómenos climáticos de El Niño y La Niña. Posteriormente, respecto a las variables locales, se consideró a la humedad relativa debido a que esta variable nos ayudará a agregarle una característica propia del lugar a la red neuronal, además de que nos permite conocer el estado de saturación de la atmósfera, y por último para la visibilidad horizontal se consideró también a la variable viento como explicativa debido a que esta nos ayudará mejor a identificar las condiciones del medio atmosférico donde se originan las reducciones de visibilidad importantes para el pronóstico del tiempo.

### 4.3. Determinación de los algoritmos apropiados para el entrenamiento y aprendizaje de la red neuronal en la arquitectura seleccionada para el pronóstico de Visibilidad Horizontal y Temperaturas Máximas y Mínimas

El análisis de la red neuronal se realizó utilizando el software WEKA 3.8, para la cual los datos deben ingresarse en un archivo con extensión *arff*. Una vez subidos los archivos se procedió con la clasificación, eligiendo en este caso *Backpropagation* que se basa en una red neuronal tipo perceptrón multicapas que optimiza tanto a la matriz de pesos como el número de neuronas ocultas. Se desarrolló una red neuronal de tres capas: una capa de entrada, una capa oculta y una capa de salida, escogiendo el total de datos para la fase de entrenamiento para su evaluación. Durante la fase de entrenamiento de la red, se determinan iterativamente los pesos de las conexiones, que conectan a las neuronas de entrada con las de la capa oculta y los pesos para la transferencia entre la capa oculta y la de salida. Se diseñaron varias pruebas de redes neuronales escogiendo diferentes números de neuronas en la capa oculta para así obtener el mejor modelo de la red neuronal para nuestro modelo.

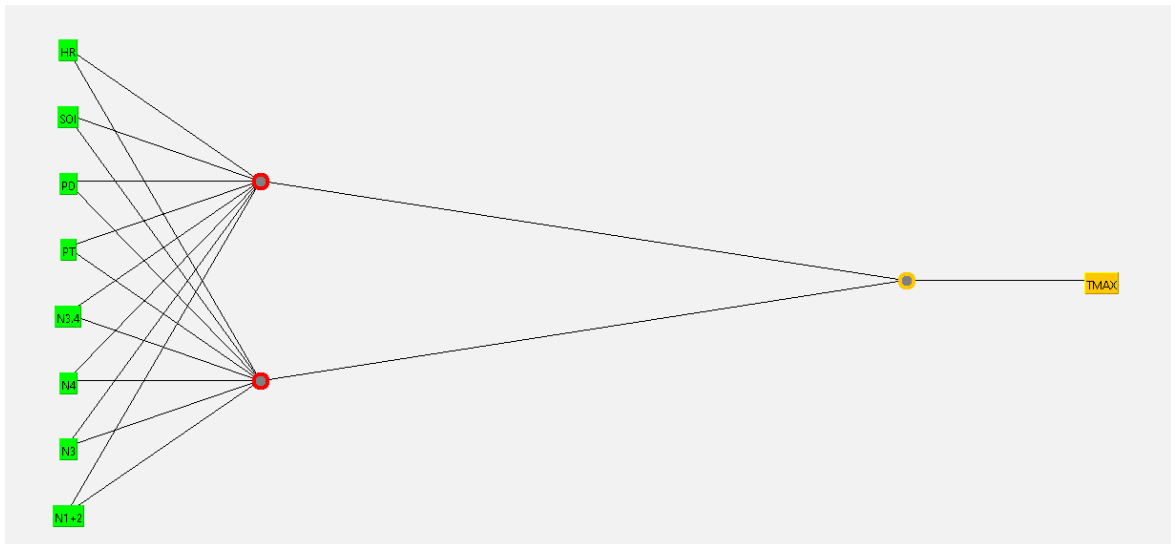
#### 4.3.1. Temperatura Máxima

Como se puede ver en la Figura 21, se empezó diseñando una red neuronal de ocho neuronas en la capa de entrada, una capa oculta de cuatro neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se escogió fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1.54°C.



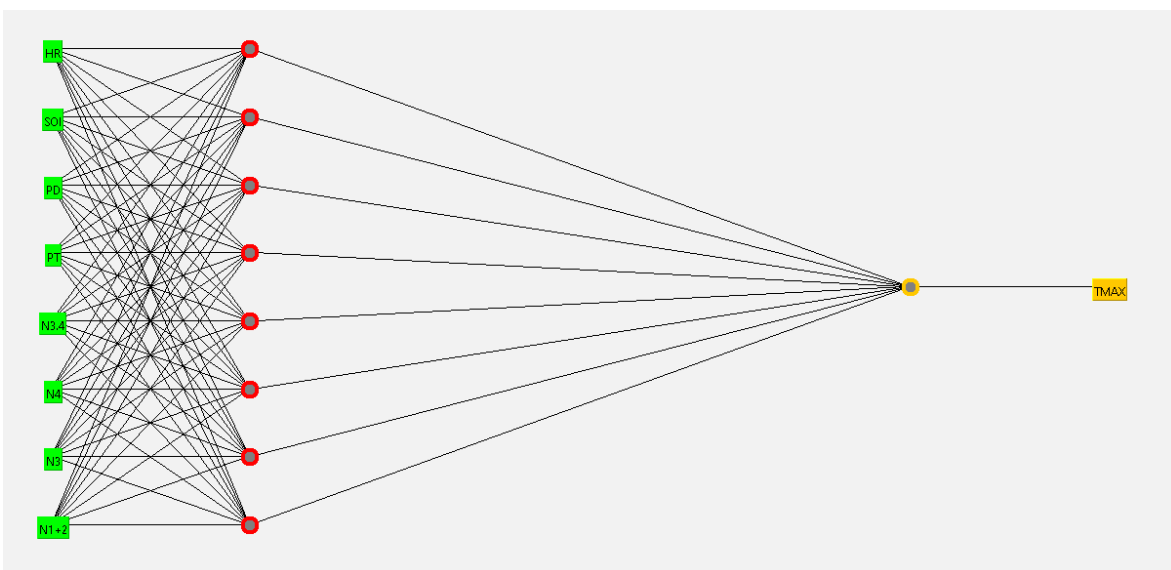
**Figura 21.** Diseño de la red neuronal de cuatro neuronas ocultas para la Temperatura Máxima.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

En la Figura 22, se tiene un diseño una red neuronal de ocho neuronas en la capa de entrada, una capa oculta de dos neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se utilizó fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1.37°C.



**Figura 22.** Diseño de la red neuronal de dos neuronas ocultas para la Temperatura Máxima.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

Por último, en la Figura 23 se muestra el diseño una red neuronal de ocho neuronas en la capa de entrada, una capa oculta de ocho neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se escogió fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1.17°C.



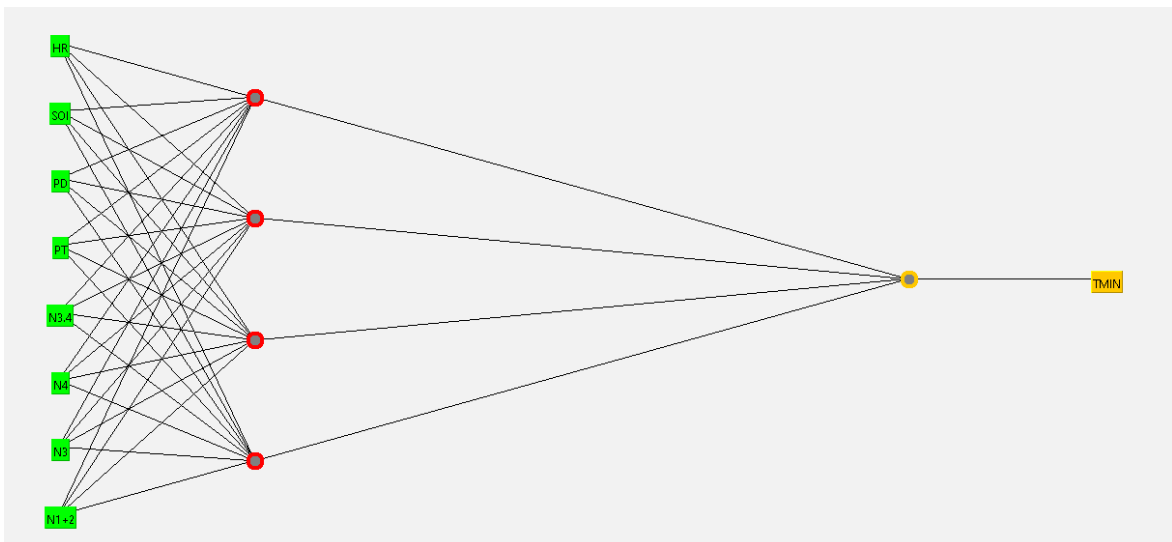
**Figura 23.** Diseño de la red neuronal de ocho neuronas ocultas para la Temperatura Máxima.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.



De los tres diseños mostrados se puede concluir que a más número de neuronas en la capa oculta el error disminuye. Por lo tanto, cada vez que se va realizando una clasificación, en un orden diferente, los pesos de las conexiones aumentan la velocidad durante el proceso, ocasionando un lento descenso del error producido en el nivel de salida, por lo tanto, la red neuronal presenta un buen comportamiento, debido al aprendizaje que tiene y a la evaluación del error que comete la red. Finalmente, el diseño seleccionado será el de ocho neuronas en la capa oculta.

#### 4.3.2. Temperatura Mínima

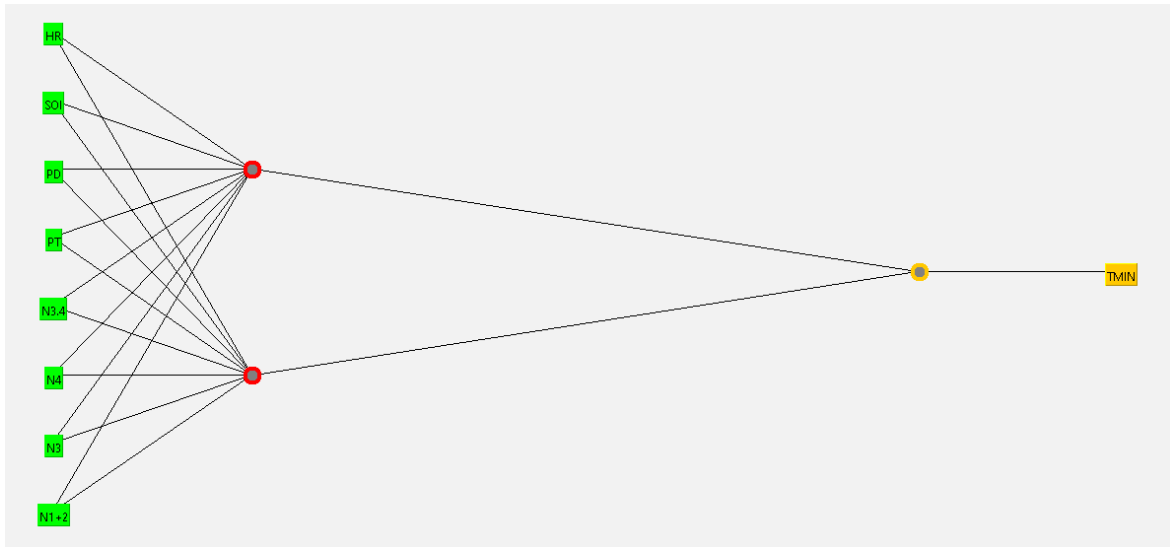
Como se ve en la Figura 24, se empezó diseñando una red neuronal de ocho neuronas en la capa de entrada, una capa oculta de cuatro neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se utilizó fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1.02°C.



**Figura 24.** Diseño de la red neuronal de cuatro neuronas ocultas para la Temperatura Mínima.

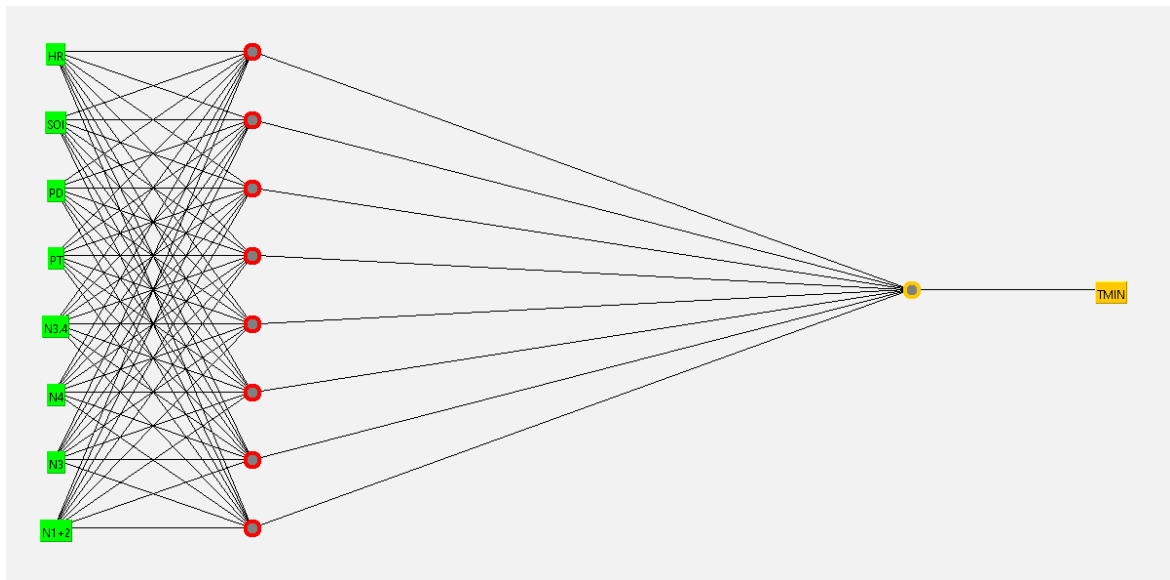
**FUENTE:** Elaboración mediante programa WEKA 3.8.

Se diseñó una red neuronal de ocho neuronas en la capa de entrada, una capa oculta de dos neuronas y una capa de salida de una neurona; mostrado en la Figura 25. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se utilizó fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1.25°C.



**Figura 25.** Diseño de la red neuronal de dos neuronas ocultas para la Temperatura Mínima.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

Por último, en la Figura 26 se muestra el diseño de una red neuronal de ocho neuronas en la capa de entrada, una capa oculta de ocho neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se escogió fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 0.99°C.



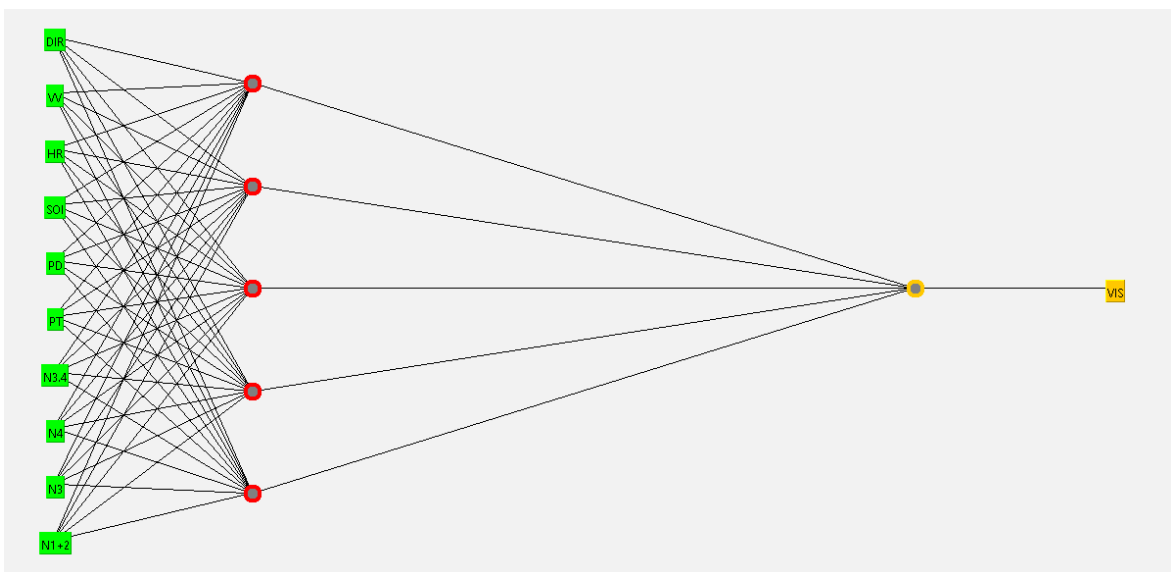
**Figura 26.** Diseño de la red neuronal de ocho neuronas ocultas para la Temperatura Mínima.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

De los tres diseños mostrados se puede concluir que a más número de neuronas en la capa oculta el error disminuye. Por lo tanto, cada vez que se va realizando una clasificación, en un orden diferente, los pesos de las conexiones aumentan la velocidad durante el proceso, ocasionando un lento descenso del error producido en el nivel de salida, por lo tanto, la red

neuronal presenta un buen comportamiento, debido al aprendizaje que tiene y a la evaluación del error que comete la red. Finalmente, el diseño seleccionado será también el de ocho neuronas en la capa oculta.

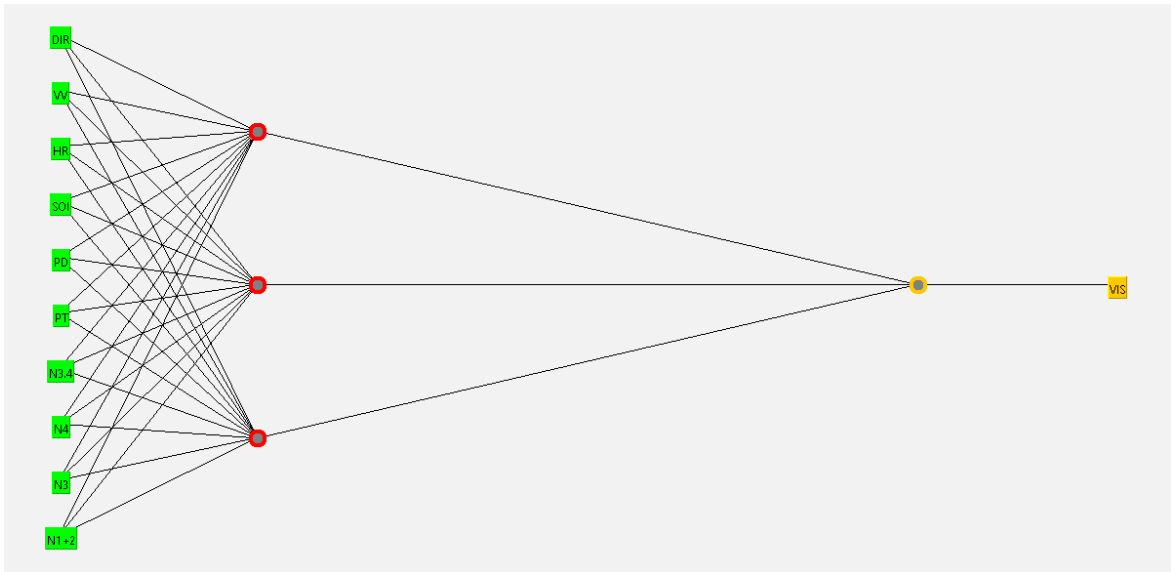
#### 4.3.3. Visibilidad Horizontal

En la Figura 27 se muestra el primer diseño de una red neuronal de diez neuronas en la capa de entrada, una capa oculta de cinco neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se escogió fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1690.07m.



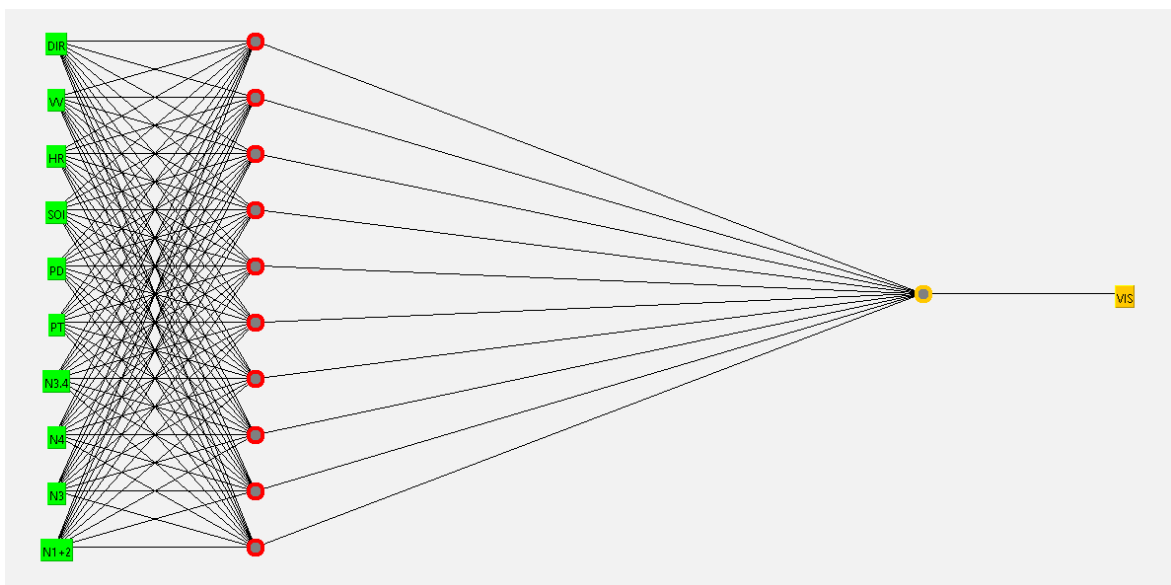
**Figura 27.** Diseño de la red neuronal de cinco neuronas ocultas para la Visibilidad Horizontal.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

Se diseñó una red neuronal de diez neuronas en la capa de entrada, una capa oculta de tres neuronas y una capa de salida de una neurona, mostrada en la Figura 28. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se escogió fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio de 1713.96m.



**Figura 28.** Diseño de la red neuronal de tres neuronas ocultas para la Visibilidad Horizontal.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

Por último, en la Figura 29 se muestra el diseño una red neuronal de diez neuronas en la capa de entrada, una capa oculta de diez neuronas y una capa de salida de una neurona. Se utilizó una función de activación sigmoidea con una tasa de aprendizaje 0.3 y una tasa de momento 0.2, y el número de iteraciones que se escogió fue de 500. Se tuvo para este diseño de modelo una raíz del error cuadrático medio 1613.27m.



**Figura 29.** Diseño de la red neuronal de diez neuronas ocultas para la Visibilidad Horizontal.  
**FUENTE:** Elaboración mediante programa WEKA 3.8.

De los tres diseños mostrados se puede concluir que a más número de neuronas en la capa oculta el error es mínimo. Por lo tanto, cada vez que se va realizando una clasificación, en un orden diferente, los pesos de las conexiones aumentan la velocidad durante el proceso,

ocasionando un lento descenso del error producido en el nivel de salida. Finalmente, el diseño Seleccionado Será El De Diez Neuronas En La Capa Oculta.

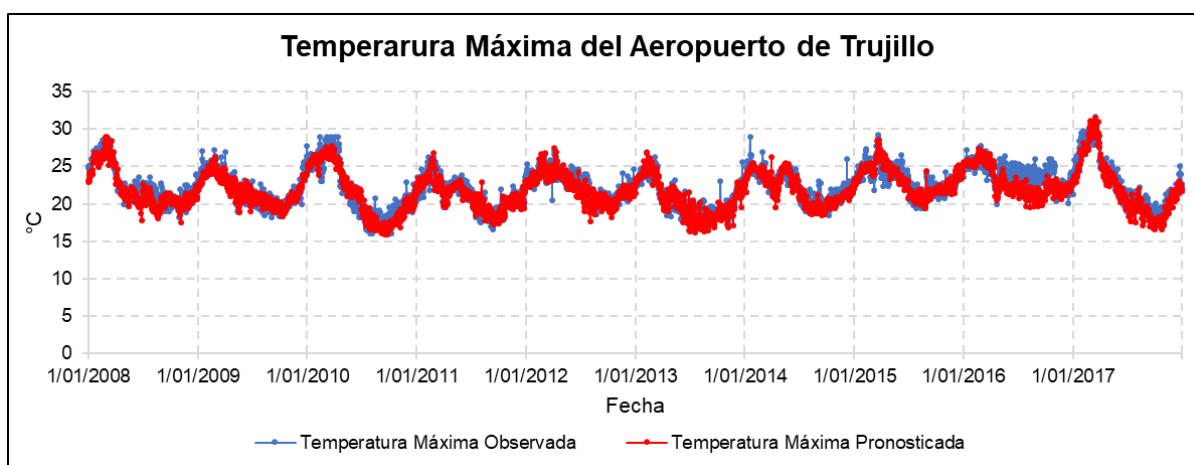
#### 4.4. Validación del modelo que optimice el pronóstico, con menor grado de error, de las variables meteorológicas Visibilidad Horizontal y Temperaturas Máximas y Mínimas.

##### 4.4.1. Temperatura Máxima

Respecto a los datos obtenidos por medio del modelo de redes neuronales artificiales de la variable temperatura máxima, se calculó la siguiente información: el ECM es  $1.36\text{ }^{\circ}\text{C}^2$  y la correlación es 0.91, siendo alta. El RECM muestra que el error es de  $1.17^{\circ}\text{C}$ , una variación pequeña. El BIAS muestra una subestimación de  $0.42^{\circ}\text{C}$ .

**Tabla 2:** Evaluación del pronóstico de la temperatura máxima obtenido por Redes Neuronales Artificiales de Retropropagación.

ECM ( $^{\circ}\text{C}^2$ )	RECM ( $^{\circ}\text{C}$ )	EAM ( $^{\circ}\text{C}$ )	EAMN	BIAS ( $^{\circ}\text{C}$ )	COR
1.36	1.17	0.89	0.04	-0.42	0.91



**Figura 30.** Temperatura Máxima del aeropuerto de Trujillo observada versus data pronosticada con Redes Neuronales Artificiales de Retropropagación.

Se puede ver en la Figura 30 la serie observada versus la serie de datos obtenidos mediante Redes Neuronales Artificiales de Retropropagación. Se percibe que el modelo refleja de manera correcta los datos con muy bajo grado de error, a excepción de algunos momentos del año 2010 y 2016 donde el modelo tiende a subestimar.

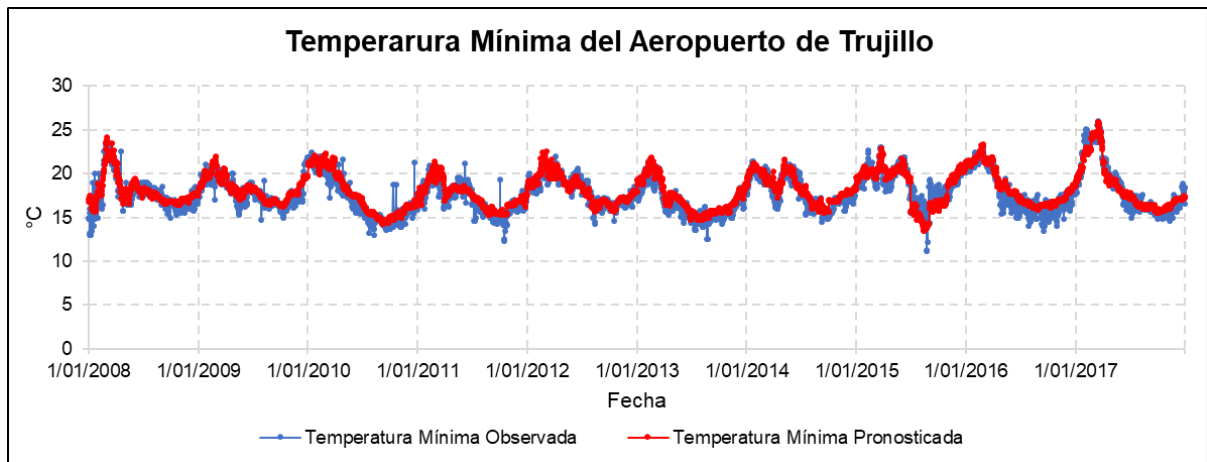
#### 4.4.2. Temperatura Mínima

Respecto a los datos obtenidos por medio del modelo de redes neuronales artificiales de la variable temperatura mínima, se calculó la siguiente información: el ECM es  $0.98 \text{ } ^\circ\text{C}^2$  y la correlación es 0.90, siendo alta. Según el RECM el error es de  $0.99^\circ\text{C}$ , que es una variación pequeña. Si se observa el BIAS se puede ver que en promedio muestra una sobrestimación de  $0.33^\circ\text{C}$ .

**Tabla 3:** Evaluación del pronóstico de temperatura mínima obtenido por Redes Neuronales Artificiales de Retropropagación.

ECM ( $^\circ\text{C}^2$ )	RECM ( $^\circ\text{C}$ )	EAM ( $^\circ\text{C}$ )	EAMN	BIAS ( $^\circ\text{C}$ )	COR
0.98	0.99	0.90	0.04	0.33	0.90

Se puede ver en la Figura 31 el gráfico de la serie observada versus la serie de datos obtenidos mediante Redes Neuronales Artificiales de Retropropagación. Se puede percibir que la serie pronosticada sigue la forma de la serie observada, sin embargo, en algunos momentos del año 2016 el pronóstico tiende a sobrestimar.



**Figura 31.** Temperatura Mínima del aeropuerto de Trujillo observada versus data pronosticada con Redes Neuronales Artificiales de Retropropagación.

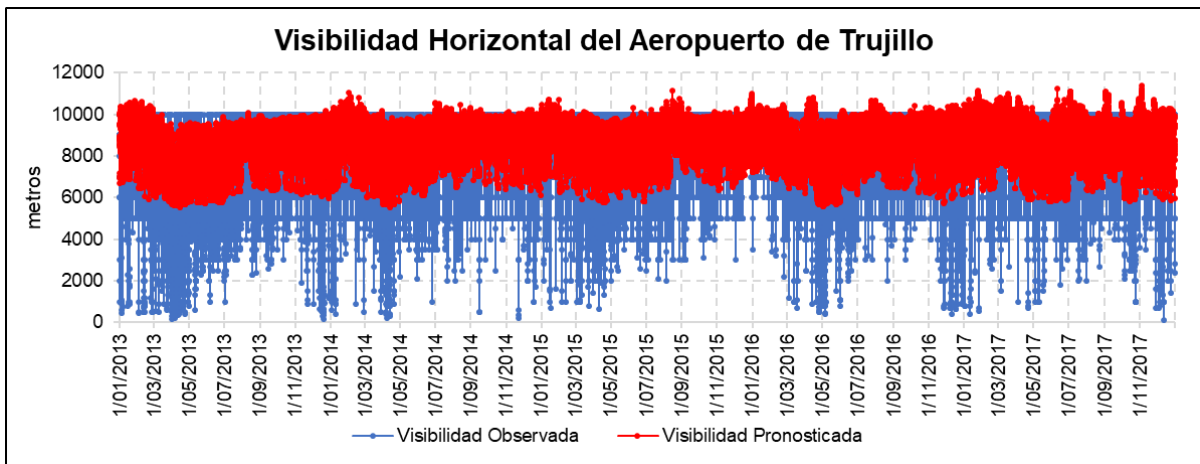
#### 4.4.3. Visibilidad Horizontal

Respecto a los datos obtenidos por medio del modelo de redes neuronales artificiales de la variable visibilidad horizontal, se calculó la siguiente información: el ECM es bastante grande, tiene un valor de  $2602607.83 \text{ m}^2$ , lo que indica que existen errores en el pronóstico; sin embargo, la correlación es alta con un valor de 0.65. En cuanto al RECM y EAM muestran un error general entre 1219.77 y 1613.26 m. El BIAS muestra que se está sobrestimando en promedio 475.04 m.

**Tabla 4:** Evaluación del pronóstico de visibilidad horizontal obtenido por Redes Neuronales Artificiales de Retropropagación.

ECM (m <sup>2</sup> )	RECM (m)	EAM (m)	EAMN	BIAS (m)	COR
2602607.83	1613.26	1219.77	0.24	475.04	0.65

Se puede ver en la Figura 30 el gráfico de la serie observada versus la serie de datos obtenidos mediante Redes Neuronales Artificiales de Retropropagación. Se muestra claramente la presencia de errores, sin embargo, podemos ver que sí nos muestra una relación siguiendo los mismos patrones de tendencia.



**Figura 32.** Visibilidad Horizontal del aeropuerto de Trujillo observada versus data pronosticada con Redes Neuronales Artificiales de Retropropagación.

## V. CONCLUSIONES

El aeropuerto internacional de Trujillo muestra durante los años 2008 al 2017 una temperatura máxima y mínima del aire media mensual con valores más altos en la estación de verano y valores más bajos durante la transición de la estación de invierno a primavera. Durante los años 2013 al 2017 se tiene una visibilidad horizontal que presenta una mayor frecuencia de valores entre 5000 a 1000 metros en horas de la madrugada y mañana en la estación de verano y primavera, y además en últimas horas de la noche en la estación de otoño e invierno; y se presenta una mayor frecuencia de valores menores a 1000 metros en horas del amanecer en la estación de verano y otoño. Debido al entorno costero en el que se encuentra la zona de estudio, los sistemas de mayor influencia son el mar peruano y la masa de agua subtropical superficial, evidenciadas en el estado térmico del agua superficial del mar, y el Anticiclón del Pacífico sur, evidenciado en los gradientes de presión que se generan por su fluctuación que originan el viento que interactúa cinemáticamente con las masas de aire, desarrollando diversas condiciones del tiempo meteorológico.

Para el desarrollo de la red neuronal, como variables de entrada se tomaron en cuenta todos los valores de la serie de diez años de datos diarios de la temperatura máxima y mínima de los años 2008 al 2017, y todos los valores de la serie de cinco años de datos horarios de la visibilidad horizontal de los años 2013 al 2017, del aeropuerto de Trujillo. Se seleccionaron como variables explicativas los índices globales de la Temperatura superficial del mar en el pacífico ecuatorial en las regiones Niño 1+2, Niño 3, Niño 4 y Niño 3.4, la presión al nivel del mar en Darwin y Tahití, y el Índice de Oscilación del Sur (SOI); se consideró también la variable meteorológica de la humedad relativa y para la visibilidad horizontal además se consideró el viento, debido a que estas variables ayudan a agregarle las condiciones del medio atmosférico propias del lugar a la red neuronal.

Se seleccionó una red neuronal artificial de retropropagación de tres capas: una capa de entrada, una capa oculta y una capa de salida, escogiendo el total de datos para la fase de entrenamiento como para su evaluación; con una función de activación sigmoidea de una



taza de aprendizaje 0.3, una tasa de momento 0.2 y un número de iteraciones de 500. Se diseñaron tres diferentes modelos de red neuronal tanto para la temperatura máxima y mínima del aire y para la visibilidad horizontal, resultando que a más número de neuronas en la capa oculta el error es mínimo; escogiendo el modelo de red neuronal artificial de retropropagación para el pronóstico de temperatura máxima y mínima de ocho neuronas en la capa de entrada, ocho neuronas en la capa oculta y una capa de salida de una neurona, y para el pronóstico de visibilidad horizontal se escogió el diseño con diez neuronas en la capa de entrada, diez neuronas en la capa oculta y una capa de salida de una neurona.

El modelo de red neuronal artificial de retropropagación para el pronóstico de temperatura máxima presentó una RECM de 1.17 °C con una correlación alta de 0.92, el modelo de red neuronal artificial de retropropagación para el pronóstico de temperatura mínima tuvo una RECM de 0.99 °C y una correlación alta de 0.90, y el modelo para el pronóstico de visibilidad horizontal presentó una RECM muy alta con un valor de 1613.26 m, lo que indica que existen errores altos en el modelo desarrollado, sin embargo, tiene una correlación alta con un valor de 0.65. Analizando la serie de datos observados versus la serie pronosticada por la red neuronal, se percibe que para la temperatura el modelo refleja de manera correcta los datos con muy bajo grado de error, a excepción de algunos momentos donde el modelo tiende a subestimar para la temperatura máxima y sobrestimar para la temperatura mínima; en cambio para la visibilidad horizontal se puede percibir claramente la presencia de errores, sin embargo, podemos ver que sí muestra una relación siguiendo los mismos patrones de tendencia con una sobrestimación.

## **VI. RECOMENDACIONES**

El estudio se enriquecería si se agregaran a las variables explicativas más datos de las variables relacionadas a la zona como nubosidad, presión atmosférica de la estación, temperatura del aire, etc. También se podría mejorar los pronósticos haciendo una revisión de calidad de las variables explicativas de los índices globales, ya que obtuvieron mediante información secundaria, asumiéndose que no tienen valores atípicos.

Se debería hacer el análisis con otro tipo de red neuronal, para comprobar otro modelo que pueda presentar mejores pronósticos en comparación con las Redes Neuronales Artificiales de Retropropagación.

Para próximos estudios podría aplicarse técnicas de imputación de datos, para obtener la mayor cantidad de información de las variables respuesta estudiadas o realizar análisis de periodos más grandes de tiempo.

## VII. BIBLIOGRAFÍA

- Alonso, F. (2000) Tema 2. Radiación, Balance energético terrestre. Temperatura. Universidad de Murcia. Recuperado de:  
<https://www.um.es/geograf/clima/tema01.html>
- Aránguiz, G., Morales, F., Nieto, J., & Silva, G. (2009) Diseño generativo aplicación en sistemas de atrapanieblas en el norte de Chile. Santiago de Chile: Universidad de Chile.
- CORPAC S.A. (2018). Nuestra actividad. 06 de diciembre del 2018. Recuperado de:  
<http://www.corpac.gob.pe/Main.asp?T=3763>
- Freedman, D., Pisani, R., Purves, R. & Adhikari, A. (1993) Estadística. 2da Edición. Traducción y edición: Bosch, A. Barcelona. 768 págs.
- González, B. (2005). Meteorología Aeronáutica. AVA, Actividades Varias Aeronáuticas, S.L. Madrid, España. 388 págs.
- Gutiérrez J., Cano R., Cofiño A. & Sordo C. (2004) Redes probabilísticas y neuronales en las ciencias atmosféricas. Instituto Nacional de Meteorología. Ministerio de Medio Ambiente. España. 279 págs.
- Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. 2da Ed. Prentice Hall. New Jersey. 833 págs.
- Hecht-Nielsen, R. (1988). Neurocomputing: Picking the Human Brain. IEEE Spectrum, 25, pp. 36-41, marzo de 1988. Reimpreso en el texto Artificial Neural Networks: Theoretical Concepts (Vemuri, V. ed.), pp. 13-18, IEEE Computer Society Press Technology Series.
- Hilera, J. & Martínez, V. (2000). Redes Neuronales Artificiales: Fundamentos, modelos y aplicaciones. Alfaomega, Santafé de Bogotá. 391 págs.
- IPCC, Grupo Intergubernamental de Expertos sobre el Cambio Climático. (2015). Cambio climático 2014: Informe de síntesis. Contribución de los Grupos de trabajo I, II y III al Quinto Informe de Evaluación del Grupo Intergubernamental de Expertos

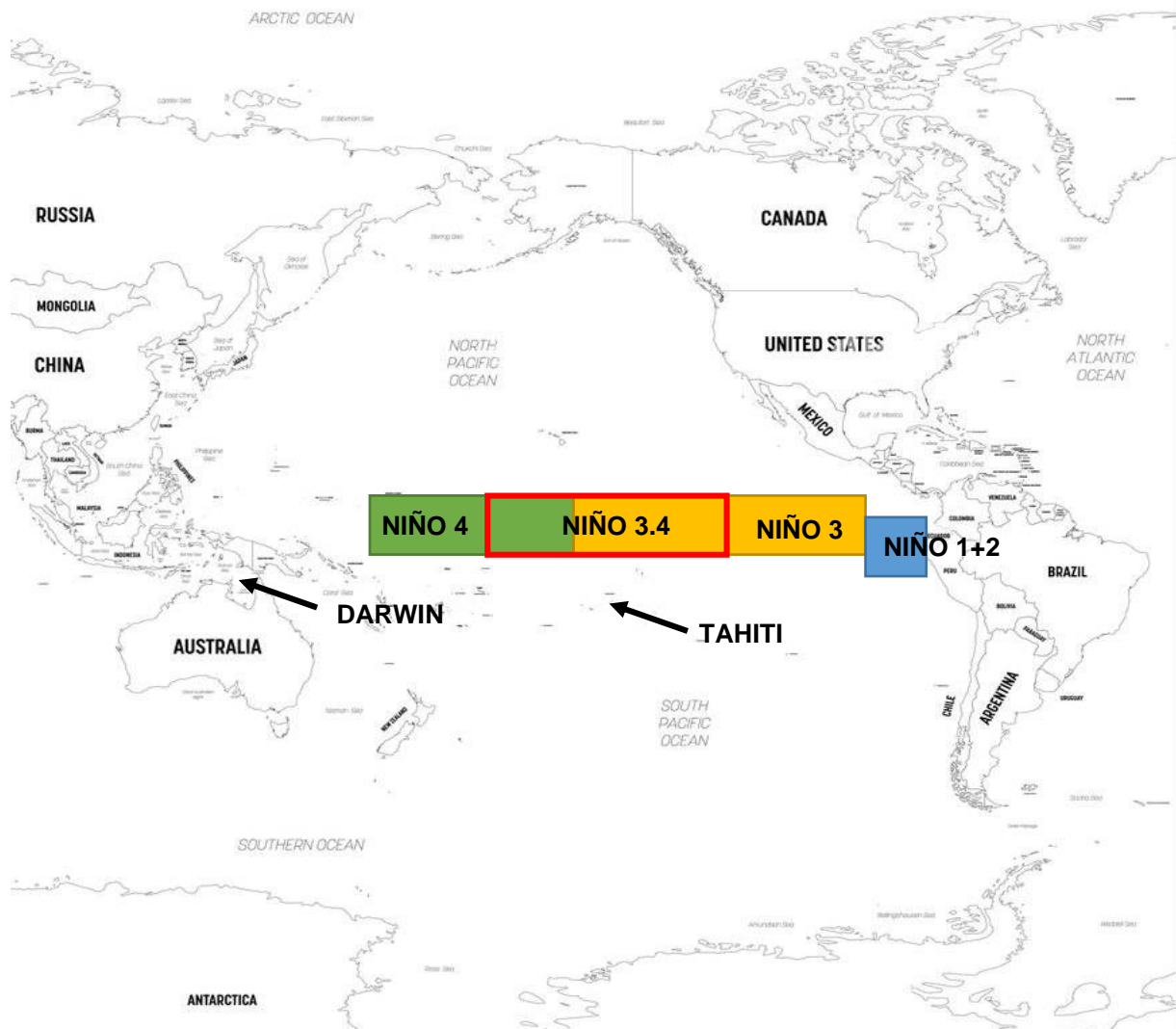
- sobre el Cambio Climático [Equipo principal de redacción, R.K. Pachauri y L.A. Meyer (eds.)]. Ginebra, Suiza, 157 págs.
- Latínez, K. (2009). Comparación de los métodos regresión multivariada adaptativa utilizando splines (mars) y redes neuronales artificiales backpropagation (RNAB) para el pronóstico de lluvias y temperaturas en la cuenca del río Mantaro. Universidad Nacional Agraria La Molina. Facultad De Economía Y Planificación. Lima, Perú.
- Marzban, C. (2000). A neural network for tornado diagnosis. *Neural Computing and Application*. Vol. 9: pp.133-141.
- Morán, G. & Alvarado, D. (2010) *Métodos de investigación*. Pearson educación. México. 80 págs.
- Ojeda, J. (2009). Aplicación de Redes Neuronales Artificiales a la Predicción y Control de Demanda de Energía Eléctrica en Empresas Industriales. Universidad Nacional de Ingeniería, Facultad de Ingeniería Industrial y Sistemas, Escuela de Postgrado. Lima, Perú.
- OMM, Organización Meteorológica Mundial. (2018). Guía de prácticas climatológicas. Ginebra, Suiza: OMM. 169 págs.
- Ovando, G., Bocco, M., & Sayago, S. (2005). Redes neuronales para modelar predicción de heladas. *Agricultura Técnica*, Vol 65(1), pp. 65-73. Sitio web: <https://dx.doi.org/10.4067/S0365-28072005000100007>
- Parker, D. (1987). Second Order Backpropagation: Implementing an Optimal  $O(n)$  Approximation to Newton's Method as an Artificial Neural Network. Submitted to Computer. Colorado, USA.
- Pérez, L. (2019). Captación de agua de niebla y análisis de la calidad para consumo humano en el asentamiento humano leandra ortega, pachacútec - ventanilla, 2016. Tesis para el título de ingeniera ambiental. Universidad Científica del Sur. Lima, Perú.
- Petterssen, S. (1956) *Weather analysis and forecasting*. New York, EEUU. McGraw-Hill Book. 2 t.
- Pielke, R. (1984) *Mesoscale Meteorological Modeling*. Academic Press. Orlando. 611 págs.

- Pinche, C. (1986). Estudio de las Condiciones Climáticas y de la Niebla en la Costa Norte de Lima. Tesis Ingeniero Meteorólogo. Universidad Nacional Agraria La Molina. Lima, Perú.
- Quispe, N. & Chira, J. (2004). Entornos sinópticos en la ocurrencia de niebla en Lima. Boletín Meteorológico e hidrológico del Perú – SENAMHI. Año IV, N°01. Pp. 54-65.
- Redacción BBC News Mundo (2019, 15 de setiembre) Corriente de Humboldt: el fenómeno por el que el mar de Chile y Perú es hasta 10 °C más frío que otras zonas del Pacífico y cómo afecta al resto del mundo. BBC News Mundo. Recuperado de: <https://www.bbc.com/mundo/noticias-49682744#:~:text=La%20corriente%20Humboldt%2C%20tambi%C3%A9n%20llamada,norte%20y%20hacia%20la%20superficie.>
- Redacción RPP (2014, 10 de junio). Trujillo: mal tiempo obliga a cancelar vuelo en aeropuerto. Radio Programas del Perú, RPP Noticias. Sitio web: <https://rpp.pe/peru/actualidad/trujillo-mal-tiempo-obliga-a-cancelar-vuelo-en-aeropuerto-noticia-698930>
- Redacción RPP (2018, 18 de abril). Suspenden vuelos en aeropuerto de Trujillo por densa neblina. Radio Programas del Perú, RPP Noticias. Sitio web: <https://rpp.pe/peru/la-libertad/suspenden-vuelos-en-aeropuerto-de-trujillo-por-densa-neblina-noticia-1117417>
- Redacción RPP (2018, 17 de mayo). Suspenden vuelos en aeropuerto internacional de Trujillo por mal tiempo. Radio Programas del Perú, RPP Noticias. Sitio web: <https://rpp.pe/peru/la-libertad/suspenden-vuelos-en-aeropuerto-internacional-de-trujillo-por-falta-de-radar-noticia-1123335>
- Resolución Directoral N° 302-2010-MTC/12 (2010, 26 de octubre). RAP 303 Servicio Meteorológico para la Navegación Aérea. Normas Legales, separata especial. Diario Oficial El Peruano, 25 de noviembre de 2010.
- Salvador, M. & Gargallo, P. (2003). Análisis Exploratorio de Datos. 10 de junio 2008, de 5campus.com, Estadística. Recuperado de: <http://www.5campus.com/leccion/aed>
- Smith, J. (2006). Diccionario del Tiempo y Clima. Edición revisada. Facts On File, Inc. New York, Estados Unidos. 262 págs.

- Stauffer, D. & Seaman, N. (1990). Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.*, 118, pp. 1250-1277.
- The Associated Press (2014, 10 de diciembre). Flight delays are costing airlines serious money. Mashable. Sitio web: <https://mashable.com/archive/cost-of-delayed-flights>
- Vistin, D. (2014). Estudio de factibilidad para el aprovechamiento de agua por medio de dos tipos de neblinómetros en las tres cuencas de la parroquia Achupallas, Cantón Alausi, provincia de Chimborazo. Tesis de pregrado. Escuela Superior Politécnica de Chimborazo. Riobamba, Ecuador:
- Werbos, P. (1974) Backpropagation Through Time: What It Does and How to Do It. 11 págs. Recuperado de: <http://www.sysc.pdx.edu/classes/Werbos-Backpropagation%20through%20time.pdf>.
- Wichter, Z. (2017, 23 de junio). El efecto del cambio climático en tus viajes en avión. *The New York Times*. Recuperado de: <https://www.nytimes.com/es/2017/06/23/vuelos-cambio-climatico-turbulencia/>

## **VIII. ANEXOS**

**Anexo 1.** Ubicación geográfica de los índices utilizados en el análisis de los datos.





**Anexo 2.** Variables globales: nombre, abreviatura y uso.

Variable	Abreviatura	Se utiliza en el modelado de	
		Vis	Tmin y Tmax
Índice de TSM en el Extremo Oriental del Pacífico (Región Niño 1+2)	N12	Sí	Sí
Índice de TSM Oriental del Pacífico (Región Niño 3)	N3	Sí	Sí
Índice de TSM Occidental del Pacífico (Región Niño 4)	N4	Sí	Sí
Índice de TSM en el Pacífico Central Oriental (Región Niño 3.4)	N34	Sí	Sí
Presión a nivel del mar en Darwin	D	Sí	Sí
Presión a nivel del mar en Tahití	T	Sí	Sí
Índice de Oscilación del Sur	SOI	Sí	Sí