

**UNIVERSIDAD NACIONAL AGRARIA**

**LA MOLINA**

**FACULTAD DE CIENCIAS**



**“PREDICCIÓN DIARIA DE TEMPERATURAS EXTREMAS Y PRECIPITACIÓN  
EN LA REGIÓN ÁNCASH MEDIANTE POSPROCESAMIENTO ESTADÍSTICO  
DEL MODELO GFS”**

Presentado por:

**ANDY JEAN PIERRE FLORES AGUILAR**

Tesis para Optar el Título Profesional de:

**INGENIERO METEORÓLOGO**

LIMA – PERÚ

2024

---

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

# PREDICCIÓN DIARIA DE TEMPERATURAS EXTREMAS Y PRECIPITACIÓN EN LA REGIÓN ÁNCASH MEDIANTE POSPROCESAMIENTO ESTADÍSTICO DEL MODELO GFS

## INFORME DE ORIGINALIDAD

16%

INDICE DE SIMILITUD

16%

FUENTES DE INTERNET

8%

PUBLICACIONES

0%

TRABAJOS DEL  
ESTUDIANTE

## FUENTES PRIMARIAS

1	<a href="https://repositorio.unal.edu.co">repositorio.unal.edu.co</a> Fuente de Internet	1 %
2	<a href="https://repositorio.lamolina.edu.pe">repositorio.lamolina.edu.pe</a> Fuente de Internet	1 %
3	<a href="https://repositorio.udec.cl">repositorio.udec.cl</a> Fuente de Internet	1 %
4	<a href="https://es.scribd.com">es.scribd.com</a> Fuente de Internet	1 %
5	<a href="https://egusphere.copernicus.org">egusphere.copernicus.org</a> Fuente de Internet	<1 %
6	<a href="https://cdn.www.gob.pe">cdn.www.gob.pe</a> Fuente de Internet	<1 %
7	<a href="https://bibliotecadigital.exactas.uba.ar">bibliotecadigital.exactas.uba.ar</a> Fuente de Internet	<1 %
8	<a href="https://zagan.unizar.es">zagan.unizar.es</a> Fuente de Internet	<1 %

**UNIVERSIDAD NACIONAL AGRARIA  
LA MOLINA**

**FACULTAD DE CIENCIAS**

**“PREDICCIÓN DIARIA DE TEMPERATURAS EXTREMAS Y PRECIPITACIÓN  
EN LA REGIÓN ÁNCASH MEDIANTE POSPROCESAMIENTO ESTADÍSTICO  
DEL MODELO GFS”**

Presentado por:

**ANDY JEAN PIERRE FLORES AGUILAR**

Tesis para Optar el Título Profesional de:

**INGENIERO METEORÓLOGO**

Sustentada y aprobada por el siguiente jurado:

---

Dr. Alexis Nicolas Ibañez Blancas  
PRESIDENTE

---

Mg. Sc. Victoria Doris Calle Montes  
MIEMBRO

---

Dr. Aldo Saturnino Moya Álvarez  
MIEMBRO

---

Mg. Sc. Weidi Flores Villanueva  
ASESOR

La presente investigación se realizó en marco del  
Proyecto Meteo-Huascarán: Ecoturismo seguro con  
monitoreo y pronóstico meteorológico automatizado con  
machine learning en el Parque Nacional Huascarán  
(Prociencia 036-2021).



## DEDICATORIA

*Dedico este trabajo a mi familia por su amor, consejos y apoyo incondicional a lo largo de mi etapa de estudiante. A mis abuelos Leopoldo y Felicitas por guiarme siempre, alentándome a ser un gran profesional y una mejor persona.*

## AGRADECIMIENTOS

*Quiero expresar mi profundo agradecimiento al PhD. Ken Takahashi Guevara por brindarme la valiosa oportunidad de formar parte del proyecto Meteo-Huascarán y por permitirme desarrollar esta investigación. Su apoyo constante y su vasta experiencia en el campo de la meteorología han enriquecido enormemente este trabajo.*

*Mi sincero agradecimiento al M.Sc. Gustavo De la Cruz Montalvo por sus invaluable enseñanzas, constante apoyo y valiosos consejos a lo largo de todo este proceso de investigación. Su orientación fue fundamental para llegar a la culminación de este proyecto.*

*También quiero extender mi gratitud al M.Sc. Weidi Flores por su asesoría y apoyo durante el desarrollo de esta investigación. Sin su guía y colaboración, no habría sido posible alcanzar este punto.*

*No puedo pasar por alto el agradecimiento al Servicio Nacional de Meteorología e Hidrología por proporcionar la información meteorológica necesaria y por facilitar los equipos tecnológicos esenciales para llevar a cabo esta investigación de manera adecuada. Además, deseo expresar mi reconocimiento a mis colegas y amigos de la subdirección Modelamiento Numérico de la Atmósfera, cuyas valiosas enseñanzas y sugerencias contribuyeron significativamente al éxito de este trabajo.*

*Finalmente, un agradecimiento especial a mis amigos, por todas las palabras de aliento, consejos en los momentos complicados y la confianza necesaria para poder terminar este trabajo.*

## ÍNDICE GENERAL

RESUMEN .....	viii
ABSTRACT .....	ix
I. INTRODUCCIÓN .....	1
II. REVISIÓN DE LITERATURA .....	3
2.1. Perfil Climático de la región Áncash .....	3
2.2. Variabilidad de la temperatura en los Andes .....	4
2.3. Variabilidad de la precipitación en los Andes .....	5
2.4. Modelamiento numérico del tiempo .....	6
2.5. Modelo Global Forecast System .....	8
2.6. Posprocesamiento estadístico .....	9
2.6.1. Enfoque Model Output Statistics .....	10
2.7. Análisis de Componentes Principales .....	11
2.8. Regresión de componentes principales .....	12
2.9. Variabilidad de alta y baja frecuencia en series de tiempo .....	13
2.10. Árboles de clasificación .....	14
2.11. Random Forests .....	15
III. METODOLOGÍA .....	18
3.1. Área de estudio .....	18
3.2. Información observada de temperaturas extremas diarias .....	20
3.3. Información observada de precipitación acumulada diaria .....	22
3.4. Información de salidas del modelo GFS .....	22
3.5. Metodología para pronóstico de temperaturas extremas diarias .....	25
3.5.1. Control de calidad e Imputación de datos .....	27
3.5.2. División de conjuntos de datos .....	27
3.5.3. Extracción de variabilidad de alta frecuencia .....	29

3.5.4.	Componentes principales de predictores y predictandos.....	29
3.5.5.	Selección de predictores y construcción de ecuaciones de regresión.....	31
3.5.6.	Validación de modelo de pronóstico temperaturas extremas .....	32
3.6.	Metodología para pronóstico de probabilidad de precipitación.....	35
3.6.1.	División de conjuntos de datos .....	37
3.6.2.	Selección de predictores .....	37
3.6.3.	Construcción de matriz de predictores .....	37
3.6.4.	Transformación de predictandos .....	39
3.6.5.	Calibración de modelos de predicción.....	40
3.6.6.	Validación de modelos .....	41
IV.	RESULTADOS Y DISCUSIÓN .....	44
4.1.	Modelo de pronóstico de temperatura mínima diaria .....	44
4.2.	Modelo de pronóstico de temperatura máxima.....	60
4.3.	Modelo de pronóstico de probabilidad de precipitación.....	75
V.	CONCLUSIONES .....	85
VI.	RECOMENDACIONES.....	87
VII.	BIBLIOGRAFÍA .....	88
VIII.	ANEXOS .....	94



## ÍNDICE DE TABLAS

<b>Tabla 1:</b> Ubicación y altitud de estaciones con información meteorológica de temperaturas extremas empleadas para la investigación .....	20
<b>Tabla 2:</b> Variables del modelo GFS con su respectivo nivel vertical .....	24
<b>Tabla 3:</b> Grados de Correlación basada en el valor de Coeficiente de Pearson .....	35
<b>Tabla 4:</b> Número de componentes principales incluidas como potenciales predictores para cada variable meteorológica proveniente de modelo numérico para la predicción de temperatura mínima .....	48
<b>Tabla 5:</b> Resultados obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura mínima en el periodo de validación con un horizonte de pronóstico de 24 horas .....	50
<b>Tabla 6:</b> Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura mínima en el periodo de validación con un horizonte de pronóstico de 48 horas .....	51
<b>Tabla 7:</b> Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura mínima en el periodo de validación con un horizonte de pronóstico de 72 horas .....	51
<b>Tabla 8:</b> Número de componentes principales incluidas como potenciales predictores para cada variable meteorológica proveniente de modelo numérico para la predicción de temperatura máxima.....	63
<b>Tabla 9:</b> Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura máxima en el periodo de validación con un horizonte de pronóstico de 24 horas .....	65
<b>Tabla 10:</b> Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura máxima en el periodo de validación con un horizonte de pronóstico de 48 horas .....	65
<b>Tabla 11:</b> Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura máxima en el periodo de validación con un horizonte de pronóstico de 72 horas .....	66

## ÍNDICE DE FIGURAS

<b>Figura 1:</b> Mapa de clasificación climática de Áncash.....	4
<b>Figura 2:</b> Representación esquemática de grillas horizontales y verticales incluidas en modelos numéricos de predicción del tiempo.....	8
<b>Figura 3:</b> Comparación de una serie de tiempo de temperatura máxima diaria con filtro paso bajo y una serie de tiempo representando la variabilidad de alta frecuencia .....	14
<b>Figura 4:</b> Diagrama de flujo de un modelo Random Forests .....	17
<b>Figura 5:</b> Mapa del Parque Nacional Huascarán.....	19
<b>Figura 6:</b> Corte transversal de la cordillera de los Andes en Perú .....	19
<b>Figura 7:</b> Mapa con ubicación de las estaciones meteorológicas empleadas para la investigación.....	21
<b>Figura 8:</b> Esquema de la metodología de investigación para pronóstico de temperaturas extremas .....	26
<b>Figura 9:</b> Serie temporal de temperatura mínima de la estación Purhuay con la aplicación de filtro de Butterworth.....	29
<b>Figura 10:</b> Esquema de la metodología de investigación para pronóstico de probabilidad de precipitación .....	36
<b>Figura 11:</b> Esquema metodológico para el desarrollo de un modelo de predicción de probabilidad de precipitaciones extremas empleando Random Forests .....	39
<b>Figura 12:</b> Modelo de matriz de confusión en clasificación binaria .....	41
<b>Figura 13:</b> Interpretación de Curva ROC .....	43
<b>Figura 14:</b> Mapas de correlación de Pearson de variabilidad de alta frecuencia de temperatura mínima y variables meteorológicas del modelo GFS correspondientes a las 12 UTC en la zona de estudio .....	46
<b>Figura 15:</b> Varianza explicada individual y acumulada de las componentes principales obtenidas de la variabilidad de alta frecuencia de temperatura mínima proveniente de estaciones meteorológicas .....	47
<b>Figura 16:</b> Series temporales correspondientes a variabilidad de alta frecuencia de temperatura mínima pronosticada y observada para cada estación meteorológica .....	54

<b>Figura 17:</b> Gráfico de barras del RMSE para el pronóstico de temperatura mínima por persistencia de 24 horas y pronóstico con el modelo empírico con horizonte de pronóstico de 24h, 48h y 72h.....	55
<b>Figura 18:</b> Diagrama de cajas de RMSE para pronóstico de temperatura mínima a diferentes horizontes de tiempo y con el método por persistencia de 24 horas.....	56
<b>Figura 19:</b> (a) Diagrama de cajas de correlación de Pearson para pronóstico alta frecuencia de temperatura mínima (b) Diagrama de cajas de correlación de Pearson para pronósticos finales de temperatura mínima.....	56
<b>Figura 20:</b> Pronóstico final de temperatura mínima en el periodo de validación .....	59
<b>Figura 21:</b> Mapas de correlación de Pearson de variabilidad de alta frecuencia de temperatura máxima y variables meteorológicas del modelo GFS correspondientes a las 18 UTC en la zona de estudio .....	61
<b>Figura 22:</b> Varianza explicada individual y acumulada de las componentes principales obtenidas de la variabilidad de alta frecuencia de temperatura máxima proveniente de estaciones meteorológicas.....	62
<b>Figura 23:</b> Series temporales correspondientes a variabilidad de alta frecuencia de temperatura mínima pronosticada y observada para cada estación meteorológica .....	69
<b>Figura 24:</b> Gráfico de barras del RMSE para el pronóstico de temperatura máxima por persistencia de 24 horas y pronóstico con el modelo empírico con horizonte de pronóstico de 24h, 48h y 72h.....	70
<b>Figura 25:</b> Box-Plot del RMSE para pronóstico de temperatura máxima a diferentes horizontes temporales y por el método por persistencia de 1 día.....	70
<b>Figura 26:</b> (a) Diagrama de cajas de correlación de Pearson para pronóstico alta frecuencia de temperatura máxima (b) Diagrama de cajas de correlación de Pearson para pronósticos finales de temperatura máxima .....	71
<b>Figura 27:</b> Pronóstico final de temperatura mínima en el periodo de validación .....	74
<b>Figura 28:</b> Curvas ROC para distintos umbrales de precipitación.....	75
<b>Figura 29:</b> Matriz de confusión para modelo de pronóstico de precipitación con umbral de 0.1 mm.....	77

<b>Figura 30:</b> Matriz de confusión para modelo de pronóstico de precipitación con umbral de 1 mm.....	79
<b>Figura 31:</b> Matriz de confusión para modelo de pronóstico de precipitación con umbral de 2 mm.....	81
<b>Figura 32:</b> Matriz de confusión para modelo de pronóstico de precipitación con umbral de 5 mm.....	83
<b>Figura 33:</b> Mapa de probabilidad de precipitación para el modelo de pronóstico con umbral de 5 mm/día.....	84

## ÍNDICE DE ANEXOS

<b>Anexo 1:</b> Gráficas de varianza explicada para variables meteorológicas filtradas del modelo numérico GFS empleadas para el modelo de pronóstico de temperatura mínima .....	94
<b>Anexo 2:</b> Gráficas de varianza explicada para variables meteorológicas filtradas del modelo numérico GFS empleadas para el modelo de pronóstico de temperatura máxima .....	97
<b>Anexo 3:</b> Mapa de pronóstico de temperaturas extremas diarias en la región Áncash .....	100

## RESUMEN

En esta investigación, se desarrollaron modelos de pronóstico para temperaturas extremas y probabilidad de precipitación en el área del Parque Nacional Huascarán y sus alrededores, en la región Áncash. Se utilizaron los datos grillados de PISCOp para calibrar los modelos de precipitación y datos de temperaturas extremas diarias recopilados de 10 estaciones en el área de estudio para calibrar los modelos de temperatura. Se aplicó el método Model Output Statistics (MOS) junto con técnicas específicas para cada tipo de predicción. Los predictores para estos modelos corresponden a pronósticos del modelo numérico Global Forecast System (GFS) con una resolución de  $0.25^\circ$ . Para las predicciones de temperaturas extremas diarias, se empleó la técnica de reducción de dimensionalidad conocida como análisis de componentes principales (PCA), junto con la regresión lineal múltiple para obtener las ecuaciones de pronóstico. En el caso de los modelos de probabilidad de precipitación, se utilizaron tanto la técnica de PCA como Random Forests para tareas de clasificación. Los modelos desarrollados para las temperaturas extremas superaron las predicciones sin posprocesamiento del modelo GFS en la zona de estudio y también superaron las predicciones por persistencia. Se observó una mayor precisión en la predicción de las temperaturas mínimas en comparación con las temperaturas máximas. Sin embargo, se identificaron ciertas limitaciones en la capacidad del modelo para prever la magnitud de eventos extremos de temperatura. En relación con la probabilidad de precipitación, los resultados variaron según el umbral establecido. Los modelos para umbrales de 0.1 mm/día y 1 mm/día demostraron ser eficaces en las predicciones, mientras que el modelo para 5 mm/día generó un alto número de falsas alarmas. Estos hallazgos indican que los modelos desarrollados son prometedores, pero también señalan áreas que podrían ser mejoradas para aumentar su precisión, especialmente en la identificación de eventos extremos.

**Palabras clave:** temperatura máxima, temperatura mínima, precipitación, MOS, PCA, GFS

## ABSTRACT

In this research, forecast models were developed for extreme temperatures and precipitation probability in the area of Huascarán National Park and its surroundings, in the Áncash region. Gridded data from PISCOp were used to calibrate the precipitation models and daily extreme temperature data collected from 10 stations in the study area were used to calibrate the temperature models. The Model Output Statistics (MOS) method was applied together with specific techniques for each type of prediction. The predictors for these models correspond to forecasts from the Global Forecast System (GFS) numerical model with a resolution of  $0.25^\circ$ . For the daily extreme temperature forecasts, the dimensionality reduction technique known as principal component analysis (PCA) was used, together with multiple linear regression to obtain the forecast equations. In the case of the precipitation probability models, both PCA and Random Forests were used for classification tasks. The models developed for extreme temperatures outperformed the non-post-processed predictions of the GFS model in the study area and also outperformed the predictions by persistence. Higher accuracy was observed in the prediction of minimum temperatures compared to maximum temperatures. However, certain limitations were identified in the model's ability to predict the magnitude of extreme temperature events. In relation to the probability of precipitation, results varied depending on the threshold set. The models for thresholds of 0.1 mm/day and 1 mm/day proved to be effective in predictions, while the model for 5 mm/day generated a high number of false alarms. These findings indicate that the developed models are promising, but also point to areas that could be improved to increase their accuracy, especially in the identification of extreme events.

**Keywords:** maximum temperature, minimum temperature, precipitation, MOS, PCA, GFS

## I. INTRODUCCIÓN

El pronóstico meteorológico tiene como objetivo primordial proporcionar información relevante acerca de las condiciones atmosféricas futuras. La importancia de estos pronósticos se refleja en numerosas actividades humanas, como el turismo, la agricultura, la aviación, entre otros, en las cuales las condiciones atmosféricas pueden tener un impacto positivo o negativo (Al-Matarneh et al., 2014).

Los modelos de predicción numérica del tiempo son herramientas fundamentales que contribuyen de manera significativa a la generación de pronósticos meteorológicos precisos. Estos modelos representan la atmósfera mediante cálculos matemáticos complejos y diversas parametrizaciones físicas, utilizando como base las condiciones iniciales o actuales obtenidas de redes de observación meteorológica, como satélites, radiosondas y estaciones meteorológicas (Calvo Sánchez, 2018). Aunque estos modelos de predicción numérica del tiempo están en constante desarrollo y experimentan mejoras continuas en su resolución espacial y temporal, aún enfrentan diversas limitaciones, las cuales se vuelven más evidentes en áreas con topografía compleja.

En este contexto, se aplican diversas técnicas de post procesamiento que calibran modelos estadísticos utilizando las salidas de los modelos de predicción numérica del tiempo. Estas técnicas permiten reducir significativamente los errores presentes en los modelos y obtener información con mayor precisión a nivel local.

La presente investigación tuvo como zona de estudio prioritaria el Parque Nacional Huascarán (PNH), ubicado en la región Áncash, en la sierra central del Perú. Esta área es conocida por su rica diversidad biológica y cultural; además de ser un destino turístico importante que ofrece una amplia gama de actividades recreativas como esquí, escalada, trekking y bicicleta de montaña (SERNANP, 2011). Dado el atractivo turístico del PNH, es fundamental garantizar condiciones seguras para los visitantes, debido a las condiciones topográficas desafiantes y la posible ocurrencia de fenómenos hidrometeorológicos adversos.



En el pasado, se han registrado incidentes en el Parque Nacional Huascarán que han afectado a turistas, subrayando la importancia de contar con pronósticos meteorológicos precisos y oportunos para minimizar al máximo los riesgos. El objetivo principal de esta tesis es desarrollar modelos empíricos para pronosticar con tres días de anticipación las variables meteorológicas de temperatura máxima, temperatura mínima y probabilidad de precipitación en la zona del Parque Nacional Huascarán. Estos modelos se construyeron utilizando variables dinámicas de pronóstico generadas por el modelo numérico Global Forecast System (GFS) durante el período 2015-2021.

La calibración de los modelos se llevó a cabo utilizando el enfoque de Model Output Statistics, en combinación con la técnica de machine learning Random Forest para el pronóstico de probabilidad de precipitación, y la técnica de regresión de componentes principales para el pronóstico de temperaturas extremas diarias. Los resultados de esta investigación tienen como objetivo respaldar la generación de pronósticos precisos en el área del PNH; asimismo la metodología puede ser aplicada sobre otros dominios del Perú con el respectivo ajuste a las condiciones meteorológicas locales.

#### **Objetivo general:**

- Desarrollar modelos empíricos de pronóstico de temperatura máxima, temperatura mínima y probabilidad de precipitación a nivel diario con un horizonte de pronóstico de hasta 3 días a partir de salidas del modelo de pronóstico Global Forecast System (GFS) en el Parque Nacional Huascarán.

#### **Objetivos Específicos**

- Desarrollar modelos de pronóstico de temperatura mínima y temperatura máxima diaria, mediante regresión lineal múltiple y componentes principales de variables meteorológicas provenientes del modelo meteorológico GFS
- Desarrollar un modelo Random Forest para el pronóstico de probabilidad de precipitación entrenado con información del modelo meteorológico GFS
- Determinar el desempeño de los modelos de pronóstico mediante conjuntos de datos independientes y las métricas estadísticas correspondientes.

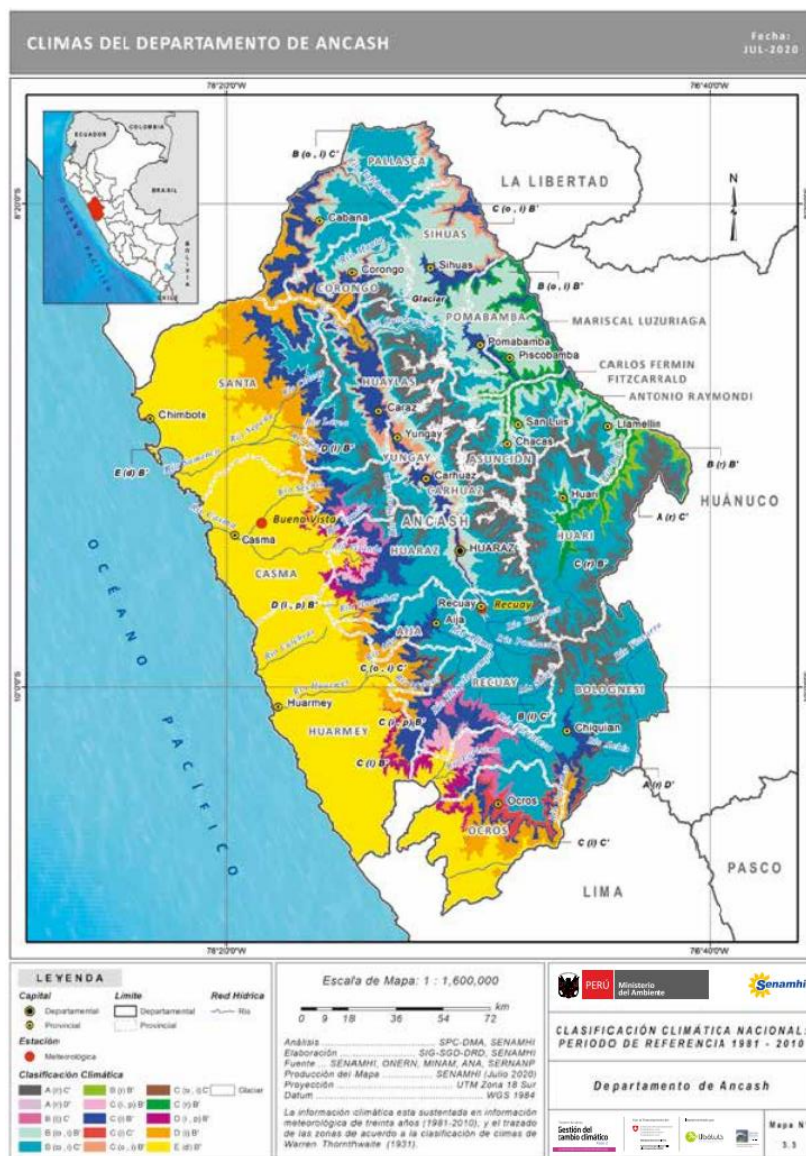
## **II. REVISIÓN DE LITERATURA**

### **2.1. Perfil Climático de la región Áncash**

El comportamiento del tiempo y el clima en el Perú y específicamente en la región de Áncash está influenciado por sistemas sinópticos a gran escala, entre los que se destacan: el Alta de Bolivia, la Vaguada del Noreste de Brasil, el Anticiclón del Pacífico Sur, el Anticiclón del Atlántico Sur y la Zona de Convergencia Intertropical (Acuña et al., 2013).

Dentro de las principales fuerzas que generan la amplia variabilidad climática en el Perú, la Cordillera de los Andes juega un papel fundamental. Este factor físico da origen a un marcado contraste altitudinal que es característico en la zona de la sierra peruana. Según señalan Castro et al. (2021), en Áncash se pueden identificar hasta 16 tipos de climas. El mapa de clasificación climática para la región de Áncash presentado en la Figura 1, muestra que, en sectores costeros de esta región, ubicada en el lado occidental (áreas con altitudes por debajo de los 1,000 m s.n.m.), prevalece un clima árido y templado, con una marcada carencia de humedad a lo largo del año. A medida que se asciende en altitud, el tipo de clima cambia hacia un patrón semiseco y templado, con inviernos secos en la vertiente occidental de la cordillera.

Estas características climáticas similares se observan en el callejón de Huaylas, hasta altitudes inferiores a los 3,500 m s.n.m. Por último, las zonas más elevadas de las cordilleras Negra y Blanca se caracterizan por un clima frío y húmedo durante todo el año, mientras que por encima de los 5,000 m s.n.m. de la Cordillera Blanca, se presenta un clima glaciario con presencia de hielo de manera permanente.



**Figura 1:** Mapa de clasificación climática de Ancash  
**FUENTE:** SENAMHI (2020)

## 2.2. Variabilidad de la temperatura en los Andes

La variabilidad de la temperatura en los Andes depende de una serie de factores. Saavedra y Takahashi (2017) resaltan la importancia del balance radiativo en la superficie, que influye en el comportamiento de la temperatura mínima. Esto se refleja en la fuerte correlación entre la temperatura mínima y la nubosidad. Esta relación se evidencia en la frecuencia de eventos de heladas en los Andes, que ocurren con mayor frecuencia en los meses de junio y julio, durante la estación seca. También señalan que la temperatura máxima presenta una estacionalidad relativamente débil en comparación con la temperatura mínima a lo largo del año.

Gran parte de la nubosidad en los Andes centrales se presenta durante el verano austral (enero, febrero y marzo) cuando los vientos del este en niveles medios y altos, impulsados por el fortalecimiento de la Alta de Bolivia, transportan el flujo ascendente del este cerca de la superficie. Este período de alta nubosidad coincide con las temperaturas más altas durante el verano austral (Sulca et al., 2018).

Imfeld et al. (2021) señalan que las temperaturas muestran un débil ciclo anual debido a las escasas variaciones en la irradiancia solar durante el año, con amplitudes diarias mayores que las amplitudes anuales. La variabilidad de la temperatura está más fuertemente influenciada por el Fenómeno El Niño-Oscilación del Sur (ENSO) y sigue las variaciones de la temperatura de la superficie del mar en el Pacífico central ecuatorial con un desfase de 1-2 meses.

Además, las variaciones en la temperatura y la precipitación también dependen de índices oceánico-atmosféricos a gran escala, como el Índice de Oscilación del Sur (SOI, por sus siglas en inglés) y la temperatura superficial del mar (Casimiro et al., 2012).

### **2.3. Variabilidad de la precipitación en los Andes**

Las condiciones atmosféricas en una región determinada experimentan cambios temporales en forma de fluctuaciones irregulares que abarcan diversas escalas, superponiéndose a los ciclos diarios y anuales. Estas variaciones abarcan desde la escala sinóptica, que está estrechamente relacionada con la meteorología, hasta variaciones intraestacionales, interanuales, interdecadales y a una escala aún mayor. Dichas fluctuaciones derivan de la variabilidad intrínseca de la atmósfera y su interacción con otros componentes del sistema terrestre (Garreaud y Aceituno, 2007).

El comportamiento estacional de las precipitaciones en los Andes peruanos permite identificar claramente dos periodos a lo largo del año, siendo estas estaciones: húmeda (con picos de precipitación en el periodo diciembre-febrero) y seca (mayo-julio). Este comportamiento se presenta bastante similar al patrón estacional de la humedad específica, la nubosidad baja y la temperatura mínima (Saavedra y Takahashi, 2017).

La existencia de estos dos periodos de abundancia y déficit de precipitaciones, se encuentra relacionada con la posición de los vientos zonales en los niveles medios y superiores de la atmósfera. Durante los meses de invierno (junio-agosto), la corriente en chorro subtropical se encuentra en su posición más septentrional y dificulta la importación regional de humedad desde la región amazónica debido a los vientos predominantes del oeste sobre el Altiplano

(el altiplano en Bolivia y parcialmente en Perú), lo que conduce a condiciones secas en el área de estudio. Hacia el verano austral, la corriente en chorro se debilita y se desplaza hacia el sur. Al mismo tiempo, la convección profunda que está conectada a la zona de convergencia intertropical (ITCZ) aumenta sobre el área amazónica. Este cambio en la corriente del chorro subtropical y la formación del Alta de Bolivia (debido al calentamiento latente del Amazonas y la liberación de calor sensible de los Andes) conducen a un aumento de los vientos del este y el transporte de humedad hacia el Altiplano dando lugar a precipitaciones sobre esta área, mientras que a escala interanual el ENSO da lugar a una precipitación reducida, principalmente en la parte sur de los Andes (Imfeld et al., 2021).

Otro fenómeno meteorológico a tener en consideración es la Oscilación de Madden-Julian (MJO, por sus siglas en inglés). Si bien se identificó que la MJO presenta influencia en el clima regional de Sudamérica durante todo el año, con grandes variaciones estacionales. En el verano austral, cuando la MJO está activa, se observan mayores probabilidades de precipitaciones intensas. Este patrón se ve favorecido por la divergencia de niveles superiores sobre Sudamérica tropical y una anomalía anticiclónica alrededor de su extremo sur, siendo ambas respuestas de circulación a la actividad de la MJO. Por otro lado, durante la estación seca en los trópicos, se observan menores probabilidades de precipitaciones superiores. Los resultados sugieren que la actividad del monzón sudamericano induce una gran modulación en la influencia de la MJO sobre la variabilidad climática en Sudamérica. También es evidente que, en todas las estaciones, las anomalías de circulación sobre Sudamérica, que a su vez inducen anomalías tanto de precipitación como de temperatura del aire en superficie, son una combinación entre la respuesta divergente de la circulación a la MJO con mayor amplitud en los trópicos y la respuesta rotacional de la circulación a la MJO, con mayor amplitud en la zona extratropical (Alvarez et al., 2016).

#### **2.4. Modelamiento numérico del tiempo**

La predicción numérica constituye un método fundamental para pronosticar las condiciones atmosféricas en un período de tiempo específico. Este enfoque se basa en el cálculo de la evolución del estado de la atmósfera, mediante la resolución de ecuaciones que rigen los procesos de la física atmosférica. Las ecuaciones principales que sustentan la predicción numérica del tiempo incluyen la ecuación de movimiento, la ecuación de continuidad, la ecuación de conservación de energía, la ecuación de conservación de vapor de agua y la ecuación de estado. Bauer et al. (2015) destacan que estas ecuaciones deben discretizarse en

el espacio y el tiempo para ser resueltas numéricamente, lo que determina qué procesos físicos puede abordar el modelo y cuáles requieren ser parametrizados.

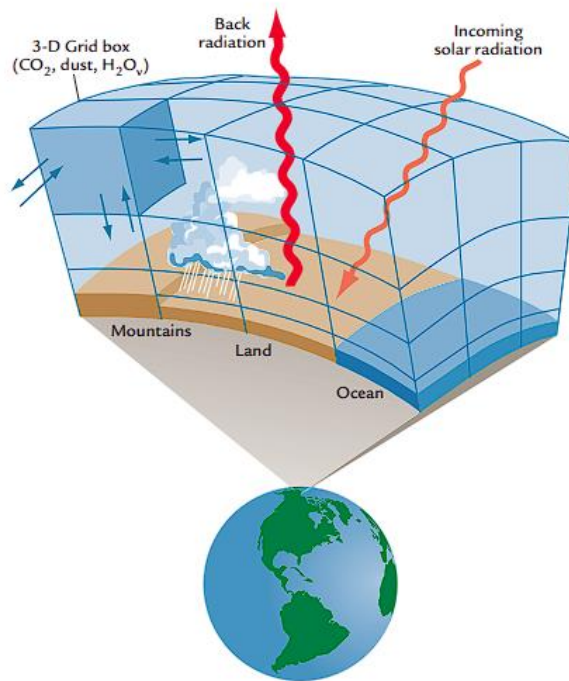
Procesos físicos como microfísica, condensación, evaporación y calentamiento/enfriamiento radiativo ocurren en escalas que no pueden ser resueltas explícitamente por el modelo, por lo que deben ser parametrizados en función de las escalas abordables por el modelo numérico. Esta parametrización optimiza la resolución numérica. Una parametrización adecuada es esencial para que los modelos numéricos representen con precisión múltiples procesos físicos a una escala de subgrilla.

La solución de las ecuaciones para predecir el estado futuro de la atmósfera se basa en condiciones iniciales proporcionadas al modelo. Estas condiciones iniciales son observaciones meteorológicas que reflejan el estado actual de la atmósfera. Kimura (2002) señala que los resultados de los modelos numéricos son altamente sensibles a estas condiciones iniciales debido a la naturaleza de las ecuaciones integradas. Pequeñas imprecisiones en las condiciones iniciales pueden amplificarse exponencialmente, lo que conduce a desviaciones significativas entre el estado pronosticado por el modelo numérico y el estado real de la atmósfera.

Collins et al. (2013) indican que las limitaciones de procesamiento y memoria en los recursos computacionales hacen que los modelos numéricos no puedan representar perfectamente la atmósfera, aproximándose a un conjunto finito de datos. Para lograr esto, utilizan grillas, un conjunto tridimensional y finito de puntos (Figura 2). La distancia entre estos puntos en el eje horizontal se denomina resolución horizontal. En los primeros modelos, esta resolución era de cientos de kilómetros, pero con el avance tecnológico y el desarrollo de los modelos, las resoluciones espaciales actuales se sitúan en el rango de decenas de kilómetros, mejorando significativamente la precisión de los modelos.

Dentro de los modelos numéricos del tiempo, se encuentran los modelos de mesoescala de área limitada, también conocidos como modelos regionales. Estos modelos predicen el futuro estado de la atmósfera en un área limitada de la Tierra, a diferencia de los modelos globales que abarcan toda la atmósfera terrestre. Una diferencia clave es que los modelos globales tienen límites solo en la parte superior e inferior de la atmósfera, mientras que los modelos regionales también tienen límites en los bordes del área de interés (Golding, 2015). Aunque los modelos regionales tienen una mejor resolución espacial que los modelos globales y requieren menos recursos computacionales para su ejecución, generalmente

dependen de la información que proporcionan los modelos globales para utilizarla como condiciones iniciales. Los modelos regionales incorporan parametrizaciones distintas para representar adecuadamente los procesos físicos (Anthes et al., 1989). La implementación de modelos regionales con propósitos operativos se basa en la necesidad de que los pronósticos numéricos capturen las características locales que influyen en la evolución atmosférica en el área de interés.



**Figura 2:** Representación esquemática de grillas horizontales y verticales incluidas en modelos numéricos de predicción del tiempo.

**FUENTE:** Ruddiman, 2013

## 2.5. Modelo Global Forecast System

El Modelo Global Forecast System (GFS) es una herramienta diseñada para la predicción meteorológica, desarrollada y gestionada por el National Centers for Environmental Prediction (NCEP) de los Estados Unidos. Este modelo se basa en las ecuaciones dinámicas primitivas y además incorpora diversas parametrizaciones para una representación adecuada de los procesos físicos atmosféricos (Yang et al., 2006).

En la actualidad, el modelo GFS se encuentra en su versión 16 y tiene la capacidad de generar pronósticos con un horizonte de proyección de hasta 16 días. El sistema realiza cuatro

corridas diarias del modelo, programadas para los momentos de las 00:00Z, 06:00Z, 12:00Z y 18:00Z.

## **2.6. Posprocesamiento estadístico**

Los modelos numéricos presentan errores que provienen de diversas fuentes, como la física subyacente en el modelo, los esquemas numéricos empleados, condiciones iniciales incorrectas, topografía compleja y las limitaciones ligadas a la incertidumbre inherente en la atmósfera. Estos desafíos resaltan la relevancia del posprocesamiento estadístico, que involucra un conjunto de técnicas estadísticas destinadas a mejorar los resultados obtenidos del modelado numérico al relacionar estadísticamente las predicciones con datos meteorológicos observados. Marzban et al. (2006) indican que entre las técnicas clásicas y ampliamente empleadas de posprocesamiento estadístico en el contexto de modelos meteorológicos dinámicos, se encuentran los enfoques clásicos, conocidos como Model Output Statistics (MOS) y perfect prog, los cuales recurren a ecuaciones de regresión. MOS y perfect prog son enfoques distintos para generar pronósticos meteorológicos. MOS se destaca por su rendimiento superior en comparación con perfect prog en términos de sesgo, varianza del error y error cuadrático medio. MOS utiliza relaciones estadísticas entre el predictando y las variables observadas, ajustándolas a la salida real del modelo numérico, lo que le permite ofrecer pronósticos más precisos con menos sesgo y varianza. En contraste, perfect prog se basa en relaciones entre el predictando y las variables observadas en momentos casi concurrentes, sin ajuste a la salida del modelo numérico. Aunque no está limitado por la disponibilidad de datos del modelo, los pronósticos de perfect prog tienden a tener sesgo y una mayor varianza en comparación con MOS.

Wilks (2011), destaca las principales razones que motivan la reinterpretación estadística o el posprocesamiento de las salidas de los modelos dinámicos, que incluyen:

- Diferencias sustanciales entre el mundo real y su representación en los modelos dinámicos, lo que introduce limitaciones en la capacidad predictiva.
- Los modelos dinámicos no reflejan con precisión el comportamiento atmosférico y sus condiciones iniciales inevitablemente difieren del estado real inicial de la atmósfera, lo que conlleva a errores en las predicciones. Si estos errores son sistemáticos, el posprocesamiento estadístico puede ayudar a mitigar y corregir los sesgos en los pronósticos.



- A pesar de que los modelos dinámicos son deterministas, la combinación de información proveniente de los pronósticos dinámicos junto con métodos estadísticos posibilita cuantificar y expresar la incertidumbre asociada a diferentes situaciones de pronóstico. De este modo, es factible generar pronósticos probabilísticos al utilizar predictores extraídos de una única integración dinámica determinista.

### **2.6.1. Enfoque Model Output Statistics**

El enfoque de posprocesamiento estadístico conocido como Model Output Statistics es uno de los primeros métodos desarrollados para mejorar las salidas de los modelos numéricos del tiempo. Fue originalmente descrito por Glahn y Lowry (1972), quienes explicaron que el principio fundamental del MOS radica en establecer relaciones estadísticas entre las variables a predecir (predictandos) y las variables predictoras derivadas de las salidas de modelos numéricos para un período de tiempo específico.

MOS tuvo sus primeras aplicaciones en el pronóstico operativo de diversas variables, como probabilidad de precipitación, viento en superficie, temperatura máxima, cantidad de nubes y probabilidad condicional de precipitación sólida. Ejemplos de aplicaciones tempranas incluyen el uso de MOS en la predicción de la temperatura mínima en las regiones alpinas de Italia, con orografía compleja (Eccel et al., 2007), y en la predicción diaria de temperatura máxima y mínima en ciudades de Australia utilizando regresión lineal múltiple (Woodcock, 1984). Otro ejemplo es su empleo en Indonesia, donde se utilizaron componentes principales de las variables predictoras de los modelos numéricos junto con técnicas de regresión ridge para pronosticar la temperatura máxima y mínima en 4 estaciones meteorológicas (Qona'ah et al., 2020).

En el contexto de Perú, De la Cruz (2016) empleó el enfoque MOS para desarrollar modelos de regresión logística destinados a la generación de pronósticos de precipitación diaria expresados en probabilidades para 40 estaciones meteorológicas distribuidas en todo el país, abarcando las tres regiones: costa, sierra y selva, que presentan condiciones meteorológicas y climáticas altamente diversas.

En general, la aplicación de técnicas de pronóstico estadístico del tiempo ha demostrado resultados favorables en la búsqueda de pronósticos más eficientes y precisos. Un antecedente adicional, aunque no emplea directamente las salidas de modelos numéricos,

demuestra la eficacia del pronóstico estadístico. Aliaga-Nestares et al. (2023), desarrollaron modelos empíricos utilizando regresión lineal múltiple para prever la temperatura máxima y mínima en Lima, Perú. Estos modelos superaron tanto el pronóstico numérico como el pronóstico operativo realizado por el servicio meteorológico nacional, a pesar de emplear solo tres predictores en cada modelo.

## 2.7. Análisis de Componentes Principales

El Análisis de Componentes Principales (PCA), también conocido en el campo de las ciencias atmosféricas como Funciones Empíricas Ortogonales (EOF), es una técnica estadística de reducción de dimensionalidad ampliamente utilizada y de larga data en diversas áreas de estudio. Esta técnica se basa en transformar, mediante combinaciones lineales, las variables originales (que a menudo se encuentran interrelacionadas) de un conjunto de datos en un nuevo conjunto de variables llamadas componentes principales. Estas componentes principales carecen de correlación entre sí, capturan la mayor variabilidad posible del conjunto original de datos y se organizan de manera que las primeras retengan la mayor parte de la variación de las variables originales (Jolliffe y Cadima, 2016).

En términos matemáticos, dada una matriz de datos original con un número determinado de observaciones para  $K$  variables, representada como el vector  $X'$ , el PCA busca encontrar  $M$  vectores  $u$ . Cada vector  $u$  es una combinación lineal de las variables originales, como se muestra en la ecuación 1:

$$u_m = e_m^T X' = \sum_{k=1}^K e_{km} X'_k, \quad m = 1, 2, \dots, M$$

Ec.1

Donde:

$u_m$ :  $m$  – ésima componente principal

$e_m$ :  $m$  – ésimo eigen-vector

$X'_k$ : vector con las observaciones correspondientes a la  $k$  – ésima variable del conjunto original de datos

Una representación matricial válida que expresa la transformación del conjunto de datos en las componentes principales, utilizando todas las componentes obtenidas, se logra mediante

la matriz ortogonal [E]. Las columnas de esta matriz son los vectores propios, y se expresa con la ecuación 2:

$$u = [E]^T X'$$

Ec.2

Donde:

[E]: matriz cuadrada de eigen-vectores con K filas y columnas

u: vector que contiene K componentes principales

X': vector conjunto de datos al cual se aplica la técnica PCA

A partir de la ecuación 2, podemos deducir la ecuación 3, que permite la reconstrucción de los datos a partir de las componentes principales:

$$X' = [E] u$$

Ec.3

Al emplear la ecuación 3, la reconstrucción de los datos es exacta si se utilizan todas las componentes principales. Sin embargo, si se emplea un número menor de componentes para la reconstrucción ( $M < K$ ), donde M corresponde a las componentes principales con la mayor varianza explicada, la reconstrucción se convierte en una aproximación (Wilks, 2011).

Es común emplear el PCA como preprocesamiento para compactar un conjunto de datos en términos de sus componentes principales, que luego pueden utilizarse en diversas aplicaciones, como regresión lineal, redes neuronales, análisis discriminante lineal, entre otros (Bro y Smilde, 2014).

## **2.8. Regresión de componentes principales**

La Regresión de Componentes Principales (PCR, por sus siglas en inglés) es una variante de la regresión lineal en la cual se utilizan las componentes principales de las variables predictoras como regresores. Según (Jolliffe, 1982), la utilización de PCR contribuye a la ortogonalización del problema de regresión, reduce el gasto computacional y disminuye la inestabilidad en las relaciones establecidas. Otro beneficio del uso de PCR radica en que, debido a que los predictores son no correlacionados entre sí, es posible agregar o eliminar

predictores basados en componentes principales de la ecuación de regresión sin afectar las contribuciones de los demás.

A pesar de sus ventajas, existen ciertas limitaciones asociadas con la PCR. Una de ellas es la necesidad de comprender adecuadamente la interpretación de la regresión, lo cual requiere una interpretación previa de las componentes principales utilizadas como predictores, en relación con el contexto del problema. Además, no hay garantía de que las componentes principales empleadas como predictores sean las que proporcionen la mejor predicción. En ocasiones, las componentes con menor varianza pueden estar más fuertemente relacionadas con la variable a predecir (Wilks, 2011).

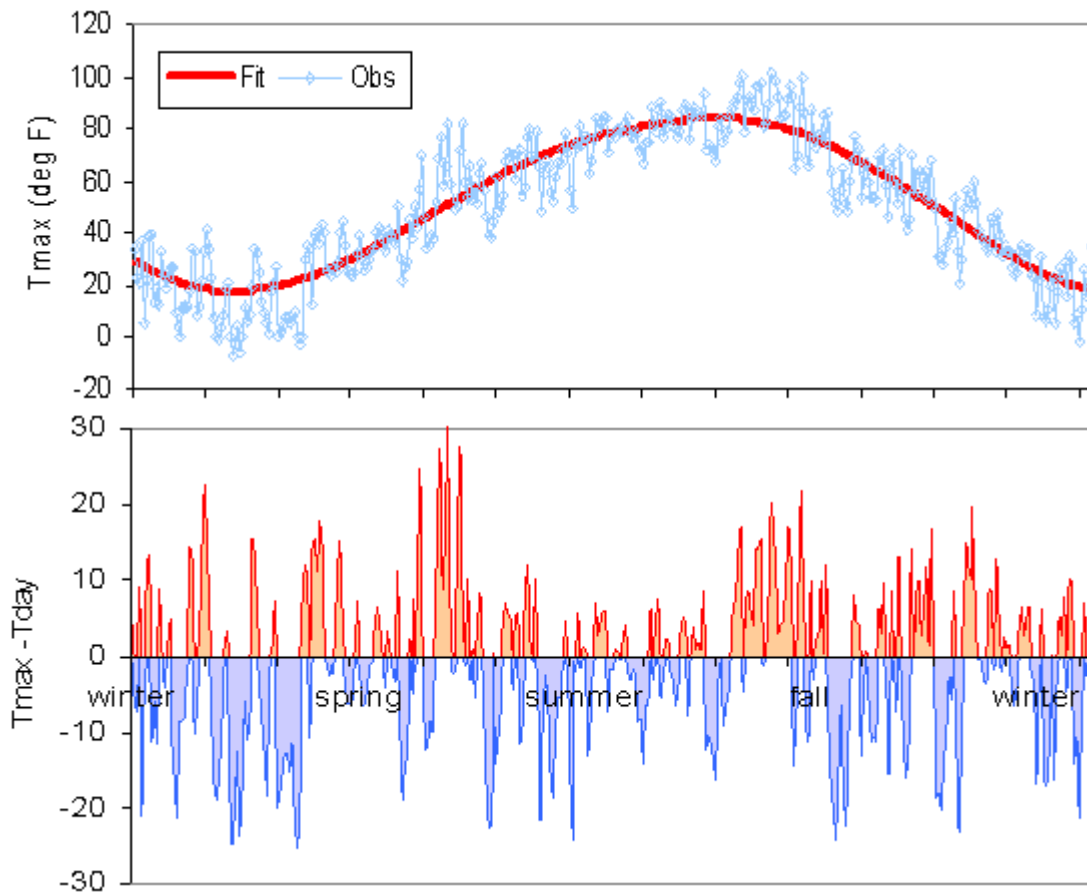
Según Deduy (2019), uno de los desafíos al aplicar PCR es la selección de las mejores variables predictoras para el problema de regresión. Para abordar este desafío, es posible recurrir a métodos clásicos de selección de variables, los cuales buscan encontrar el mejor subconjunto de variables entre todas las combinaciones posibles. Algunos de los algoritmos clásicos de selección de variables aplicables en PCR, mencionados por Navarro (2009), son los métodos de selección escalonada Forward Selection, Backward Selection y Stepwise Selection.

## **2.9. Variabilidad de alta y baja frecuencia en series de tiempo**

Una serie de tiempo es un conjunto de observaciones realizadas sistemáticamente a lo largo del tiempo, que puede estar influenciado por diversas fuentes de variación, como efectos estacionales, tendencias, cambios cíclicos y fluctuaciones irregulares. Una serie temporal puede dividirse en dos componentes: la variabilidad de alta frecuencia y la variabilidad de baja frecuencia. Cuando se busca obtener valores suavizados a partir de una serie de tiempo, el objetivo es eliminar las fluctuaciones locales representadas por la variabilidad de alta frecuencia. Para lograrlo, se aplica un filtro conocido como filtro paso bajo. Por otro lado, en situaciones en las que se desea eliminar las fluctuaciones a largo plazo representadas por la variabilidad de baja frecuencia, se recurre a un filtro paso alto o "high-pass filter" (Chatfield, 1975).

La Figura 3 muestra la descomposición de una serie de tiempo de temperatura máxima diaria en sus componentes de variabilidad de alta y baja frecuencia. Esta descomposición se logra

restando el filtro paso bajo a la serie de tiempo original, lo que resulta en la variabilidad de alta frecuencia a nivel diario.



**Figura 3:** Comparación de una serie de tiempo de temperatura máxima diaria con filtro paso bajo (figura superior) y una serie de tiempo representando la variabilidad de alta frecuencia (figura inferior)

FUENTE: NOAA (s/f)

## 2.10. Árboles de clasificación

Usualmente los árboles de clasificación presentan la limitante de sobreajuste respecto a los datos con los cuales fueron construidos; por lo cual, su capacidad de generalización y predicción es bastante limitada. Este problema se da porque los árboles de clasificación crecen más de lo necesario, para lo cual luego de la construcción necesitan ser podados hasta alcanzar un tamaño ideal, en búsqueda de evitar el problema de sobreajuste y definir la complejidad que presentará el árbol de decisión, dado que según Rokach y Maimon (2006) el método y el criterio de interrupción empleados para la poda, definen directamente la complejidad del árbol. A pesar de su importancia en el diseño de árboles de clasificación individuales, el proceso de poda se omite en la construcción de modelos Random Forests.

Los árboles de clasificación se construyen mediante particiones binarias sucesivas de los datos, tal como se ilustra en los ejemplos de árboles de clasificación mostrados en la figura 4. Un árbol de clasificación consta de una raíz que inicia el proceso de partición, dividiendo los nodos en nodos no terminales (que se subdividen aún más) o nodos terminales (que no se subdividen). Los nodos no terminales se dividen en dos nodos hijos, en función de un punto de división correspondiente a uno de los predictores. Esta división se realiza de manera que las observaciones con valores de predictor menores que el punto de división se asignan a la rama izquierda, mientras que las demás se ubican en la rama derecha (Cutler et al., 2012).

Para determinar cómo dividir un nodo, se evalúan diferentes divisiones posibles en cada variable predictora y se elige la división "óptima" según algún criterio. Según Tan et al. (2005), los indicadores de calidad de división se basan en la impureza de los nodos descendientes. Entre estos indicadores, el índice Gini es comúnmente utilizado, como se muestra en la ecuación 4.

$$Gini(t) = 1 - \sum_{i=0}^{c-1} [p(i|t)]^2$$

Ec.4

Donde:

$c$ : número de clases

$p(i|t)$ : fracción de observaciones correspondientes a la clase  $i$  en el nodo  $t$

Sin embargo, los árboles de clasificación suelen enfrentar el problema del sobreajuste, adaptándose en exceso a los datos de entrenamiento. Para abordar este problema, es necesario podar el árbol después de su construcción, ajustando su tamaño y complejidad para mejorar su capacidad de generalización y predicción.

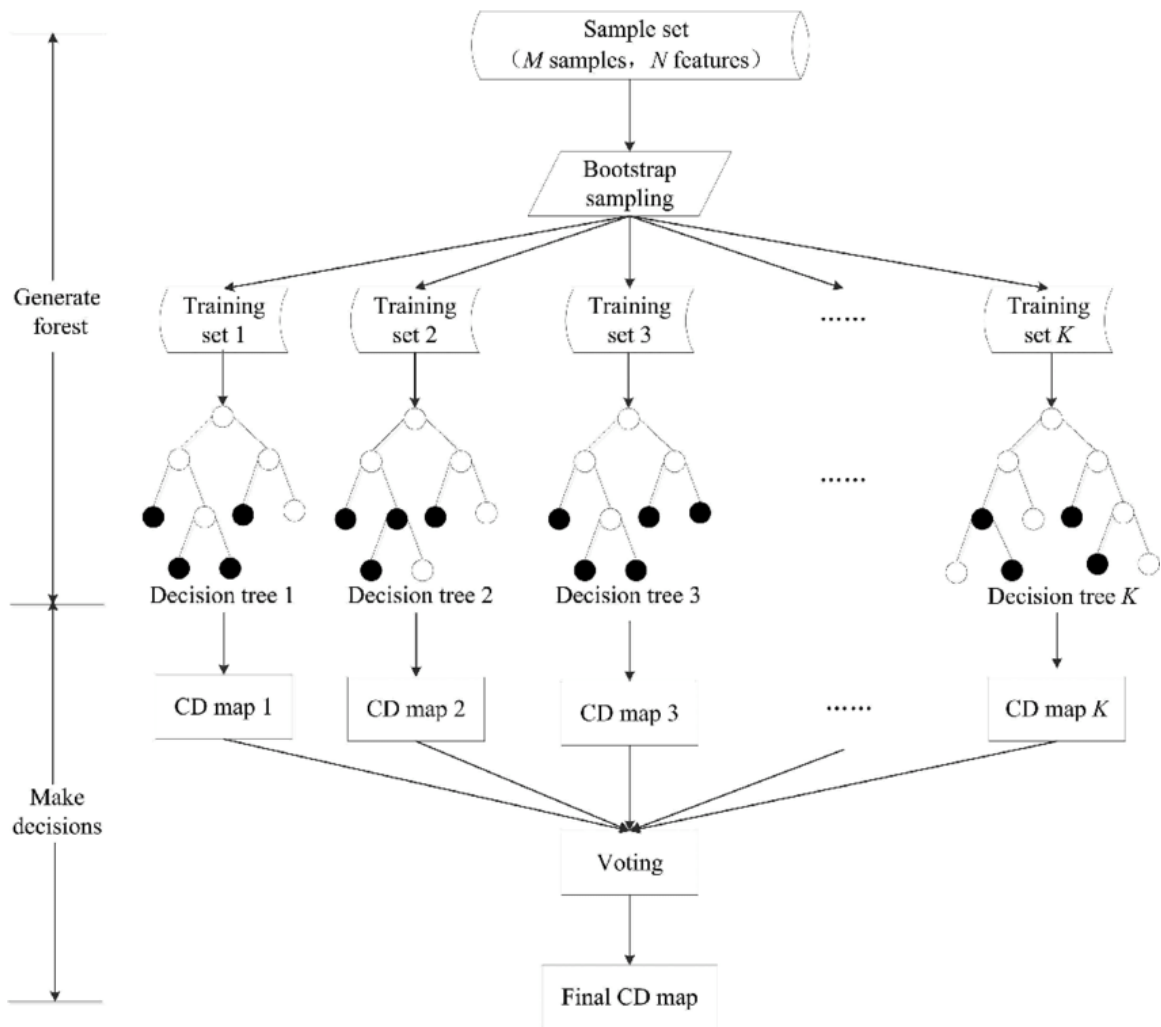
## 2.11. Random Forests

La técnica de Random Forest constituye un algoritmo de aprendizaje automático empleado principalmente en tareas de clasificación y regresión. La idea fundamental en la que se basa Random Forests fue propuesta por Ho (1995), quien sugirió la generación de múltiples árboles de decisión en subespacios seleccionados aleatoriamente del espacio de características. Posteriormente, Breiman (2001) desarrolló y patentó el algoritmo Random Forest definiéndolo como una combinación de árboles predictores, en la cual cada árbol depende de los valores de un vector aleatorio muestreado de forma independiente y con la

misma distribución para todos los árboles del bosque (conjunto total de árboles que componen el modelo). Como se mencionó previamente, los árboles de decisión tienden a ser propensos al sobreajuste, ajustándose en exceso al ruido en lugar de capturar las relaciones fundamentales presentes en el conjunto de entrenamiento. En contraste, Random Forests logra reducir significativamente el error de sobreajuste generando un conjunto diversificado de árboles de decisión que no están fuertemente correlacionados entre sí. Utilizar el mismo conjunto de datos para construir cada árbol resultaría en un bosque de árboles idénticos, lo que no aportaría ventajas respecto a un solo árbol. Por lo tanto, se aplican dos procesos adicionales conocidos como empaquetado de árboles (tree bagging) y empaquetado de características (feature bagging), para producir árboles únicos que compongan el modelo Random Forests (Herman y Schumacher, 2018).

De manera simplificada, el proceso detrás del algoritmo de Random Forests implica la creación de un bosque que consta de múltiples árboles de decisión. Cada árbol se construye a partir de diferentes conjuntos de datos obtenidos mediante un muestreo aleatorio con reemplazo de las observaciones originales. A pesar de la aleatoriedad introducida mediante este proceso, el sobreajuste puede ser un problema latente. Para evitar este problema, se introduce aleatoriedad al considerar solo un subconjunto de las características disponibles al tomar decisiones en cada nodo. Este subconjunto varía en tamaño y contiene una selección aleatoria de las características originales. Al combinar las predicciones de todos estos árboles individuales, ya sea por votación o ponderación, se obtiene el resultado final del modelo. Este enfoque asegura que las predicciones sean robustas y menos propensas al sesgo y sobreajuste (Cutler et al., 2012; Herman y Schumacher, 2018).

Dado que el muestreo realizado mediante bootstrapping no utiliza todas las observaciones del conjunto de datos de entrenamiento, este conjunto de observaciones no utilizado en la construcción de los árboles se conoce como "Out-Of-Bag Data", el cual se emplea en el cálculo del error del modelo.



*Figura 4: Diagrama de flujo de un modelo Random Forests*

**FUENTE:** Feng et al., 2018



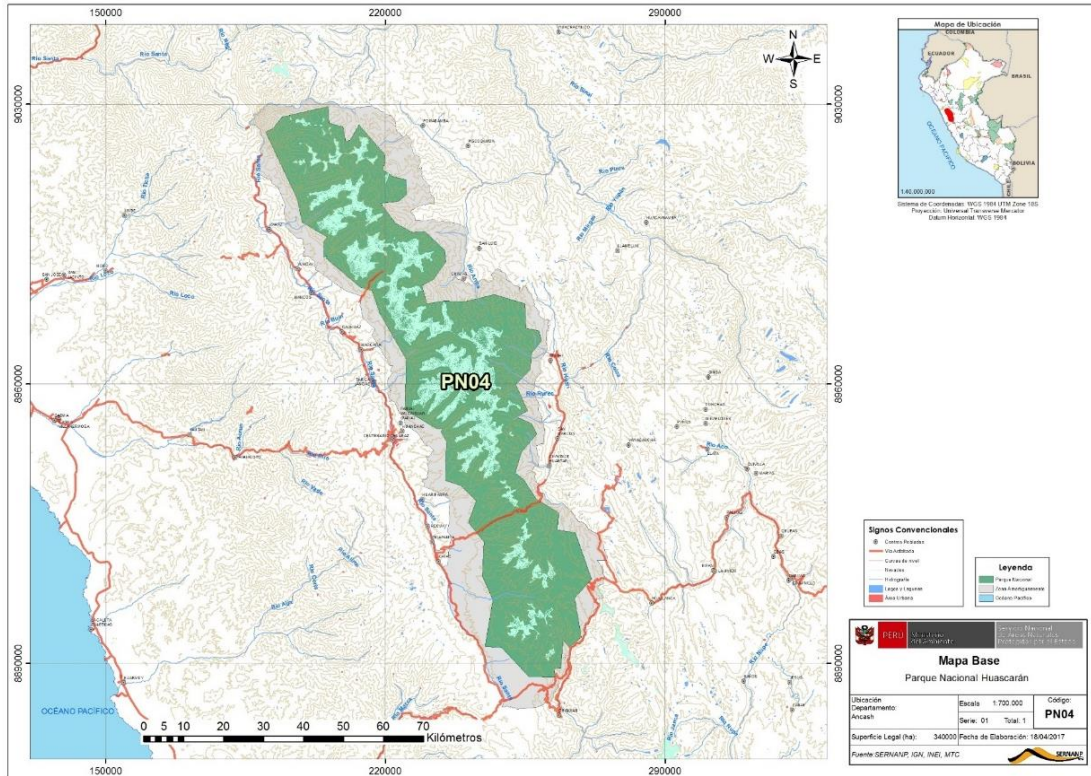
### **III. METODOLOGÍA**

#### **3.1. Área de estudio**

La presente investigación se desarrolló en el Parque Nacional Huascarán (PNH) y sus áreas circundantes, como se ilustra en la Figura 5. Ubicado en la sierra central de la región Áncash, el PNH es reconocido como uno de los enclaves de conservación más emblemáticos del Perú, destacando por su riqueza biológica, belleza paisajística y valor cultural. Su extensión abarca alrededor de 340,000 ha repartidas entre las provincias de: Huaylas, Yungay, Carhuaz, Huaraz, Recuay, Bolognesi, Pomabamba, Huari, Mariscal Luzuriaga y Asunción. La altitud varía entre los 2,400 y los 6,768 m s.n.m., conformando un terreno abrupto en el cual se encuentran glaciares, lagunas y ríos que se entrelazan en las cuencas de los ríos Santa, Marañón y Pativilca. Los servicios ecosistémicos del PNH están principalmente orientados hacia el suministro de recursos hídricos y atributos paisajísticos. Dentro de este paraje se llevan a cabo diversas actividades como la agricultura, pastoreo, minería, turismo y proyectos de embalses para la generación de energía hidroeléctrica.

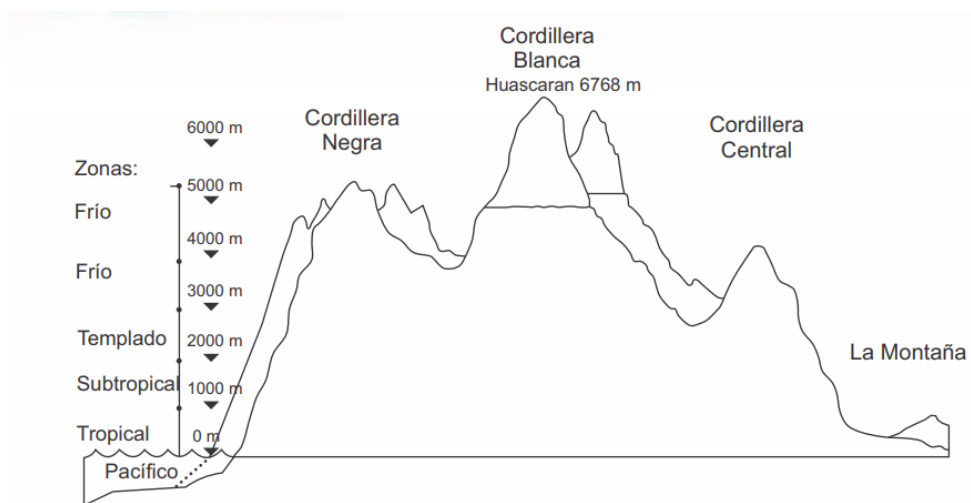
Desde una perspectiva geomorfológica, las partes alta y media del PNH se caracterizan por la presencia de valles interandinos, mientras que en la región correspondiente a la Cordillera Blanca (la cadena montañosa tropical más alta del mundo, con aproximadamente 180 km de longitud y orientación Norte-Sur) se identifican valles, quebradas, altiplanos y glaciares, estos últimos situados entre los 4,800 y 6,768 m s.n.m. Al oeste de la Cordillera Blanca se localiza la Cordillera Negra, que se diferencia por su menor altitud (con su pico más alto a 5,187 m s.n.m.), menores niveles de precipitación y la ausencia de nieve. La Figura 6 proporciona una visualización de las características altitudinales de la región de estudio, incluyendo la disposición espacial de la Cordillera Negra y la Cordillera Blanca mediante un corte transversal de la Cordillera de los Andes. Es relevante mencionar el Callejón de Huaylas, un valle aluvial enclavado entre las Cordilleras Negra y Blanca, que alberga la ciudad de Huaraz, capital de la región Áncash y un importante destino turístico. En cuanto al clima, el PNH experimenta dos estaciones bien definidas: una estación fría con niveles

reducidos de precipitación entre los meses de mayo y octubre, y un período cálido y lluvioso que abarca de noviembre a abril (con picos de precipitación en enero y marzo).



**Figura 5:** Mapa del Parque Nacional Huascarán

**FUENTE:** SERNANP, 2019



**Figura 6:** Corte transversal de la cordillera de los Andes en Perú

**FUENTE:** Sevink, 2009

### 3.2. Información observada de temperaturas extremas diarias

Para el desarrollo de los modelos de pronóstico de temperaturas extremas, se recopiló información diaria de temperatura mínima y temperatura máxima de estaciones meteorológicas pertenecientes a SENAMHI, a la Universidad Nacional Santiago Antúnez de Mayolo (UNASAM) y a la Autoridad Nacional del Agua (ANA) ubicadas dentro del área de estudio (Figura 7). Los detalles de las coordenadas geográficas y la elevación con respecto al nivel del mar de estas estaciones se encuentran enlistados en la Tabla 1.

**Tabla 1:** Ubicación y altitud de estaciones con información meteorológica de temperaturas extremas empleadas para la investigación

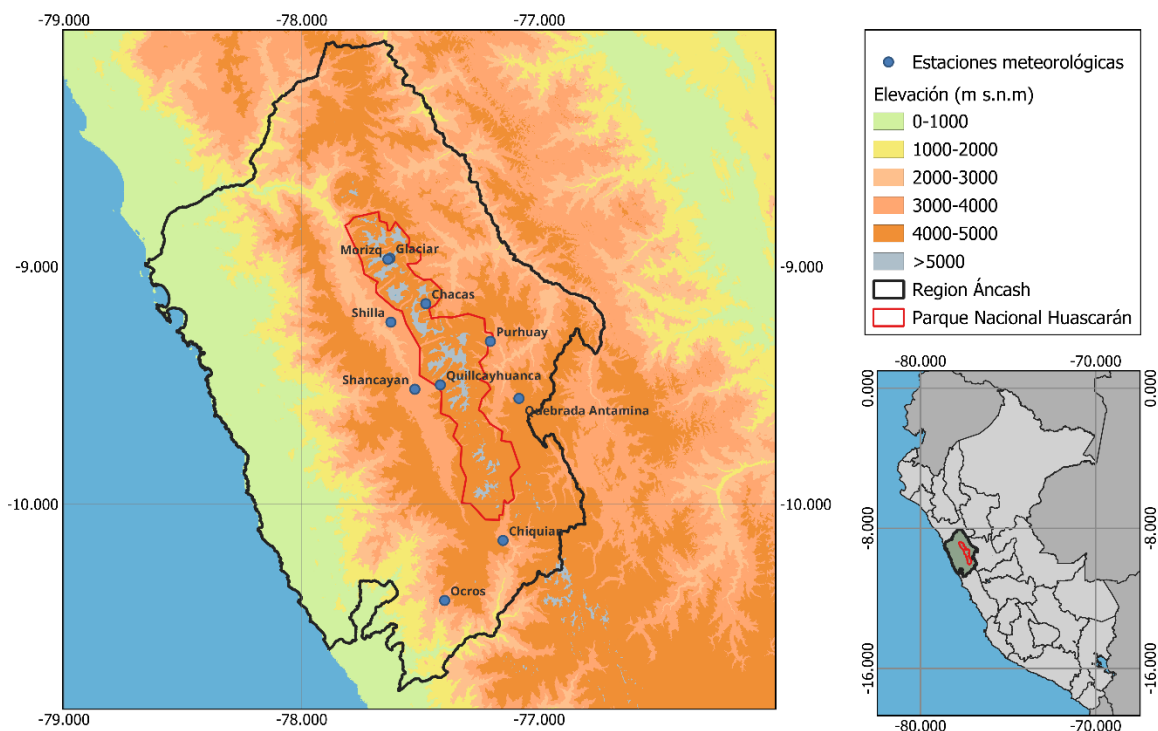
<b>Estación</b>	<b>Longitud (°)</b>	<b>Latitud (°)</b>	<b>Altitud (m s.n.m.)</b>
MORIZQ	-77.6371	-8.9699	4,814
GLACIAR	-77.6287	-8.9652	4,918
CHACAS	-77.4779	-9.1564	4,040
SHILLA	-77.6248	-9.2341	3,058
PURHUAY	-77.2061	-9.3148	3,499
QUILLCAYHUANCA	-77.4166	-9.4981	3,817
SHANCAYAN	-77.5240	-9.5165	3,078
CHIQUIAN	-77.1528	-10.1527	3,412
OCROS	-77.3985	-10.4050	3,253
QUEBRADA ANTAMINA	-77.0857	-9.5550	4,042

Los registros meteorológicos abarcaron el periodo comprendido entre enero de 2015 y diciembre de 2018. En el caso de la temperatura mínima diaria, se consideró la temperatura más baja registrada entre las 00:00 Z y las 12:00 Z (7:00 p.m. hora local del día previo hasta las 7:00 a.m. hora local del día actual). Por otro lado, los datos de temperatura máxima diaria correspondieron a la temperatura más alta registrada entre las 12:00 Z y las 00:00 Z (7:00 a.m. hora local hasta las 7:00 p.m. hora local) de una fecha determinada.

Se realizó un exhaustivo control de calidad de los datos de temperaturas extremas diarias provenientes de las estaciones meteorológicas, teniendo en cuenta las variables de interés disponibles. Durante este proceso, se verificó la coherencia física de los datos y se eliminaron registros que presentaban información errónea. Una vez concluido este procedimiento, se seleccionó un conjunto de estaciones que ofrecían datos de alta calidad. Esta elección se

basó en criterios de consistencia y en la presencia de un periodo de mediciones uniforme, evitando días con escasez de información sobre temperaturas extremas.

Para abordar vacíos de información en los registros de las estaciones meteorológicas seleccionadas, se emplearon técnicas de imputación de datos con el propósito de completar los vacíos y garantizar la integridad de las series temporales utilizadas en el análisis.



**Figura 7:** Mapa con ubicación de las estaciones meteorológicas empleadas para la investigación

### **3.3. Información observada de precipitación acumulada diaria**

Para el desarrollo de los modelos de pronóstico de probabilidad de precipitación, se utilizaron datos observados provenientes del conjunto de datos Peruvian Interpolated data of the SENAMHI's Climatological and Hydrological Observations (PISCO, por sus siglas en inglés). Específicamente, se utilizó el producto PISCO de precipitación (PISCOp), el cual fue documentado por Aybar et al. (2020). La base de datos PISCOp se obtuvo al combinar información proveniente de estaciones meteorológicas en superficie, climatologías, reanálisis e información satelital de estimación de precipitación.

La combinación de estas fuentes de datos permitió generar una base de datos grillada de alta resolución espacial que abarca todo el territorio peruano. Esta base de datos cubre un período temporal que va desde el 1 de enero de 1981 hasta el 31 de diciembre de 2016. Aunque existen versiones de PISCOp con diferentes resoluciones temporales (diaria y mensual), en este estudio se utilizó la información a nivel diario, es decir, los valores acumulados de precipitación durante un período de 24 horas, desde las 12 UTC de un día hasta las 12 UTC del día siguiente.

Es importante señalar que, si bien la fecha límite del conjunto de datos descrito es el 31 de diciembre de 2016, se utilizó una versión de PISCOp con fines operativos que se actualiza de manera constante. Esta versión actualizada fue proporcionada por el SENAMHI específicamente para el desarrollo de la presente investigación.

La información detallada de PISCOp, incluyendo su acceso y descarga, se encuentra disponible en el siguiente enlace:

<https://iridl.ldeo.columbia.edu/SOURCES/.SENAMHI/.HSR/.PISCO/.Prec/.v2p1/?Set-Language=es>

### **3.4. Información de salidas del modelo GFS**

Los datos de modelo numérico utilizados como predictores en el desarrollo de los modelos de pronóstico (tanto para temperaturas extremas como para probabilidad de precipitación) corresponden al modelo global de predicción del tiempo GFS 0.25°. El modelo GFS 0.25° realiza cuatro ejecuciones diarias a las 00:00 UTC, 06:00 UTC, 12:00 UTC y 18:00 UTC. Para este estudio, se emplearon exclusivamente las ejecuciones correspondientes a las 00:00 UTC, con un horizonte de pronóstico de hasta 84 horas. Esta elección se fundamenta en el

objetivo de generar pronósticos con anticipación de 1, 2 y 3 días, con una resolución temporal de 6 horas.

La información del modelo GFS utilizada se obtuvo del NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive (NCEP, 2015). Estos registros son pronósticos previos generados por el modelo GFS y están disponibles en el portal web del Research Data Archive (RDA). El periodo abarca desde 2015 hasta 2021, con una excepción: para el pronóstico de temperaturas extremas se limitó el periodo a 2015-2018 debido a la disponibilidad de datos observados. La información histórica del GFS 0.25° se encuentra en el enlace siguiente: <https://rda.ucar.edu/datasets/ds084.1/>

Dado que el modelo GFS tiene un alcance global, se procesó la información para abarcar un dominio geográfico correspondiente a las coordenadas entre 22°S y 2° N de latitud, y entre 85°W y 64°W de longitud. Posteriormente, se extrajeron las variables sinópticas pertinentes en niveles de presión significativos, según lo especificado en la Tabla 2.

Además, se procesó la información del modelo numérico para extraer las temperaturas extremas diarias, es decir, las temperaturas máximas y mínimas que el modelo GFS pronostica para los periodos de 24, 48 y 72 horas. La generación de la variable temperatura mínima implicó la extracción de la menor temperatura pronosticada para cada día, mientras que para la variable temperatura máxima se seleccionó la mayor temperatura pronosticada para cada día. Estas temperaturas extremas diarias fueron esenciales para realizar comparaciones y evaluaciones con las predicciones generadas por los modelos empíricos desarrollados en este estudio, así como para determinar la precisión de las predicciones directas del modelo GFS sin ningún procesamiento adicional. Aunque esta información no se empleó en la construcción de los modelos de pronóstico, sirvió para comparar las predicciones directas de temperaturas extremas del modelo GFS sin procesamiento adicional, con las predicciones generadas por los modelos empíricos desarrollados en este estudio.

**Tabla 2:** Variables del modelo GFS con su respectivo nivel vertical

<b>Variables atmosféricas</b>	<b>Niveles</b>
Temperatura (°C)	500 hPa, 400 hPa, 200 hPa
Humedad relativa (%)	500 hPa, 400 hPa, 200 hPa
Componente zonal del viento (m/s)	500 hPa, 400 hPa, 200 hPa
Componente meridional del viento (m/s)	500 hPa, 400 hPa, 200 hPa
Humedad Específica (Kg/Kg)	500 hPa, 400 hPa, 200 hPa
Temperatura a 2m del suelo (°C)	Variable en superficie
Humedad relativa a 2m del suelo (%)	Variable en superficie
Humedad especifica a 2m del suelo (Kg/Kg)	Variable en superficie
Presión reducida a nivel del mar (Pa)	Variable en superficie
Precipitación (mm)	Variable en superficie
Agua Precipitable (Kg/m <sup>2</sup> )	Variable en toda la columna de atmósfera

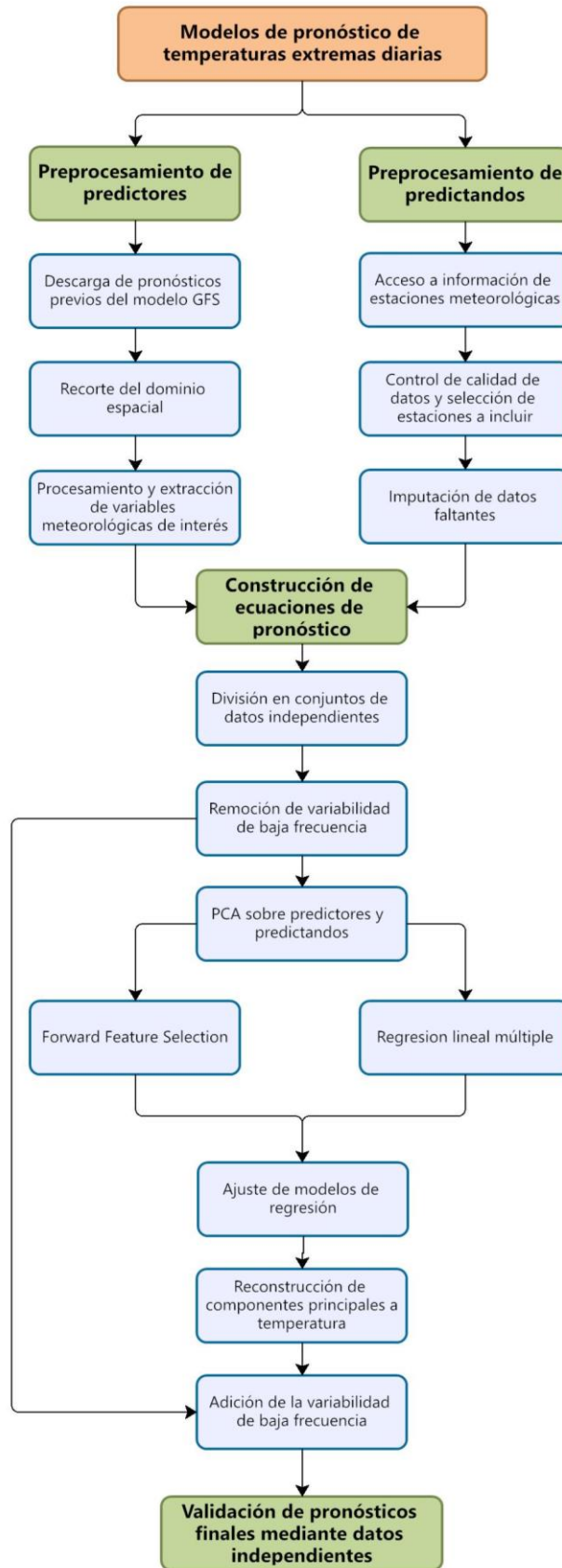
A continuación, se detallan los procedimientos desarrollados para lograr los objetivos establecidos en el marco de esta investigación. Se adoptaron enfoques metodológicos diferentes para desarrollar los modelos de pronóstico de temperaturas extremas diarias y la probabilidad de precipitación. El pronóstico de las temperaturas extremas diarias se abordó como un problema de regresión, mientras que el pronóstico de la probabilidad de precipitación se trató como un problema de clasificación. Para evaluar los modelos empíricos de pronóstico desarrollados en función de la variable objetivo, se emplearon métricas estadísticas adecuadas. Finalmente, se llevaron a cabo pruebas de pronóstico utilizando conjuntos de datos de prueba para validar la efectividad y aplicabilidad de los modelos desarrollados en tareas de pronóstico. En las Figura 8 y 10 se presenta el esquema

metodológico empleado en el desarrollo de los modelos de pronóstico de temperaturas extremas y probabilidad de precipitación diaria, respectivamente.

### **3.5. Metodología para pronóstico de temperaturas extremas diarias**

Para el pronóstico de las temperaturas extremas diarias, se generaron ecuaciones de regresión lineal múltiple enfocadas en la variabilidad de alta frecuencia de los predictores y predictandos. La Figura 8 ilustra de manera esquemática los procedimientos aplicados para obtener los modelos de pronóstico. Este enfoque metodológico se aplicó tanto para las temperaturas máximas como para las temperaturas mínimas.





**Figura 8:** Esquema de la metodología de investigación para pronóstico de temperaturas extremas

### **3.5.1. Control de calidad e Imputación de datos**

Una vez recopilada la información meteorológica observada, el primer paso consistió en llevar a cabo un riguroso proceso de control de calidad para asegurar la consistencia y representatividad de los datos en relación con la zona de estudio. El proceso de control de calidad de los datos se basó en una serie de técnicas diseñadas para identificar y abordar posibles problemas en los registros de temperatura. Estas técnicas incluyeron la aplicación de límites duros y blandos para ciertas variables meteorológicas, permitiendo detectar y eliminar valores anómalos y erróneos que podrían tener un impacto negativo en la calibración de los modelos de predicción.

Además de establecer límites, se llevaron a cabo pruebas de coherencia temporal para verificar que los datos de temperatura mantuvieran patrones lógicos a lo largo del tiempo, evitando cambios bruscos asociados a factores externos. Asimismo, se realizaron pruebas de coherencia espacial para asegurarnos de que los datos de temperatura se alinearan con las características geográficas de la zona de estudio, tomando en consideración factores como la altitud y la proximidad a estaciones vecinas.

Un componente esencial del proceso fue la inspección visual de los datos de temperatura. Esto permitió identificar posibles anomalías o irregularidades que podrían no haber sido detectadas por las técnicas automatizadas de control de calidad. En conjunto, estas técnicas garantizaron la integridad y confiabilidad de los datos de temperatura utilizados en el análisis.

Tras el proceso de control de calidad, se procedió a abordar los vacíos de información en los registros de las estaciones meteorológicas. Si bien los días con datos faltantes no representaban un porcentaje significativo en relación al total de mediciones, se emplearon técnicas de imputación de datos basadas en regresión lineal con información proveniente de estaciones cercanas. Esta imputación permitió completar los registros y aseguró que el conjunto de datos utilizado para el desarrollo de los modelos de pronóstico fuese completo y coherente.

### **3.5.2. División de conjuntos de datos**

La división de conjuntos de datos es un paso esencial en el desarrollo de modelos de pronóstico basados en algoritmos de machine learning. En la presente investigación, se llevó a cabo esta división para crear tres conjuntos de datos independientes, cada uno con un propósito específico:

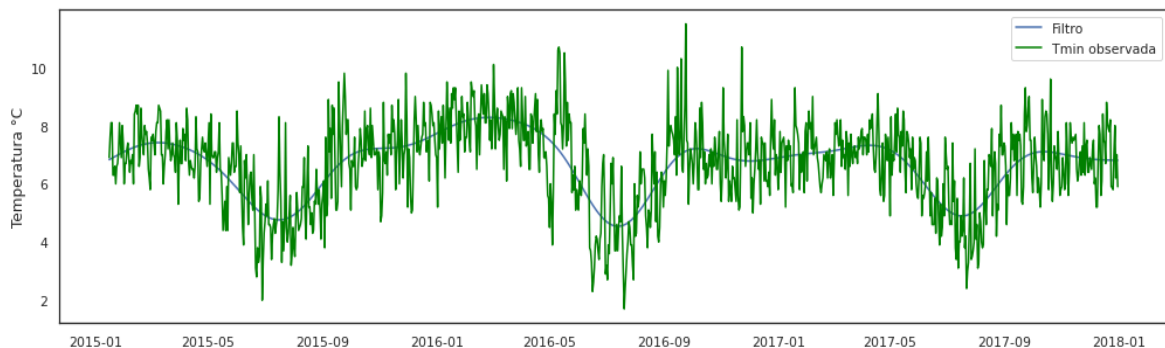
- **Conjunto de datos de entrenamiento:** representan el conjunto de observaciones a partir del cual se entrena el modelo; es decir, los datos con los cuales se construye la función de estimación que se empleará para realizar las predicciones, además se empleó en el proceso de selección de variables que alimentarán al modelo. Una notación formal para representar el conjunto de datos de entrenamiento, la cual fue dada por (James et al., 2013) indica que  $x_{ij}$  representa la  $i$ -ésima observación de la  $j$ -ésima variable predictora, donde  $i = 1, 2, \dots, n$  y  $j = 1, 2, \dots, p$ . Por otra parte  $y_i$  representa la  $i$ -ésima observación de la variable respuesta o predictando, con lo cual la data de entrenamiento se encuentra representada por  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$
- **Conjunto de datos de testeo:** Este conjunto está compuesto por observaciones independientes de los datos de entrenamiento y se utiliza para evaluar el rendimiento del modelo. La evaluación con datos de validación ayuda a identificar problemas de sobreajuste y permite ajustar los hiperparámetros del modelo para mejorar su capacidad de predicción.
- **Conjunto de datos de validación:** Una vez que el modelo ha sido entrenado y ajustado, se utiliza el conjunto de datos de testeo para evaluar su rendimiento final. En esta fase, se examina cómo el modelo se ajusta a nuevas observaciones. Formalmente, un modelo entrenado con el conjunto de datos  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  y representado por la función de estimación  $\hat{f}$ , se evalúa comparando  $\hat{f}(x_0)$  con  $y_0$ , donde  $(x_0, y_0)$ , es un elemento del conjunto de datos de testeo (James et al., 2013).

En el contexto del pronóstico de temperaturas extremas diarias, se utilizaron dos fuentes de información: variables del modelo numérico como predictores y datos de estación como predictandos. Dado que estas fuentes tenían periodos de disponibilidad de datos diferentes, se interceptaron para obtener un intervalo total de datos desde enero de 2015 (inicio de las corridas del modelo GFS 0.25°) hasta diciembre de 2018 (último mes con datos disponibles de las estaciones meteorológicas). Este intervalo de tiempo representó el período total de estudio en el que se desarrollaron los modelos empíricos de pronóstico para las temperaturas extremas diarias.

### 3.5.3. Extracción de variabilidad de alta frecuencia

El siguiente paso consistió en la extracción de la variabilidad de alta frecuencia en predictores y predictandos. En el caso de los predictandos, se separaron las series temporales de temperatura observada para cada estación meteorológica en su variabilidad de alta y baja frecuencia. Esto se logró mediante la aplicación de un filtro de Butterworth de tercer orden con un período de corte (cutoff) de 60 días. El filtro se utilizó para eliminar la estacionalidad y la variabilidad de baja frecuencia, dejando únicamente la variabilidad de alta frecuencia en las series de tiempo.

Por otro lado, en cuanto a los predictores, que consistían en información grillada del modelo numérico, se aplicó el mismo filtro de Butterworth con las mismas características de orden y período de corte a cada punto de grilla dentro del dominio para todos los predictores disponibles. Para ello, se generó una serie temporal en cada uno de los puntos de grilla y se aplicó el filtro, con el objetivo de preservar únicamente la variabilidad de alta frecuencia en los datos de los predictores. La figura 9 muestra un ejemplo de la aplicación del filtro sobre la serie temporal de temperatura mínima de la estación meteorológica Purhuay.



**Figura 9:** Serie temporal de temperatura mínima de la estación Purhuay con la aplicación de filtro de Butterworth.

### 3.5.4. Componentes principales de predictores y predictandos

Luego de aplicar el filtro a los predictores y predictandos, se empleó la técnica de reducción de dimensionalidad conocida como análisis de componentes principales o PCA (Principal Component Analysis) sobre ambos conjuntos de datos. Antes de calcular las componentes principales, se realizó un preprocesamiento que consistió en la estandarización de los datos. La estandarización es crucial en PCA, ya que esta técnica implica una transformación lineal

de las variables originales y puede ser sensible a las unidades o escalas de medición de las variables (Jolliffe y Cadima, 2016). Por lo tanto, se estandarizaron los conjuntos de datos para asegurarse de que todas las variables tuvieran la misma escala antes de aplicar PCA.

La estandarización de los conjuntos de datos se realizó utilizando la siguiente ecuación:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Ec.5

Donde:

$z_{ij}$  = valor estandarizado correspondiente a la i-ésima observación de la j-ésima columna

$x_{ij}$  = valor original correspondiente a la i-ésima observación de la j-ésima columna

$\bar{x}_j$  = promedio de la j-ésima columna

$s_j$  = desviación estándar correspondiente a la j-ésima columna

Una vez estandarizados los datos, se procedió a calcular las componentes principales. Para los predictandos, que consistían en las series temporales de variabilidad de alta frecuencia de temperatura observada para cada estación meteorológica, se creó una matriz de datos donde cada fila representaba una observación de temperatura filtrada de un día, y cada columna representaba una estación meteorológica. Esta matriz fue estandarizada y luego sometida al análisis de componentes principales para obtener las componentes principales de los predictandos.

En el caso de los predictores, que tenían tres dimensiones (latitud, longitud y tiempo), se realizó reorganizó los datos en una matriz de dos dimensiones. Cada columna de la nueva matriz representaba un punto de grilla, es decir, una ubicación dentro del dominio geográfico. Luego de este arreglo, se estandarizaron los datos y se aplicó el análisis de componentes principales para obtener las componentes principales de los predictores.

Es importante mencionar que, aunque la aplicación de PCA tiene como objetivo reducir la dimensionalidad manteniendo la mayor varianza explicada, la cantidad de componentes principales obtenidos es igual al número de variables originales. Por lo tanto, se emplearon criterios para limitar el número de componentes principales a utilizar.

Qona'ah et al. (2020) indican 3 criterios para encontrar el número de PC a retener:

- Un método visual mediante *scree plot*, detallado por Johnson y Wichern (2007), donde se grafican los eigenvalues ordenados de mayor a menor para cada una de las PC. Determinar el número apropiado de componentes se realiza de manera visual, buscando una curvatura en el scree plot a partir de la cual los siguientes eigenvalues son relativamente pequeños y sin cambios significativos uno de otro.
- Se selecciona las PC cuyos eigenvalues sean mayores a 1
- Se seleccionan las PC hasta alcanzar un porcentaje acumulado de varianza explicada del 80%; es decir, la suma del porcentaje de varianza explicada de cada una las PC seleccionadas deben superar el umbral establecido.

En este estudio, se utilizó el criterio de retener las componentes principales hasta alcanzar un 80% de varianza acumulada explicada, un enfoque comúnmente utilizado en análisis de componentes principales.

La estandarización de los datos y el cálculo de las componentes principales se llevaron a cabo utilizando las funcionalidades proporcionadas por la librería Scikit-Learn en el lenguaje de programación Python.

### **3.5.5. Selección de predictores y construcción de ecuaciones de regresión**

Después de transformar los predictores y los predictandos en sus componentes principales mediante el filtro y el análisis de componentes principales (PCA), el siguiente paso consistió en la construcción de ecuaciones de regresión lineal múltiple. Estas ecuaciones se diseñaron utilizando las componentes principales de los diferentes predictores derivados del modelo numérico como variables independientes, mientras que las componentes principales de los datos observados se utilizaron como variable dependiente. Este proceso requirió el desarrollo de un conjunto de ecuaciones de regresión igual al número de componentes principales de los predictandos que se retuvieron.

Dado que había una amplia variedad de posibles predictores, compuestos por las componentes principales de diversas variables del modelo numérico, fue necesario llevar a cabo un proceso de selección de predictores. Según Deng (1998), en problemas de aprendizaje automático que buscan encontrar una función para estimar la variable de respuesta  $Y$  a partir de variables predictoras  $X = \{x_1, x_2, \dots, x_M\}$ , se opta por seleccionar un subconjunto de predictores  $m$ , donde  $m < M$ . Esta selección se debe a que utilizar todos los

predictores podría conducir al sobreajuste, en el que la función se ajusta excepcionalmente bien a los datos de entrenamiento, pero no puede generalizar correctamente con datos fuera de ese conjunto. Además, esto también tendría un costo computacional considerable.

El método aplicado para seleccionar los predictores fue el "forward feature selection" (selección de características hacia adelante). Según Wilks (2011) este enfoque comienza con una ecuación de predicción no informativa  $y = b_0$ , que representa el promedio de las observaciones de los predictandos. Luego, se evalúan los  $M$  posibles predictores en función de su relación con el predictando, generando así  $M$  ecuaciones de regresión lineal simple. Se selecciona el predictor que mejor se ajuste, lo que resulta en una nueva ecuación de la forma  $y = b_0 + x_1 b_1$ , donde  $x_1$  es el primer predictor seleccionado y  $b_0$  toma un valor diferente al de la ecuación original.

A continuación, se prueban ecuaciones de regresión lineal múltiple utilizando los  $M - 1$  predictores restantes. Estas ecuaciones incluyen necesariamente el predictor  $x_1$  seleccionado anteriormente. Por lo tanto, la ecuación tomará la forma  $y = b_0 + x_1 b_1 + x_2 b_2$ , donde en este paso se selecciona el predictor  $x_2$  debido a que su inclusión mejora el coeficiente de determinación  $R^2$ , reduce el Error Cuadrático Medio (MSE) u otra métrica estadística que se utilice para medir la mejora del modelo. Este proceso se repite hasta que no haya más mejoras en el rendimiento del modelo con la adición de nuevos predictores o se cumpla un criterio de detención, como un número máximo de predictores. En este estudio, se estableció un límite de 10 predictores como criterio de detención.

Se aplicó la misma metodología de construcción de ecuaciones de regresión para cada una de las componentes principales de los predictandos. Esto condujo a la obtención de múltiples ecuaciones de regresión, que posteriormente se emplearon para pronosticar las componentes principales y, finalmente, obtener el pronóstico completo.

### **3.5.6. Validación de modelo de pronóstico temperaturas extremas**

Una vez se obtuvieron las ecuaciones de pronóstico para las temperaturas extremas diarias, se procedió a evaluar su desempeño mediante la utilización del conjunto de datos de prueba correspondiente al año 2018. En este proceso, se emplearon dos métricas estadísticas clave: la Raíz Cuadrada del Error Cuadrático Medio (RMSE) y el coeficiente de correlación de

Pearson ( $r$ ). Estas métricas permiten cuantificar la precisión y la relación entre los valores pronosticados y los valores observados.

Para llevar a cabo esta evaluación, se utilizó cada modelo de pronóstico desarrollado para generar predicciones, las cuales se compararon con los valores reales observados en el conjunto de datos de prueba. Además, se calculó el pronóstico directo proporcionado por el modelo numérico GFS sin aplicar ningún posprocesamiento. Para esto, se determinó el punto de grilla del GFS más cercano a cada estación meteorológica utilizando distancias euclidianas. A partir de este punto, se extrajeron los valores máximos y mínimos de temperatura diaria previstos por el modelo GFS.

Con el fin de establecer una medida de referencia adicional para evaluar el rendimiento de los modelos de pronóstico desarrollados, se aplicó el método de pronóstico por persistencia. Este método se basa en la premisa de que las condiciones meteorológicas mantendrán las mismas características durante un intervalo de horas en el futuro, tal como lo hicieron en el mismo intervalo de tiempo en el pasado. Se asume que no habrá cambios en las condiciones atmosféricas en ese período (Valderá-Figueredo et al., 2018). Es importante señalar que este método es aplicable en el contexto de pronósticos meteorológicos a corto plazo.

En resumen, se emplearon diversas métricas estadísticas para medir el desempeño y los errores de los modelos de pronóstico. Estos métodos permitieron evaluar cuán bien se ajustan las predicciones de los modelos a los valores observados y cómo se comparan con el pronóstico directo del modelo numérico GFS y el método de pronóstico por persistencia.

### **Raíz Cuadrada del Error Cuadrático Medio**

La Raíz Cuadrada del Error Cuadrático Medio (RMSE) es una métrica estadística ampliamente utilizada para evaluar el rendimiento de modelos de pronóstico cuantitativo. Esta métrica proporciona una medida del error entre los valores pronosticados por el modelo y los valores observados. Conforme a la literatura, el uso del RMSE es apropiado ya que refleja de manera efectiva los errores del modelo, dado que estos errores generalmente siguen una distribución normal (Chai y Draxler, 2014). No obstante, es importante destacar que el RMSE es altamente sensible a la presencia de valores atípicos (outliers), los cuales pueden afectar significativamente su cálculo. En situaciones prácticas, es posible considerar la eliminación de valores atípicos al calcular el RMSE, especialmente cuando la cantidad de muestras es limitada.



La expresión matemática del RMSE se define considerando un conjunto de n muestras:

$$e_i = Y_i - F_i$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2}$$

Ec.6

Donde:

$Y_t$ : i-ésimo valor observado

$F_t$ : i-ésimo valor estimado por el modelo de predicción

$e_t$ : i-ésimo error de predicción

En resumen, el RMSE proporciona una medida cuantitativa de cuán bien el modelo de pronóstico se ajusta a los valores reales. Es una métrica útil para evaluar la precisión general del modelo, pero se debe considerar su sensibilidad a valores atípicos al interpretar los resultados.

### **Coefficiente de correlación de Pearson**

El coeficiente de correlación de Pearson ( $r$ ) es una medida estadística que cuantifica la relación lineal entre dos variables cuantitativas. El valor de  $r$  oscila entre -1 y 1, y su signo indica si las variables exhiben una correlación positiva o negativa. Aunque no hay umbrales fijos para categorizar el grado de correlación, se ofrecen algunos valores de referencia en la Tabla 3.

La fórmula para calcular  $r$  se basa en tres sumatorias: la suma de los cuadrados de las diferencias entre los valores observados y sus promedios ( $SS_{XX}$  y  $SS_{YY}$ , respectivamente) y la suma de las diferencias cruzadas entre los valores observados y estimados por el modelo ( $SS_{XY}$ ). La ecuación se presenta a continuación:

$$SS_{XX} = \sum (x_i - \bar{x})^2$$

$$SS_{YY} = \sum (y_i - \bar{y})^2$$

$$SS_{XY} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

$$r = \frac{SS_{XY}}{\sqrt{(SS_{XX})(SS_{YY})}}$$

Ec.7

Donde:

$x_i$ : i-ésimo valor observado

$y_i$ : i-ésimo valor estimado por el modelo de predicción

$\bar{x}$ : valor promedio de valores observados

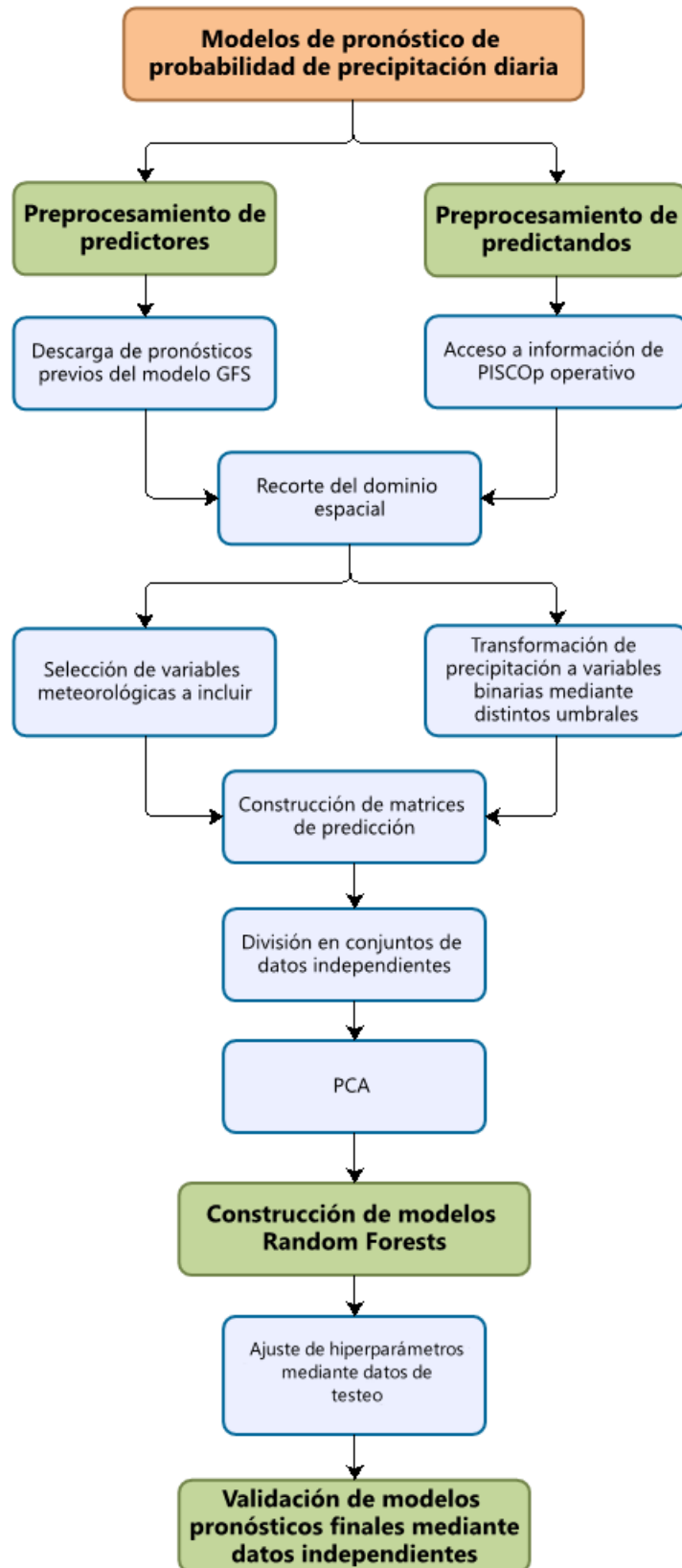
$\bar{y}$ : valor promedio de valores estimados por el modelo de predicción

**Tabla 3:** Grados de Correlación basada en el valor de Coeficiente de Pearson (Mukaka, 2012)

Valores de $r$	Grado de correlación entre variables
$0.9 < r \leq 1$ ( $-0.9 > r \geq -1$ )	Correlación positiva (negativa) muy elevada
$0.7 < r \leq 0.9$ ( $-0.7 > r \geq -0.9$ )	Correlación positiva (negativa) elevada
$0.5 < r \leq 0.7$ ( $-0.5 > r \geq -0.7$ )	Correlación positiva (negativa) moderada
$0.3 < r \leq 0.5$ ( $-0.3 > r \geq -0.5$ )	Correlación positiva (negativa) baja
$-0.3 \leq r \leq 0.3$	Correlación poco significativa

### 3.6. Metodología para pronóstico de probabilidad de precipitación

La metodología aplicada al pronóstico de probabilidad de precipitación a nivel diario se basó en la utilización de modelos de pronóstico desarrollados mediante el algoritmo Random Forests, diseñados para tareas de clasificación. Estos modelos fueron creados considerando diferentes umbrales de precipitación acumulada diaria. La Figura 10 ilustra de manera esquemática los procedimientos llevados a cabo para la obtención de los modelos de probabilidad de precipitación.



**Figura 10:** Esquema de la metodología de investigación para pronóstico de probabilidad de precipitación

### 3.6.1. División de conjuntos de datos

En la sección anterior, se destacó la relevancia de dividir los conjuntos de datos disponibles en función de su papel en el desarrollo de modelos de predicción. Con el objetivo de construir modelos de pronóstico de probabilidad de precipitación diaria, se recopilaron dos fuentes de información esenciales: las variables meteorológicas del modelo numérico GFS, que actuaron como predictores, y los datos grillados de precipitación diaria del PISCOp, que actuaron como predictandos. El periodo de tiempo con información disponible abarcó desde enero de 2015 hasta diciembre de 2021, estableciendo el marco temporal total para el estudio de pronóstico de probabilidad de precipitación. Este periodo se dividió en los siguientes conjuntos de datos:

- **Conjunto de datos de entrenamiento:** Se extendió desde enero de 2015 hasta diciembre de 2019.
- **Conjunto de datos de testeo:** Cubrió el periodo de enero a diciembre del año 2020
- **Conjunto de datos de validación:** Abarcó el periodo de enero a diciembre del año 2021

### 3.6.2. Selección de predictores

Dado el gran número de posibles predictores derivados del modelo global, fue necesario realizar una selección de aquellos que tendrían un mayor impacto en la predicción de la ocurrencia de precipitaciones en la zona de estudio. Para esta tarea, se utilizó el coeficiente de correlación de Pearson con el propósito de identificar las variables con mayor influencia en este fenómeno meteorológico. Se llevó a cabo un proceso de evaluación de las variables predictoras seleccionadas para garantizar que estuvieran físicamente coherentes con la ocurrencia de precipitaciones. Esto se debió a que, aunque algunas variables podrían presentar una alta correlación con la variable objetivo, podrían carecer de una relación directa con las precipitaciones, lo que las haría inadecuadas como predictores.

### 3.6.3. Construcción de matriz de predictores

Después de identificar los predictores relevantes a través de su correlación con la variable objetivo, se procedió a definir la estrategia para construir la matriz de datos esencial para entrenar el modelo. Para esta etapa, se adoptó una metodología similar a la presentada por

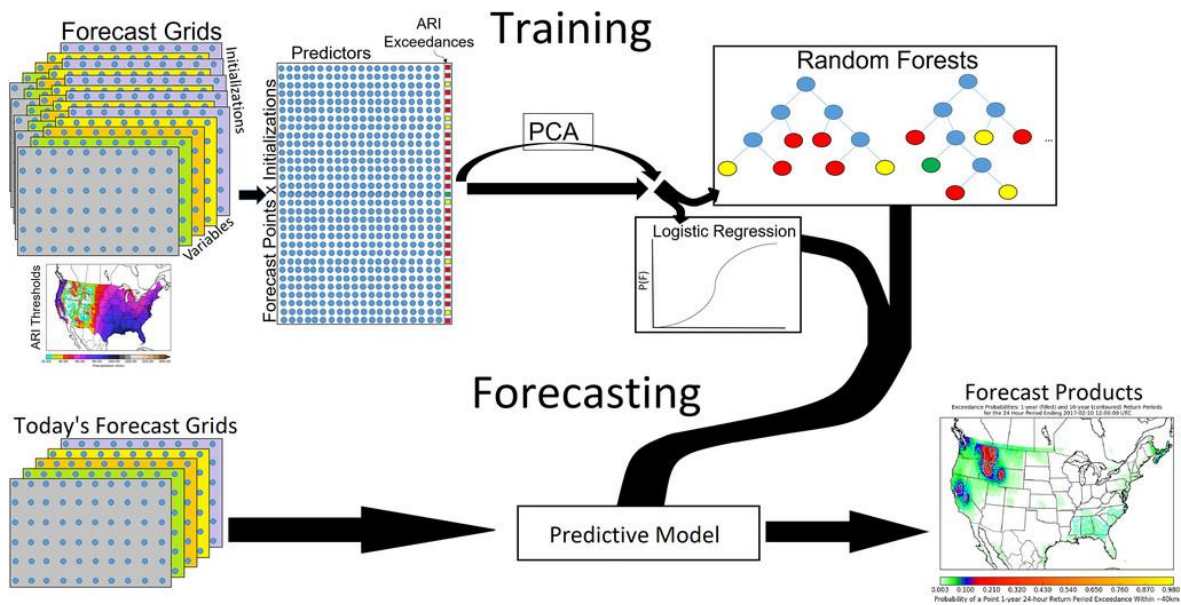
Herman y Schumacher (2018) en el contexto del pronóstico de precipitaciones extremas, como se ilustra en el esquema de la Figura 11.

Los predictores meteorológicos seleccionados fueron reorganizados de manera que cada punto de grilla dentro del dominio geográfico representara un punto de pronóstico. Desde este enfoque, se extrajo información correspondiente a las  $N$  variables meteorológicas previamente identificadas como relevantes. Además, se consideraron valores de pronóstico del modelo GFS para un radio  $R = 2$ , lo que equivale a seleccionar un cuadrante con los  $(2R + 1)^2$  puntos más cercanos al punto de pronóstico. Dado que la variable objetivo consiste en el acumulado diario de precipitación, se incluyó información de pronóstico para distintas horas del día. Esto se realizó con una resolución temporal de 6 horas, generando  $T = 5$  intervalos de pronóstico por día.

Los horizontes de pronóstico tomados en consideración para el primer día fueron +12h, +18h, +24h, +30h y +36h; para el segundo día, se consideraron +36h, +42h, +48h, +54h y +60h; finalmente, para el tercer día se tuvieron en cuenta +60h, +66h, +72h, +78h y +84h. En total, se construyó la matriz de predicción con una cantidad significativa de columnas de predictores, representando la información espacial y temporal de manera detallada.

Sin embargo, debido a la elevada cantidad de columnas en la matriz de predicción, la posibilidad de sobreajuste y los altos requerimientos computacionales, se optó por aplicar una técnica clásica de reducción de dimensionalidad: el Análisis de Componentes Principales (PCA). Dado que las columnas de predictores presentaban fuertes relaciones espaciales y temporales, PCA se convirtió en una herramienta valiosa para reducir la cantidad de dimensiones y conservar la información esencial.

En términos adicionales, la cantidad de filas en la matriz de predicción se definió como el producto entre la cantidad de puntos de pronóstico y los días de pronóstico considerados en el estudio. Esta estrategia permitió construir una matriz de datos robusta y representativa para el entrenamiento del modelo, fusionando información espacial y temporal de manera coherente y efectiva.



**Figura 11:** Esquema metodológico para el desarrollo de un modelo de predicción de probabilidad de precipitaciones extremas empleando Random Forests

**FUENTE:** Herman y Schumacher, 2018

### 3.6.4. Transformación de predictandos

Dado el enfoque de clasificación que se adoptó para desarrollar los modelos de pronóstico de probabilidad de precipitación en esta investigación, fue necesario transformar los datos de PISCOp utilizados como predictandos. Estos datos representan montos acumulados diarios de precipitación para cada punto de grilla y son continuos en naturaleza. Sin embargo, para abordar el problema de clasificación, se transformaron en variables binarias opuestas, típicamente representadas por los valores 0 y 1.

Esta transformación se llevó a cabo asignando el valor 1 a los puntos de pronóstico en los cuales el monto acumulado de precipitación superara un umbral predefinido. Por otro lado, se asignó el valor 0 a los puntos de grilla en los cuales el monto acumulado no superara dicho umbral. En resumen, los modelos de pronóstico se entrenaron considerando variables binarias como predictandos, lo que permitió que los modelos clasificaran entre las categorías de 0 y 1. Además, los modelos proporcionaron la probabilidad correspondiente a cada punto de pronóstico de exceder el umbral de precipitación específico utilizado para entrenar el modelo respectivo.

Se entrenaron cuatro modelos de precipitación independientes, cada uno empleando un umbral específico. El primer umbral utilizado fue de 0.1 mm/día, designado exclusivamente para indicar la presencia o ausencia de precipitación en los pronósticos diarios. Este valor corresponde al percentil 20 de los datos PISCOp en la zona de estudio.

El segundo umbral corresponde a 1 mm/día, utilizado como indicador de precipitación ligera y representando el percentil 50 de la precipitación en el dominio espacial considerado.

Los dos últimos umbrales, establecidos arbitrariamente en 2 y 5 mm/día, se seleccionaron para indicar precipitación de moderada intensidad y se fijaron en valores enteros con el propósito de facilitar la interpretación en los productos finales generados. Estos valores corresponden a los percentiles 63 y 83, respectivamente.

Esta selección de umbrales permite abordar diversas categorías de eventos de precipitación, desde la ausencia hasta intensidades moderadas, proporcionando así información detallada sobre la magnitud de las condiciones meteorológicas previstas.

### **3.6.5. Calibración de modelos de predicción**

La calibración de los modelos se efectuó mediante el algoritmo Random Forests Classifier. Para configurar inicialmente los modelos de pronóstico, se emplearon los conjuntos de datos de entrenamiento, y se utilizó el conjunto de datos de testeo como punto de referencia para encontrar la combinación más apropiada de hiperparámetros para cada uno de los modelos de pronóstico correspondientes a distintos umbrales de precipitación.

Durante el entrenamiento del modelo con el umbral más alto de 10 mm/día, se identificó la necesidad de introducir un hiperparámetro adicional. Esto se debió a una limitación en la que la cantidad de puntos de pronóstico con valores inferiores a dicho umbral era significativamente mayor que la de sus contrapartes superiores. En otras palabras, se observó un desbalance en los valores del predictando. Para abordar esta situación, fue necesario asignar un mayor peso e importancia a los datos menos frecuentes durante la fase de entrenamiento del modelo Random Forests. Esto permitió al algoritmo identificar y considerar de manera efectiva los casos de precipitación menos comunes.

El ajuste de los hiperparámetros del modelo se llevó a cabo mediante la comparación de los resultados obtenidos en los datos de validación, evaluando distintas configuraciones. Los hiperparámetros específicos utilizados para entrenar los diferentes modelos de pronóstico, se empleó la medida de impureza Gini, una profundidad máxima de árbol de 5 y 100

estimadores (cantidad de árboles de decisión) para cada modelo. Se aplicó pesos a las clases únicamente en los modelos de 2 y 5 mm/día para compensar el desbalanceo de datos.

### 3.6.6. Validación de modelos

Después de entrenar y ajustar los modelos Random Forests para el pronóstico de probabilidad de precipitación, se procedió a evaluar su rendimiento utilizando el conjunto de datos de testeo. Para medir la eficacia de los pronósticos, se emplearon diversas métricas estadísticas: Precisión, Exactitud, Exhaustividad, F1-score, AUC, así como matrices de confusión para cada uno de los modelos entrenados con diferentes umbrales de precipitación.

A continuación, se describen las métricas y gráficos utilizados para medir el rendimiento y la calidad de los modelos de pronóstico.

**Matriz de confusión:** La matriz de confusión es una herramienta que contiene información sobre los valores reales y los valores predichos por un modelo de clasificación (Kohavi y Provost, 1998). Se utiliza ampliamente en análisis predictivos y para evaluar el rendimiento de modelos basados en machine learning. La estructura de una matriz de confusión aplicada a un modelo de clasificación binaria se muestra en la Figura 12.

		Valores Predichos	
		Positivo	Negativo
Valores Reales	Positivo	TP	FN
	Negativo	FP	TN

**Figura 12:** Modelo de matriz de confusión en clasificación binaria

Donde:

TP: representa la cantidad de verdaderos positivos

TN: representa la cantidad de los verdaderos negativos

FN: representa la cantidad de los falsos negativos

FP: representa la cantidad de los falsos positivos



**Exactitud:** La exactitud (*accuracy*) muestra el porcentaje de predicciones correctas realizadas por el modelo (Sokolova et al., 2006). Aunque es un indicador de la eficacia general del modelo, no es confiable en conjuntos de datos desbalanceados. Se calcula con la ecuación 8.

$$accuracy = \frac{tp + tn}{tp + fp + fn + tn}$$

Ec.8

**Precisión:** Este indicador se calcula dividiendo el número de predicciones correctas de una clase determinada (generalmente la clase positiva) por la cantidad total de instancias de esa clase. Evalúa la capacidad predictiva del algoritmo y se calcula mediante la ecuación 9.

$$precision = \frac{tp}{tp + fp}$$

Ec.9

**Exhaustividad:** También conocida como sensibilidad o *recall*, la exhaustividad muestra la proporción de observaciones positivas correctamente identificadas por el modelo en relación con todas las observaciones positivas. Calcula la efectividad del algoritmo para una sola clase y se obtiene mediante la ecuación 10.

$$recall = \frac{tp}{tp + fn}$$

Ec.10

**F-score:** Este estadístico es el promedio ponderado de precisión y exhaustividad, dependiendo de la función de peso  $\beta$ . El F1-score es equivalente a la media armónica de los indicadores mencionados anteriormente (Dalianis, 2018). En esta investigación, se utilizó el F1-score como medida combinada de la exhaustividad y precisión de los modelos de pronóstico desarrollados.

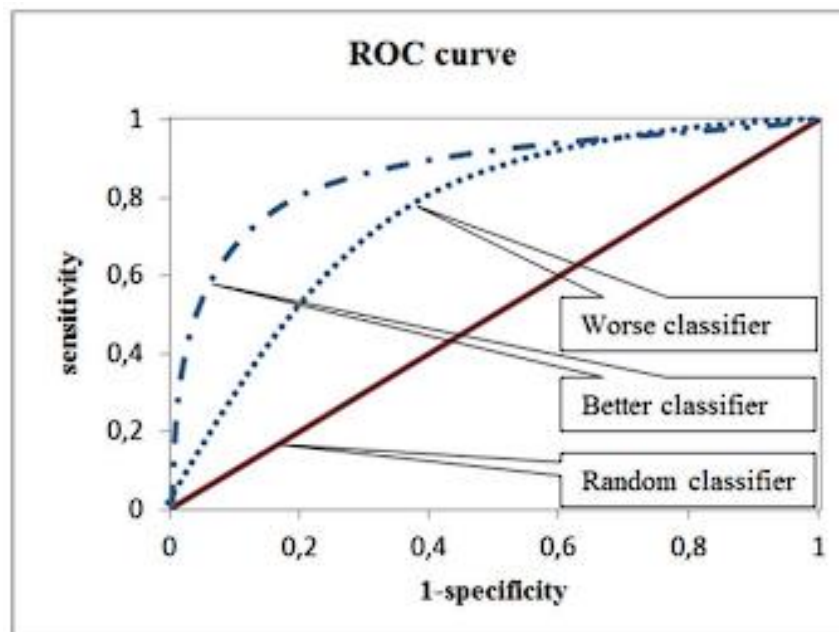
$$F_{\beta} - score = \frac{(\beta^2 + 1) \times precision \times recall}{\beta^2 \times precision + recall}$$

$$F_1 - score = \frac{2 \times precision \times recall}{precision + recall}$$

Ec.11

**Análisis de curva ROC y AUC:** La curva ROC (Receiver Operating Characteristic) es una herramienta estadística para evaluar la capacidad discriminativa de modelos de clasificación binaria. De acuerdo a Martínez Pérez y Pérez Martín (2023), se utiliza para determinar puntos de corte, evaluar la discriminación del modelo y comparar la capacidad de diferentes modelos. El Área bajo la curva (AUC) mide la capacidad del modelo para diferenciar entre clases positivas y negativas, representando la probabilidad de que el modelo clasifique correctamente.

Un modelo con una capacidad de clasificación perfecta tendría una curva ROC en la esquina superior izquierda del gráfico. Un modelo sin capacidad de clasificación se representaría por la diagonal (Figura 13). Una curva con un AUC de 0.75 sugiere una capacidad de discriminación moderada.



**Figura 13:** Interpretación de Curva ROC

**FUENTE:** Gajowniczek et al. (2014)

## **IV. RESULTADOS Y DISCUSIÓN**

En la presente sección se presentan los resultados y discusiones obtenidos de acuerdo a los objetivos generales y específicos planteados para el desarrollo de la presente investigación.

### **4.1. Modelo de pronóstico de temperatura mínima diaria**

En esta sección, se analizó la selección de predictores para el desarrollo del modelo de pronóstico de temperatura mínima diaria. Se generaron mapas de correlación espacial de Pearson para evaluar la relación entre la variabilidad de alta frecuencia de la temperatura mínima y las variables meteorológicas filtradas del modelo GFS correspondientes a las 12 UTC en el dominio de la región Áncash (Figura 14). El objetivo fue identificar las variables del modelo que presentaran una correlación significativa con la variable respuesta para su inclusión como potenciales predictores en los modelos empíricos de pronóstico.

Debido a la elevación de la zona de estudio, la cual se encuentra por encima de los 3,000 m s.n.m., se tomaron en consideración variables de modelo numérico partiendo desde los 500 hPa a niveles de presión superiores, además de variables a 2 m del suelo de acuerdo a la topografía del modelo.

El análisis de correlación de los mapas, reveló que la humedad específica en los niveles de 500 hPa y 400 hPa mostró una buena correlación con la alta frecuencia de la temperatura mínima, siendo esta más intensa en el nivel de 500 hPa. Los patrones espaciales de estas variables fueron similares a los observados en la humedad relativa de los mismos niveles de presión, lo que puede ser atribuido a su comportamiento similar como indicadores de la humedad atmosférica.

En cuanto al viento meridional, su correlación con la variable respuesta fue limitada, excepto en el nivel de 500 hPa, donde se observó una correlación negativa notable tratándose de valores de alta frecuencia, en la zona de Sierra con valores en el rango de 0.25 a -0.5. Esto llevó a considerar al viento meridional en el nivel de 500 hPa como el único predictor relevante entre las variables de viento.

El viento zonal en el nivel de 500 hPa mostró la correlación negativa más intensa, con valores en el rango de -0.6 a -0.4. Esta correlación sugiere que los patrones de viento en niveles medios de la atmósfera pueden influir en las variaciones de temperatura mínima en la zona de estudio.

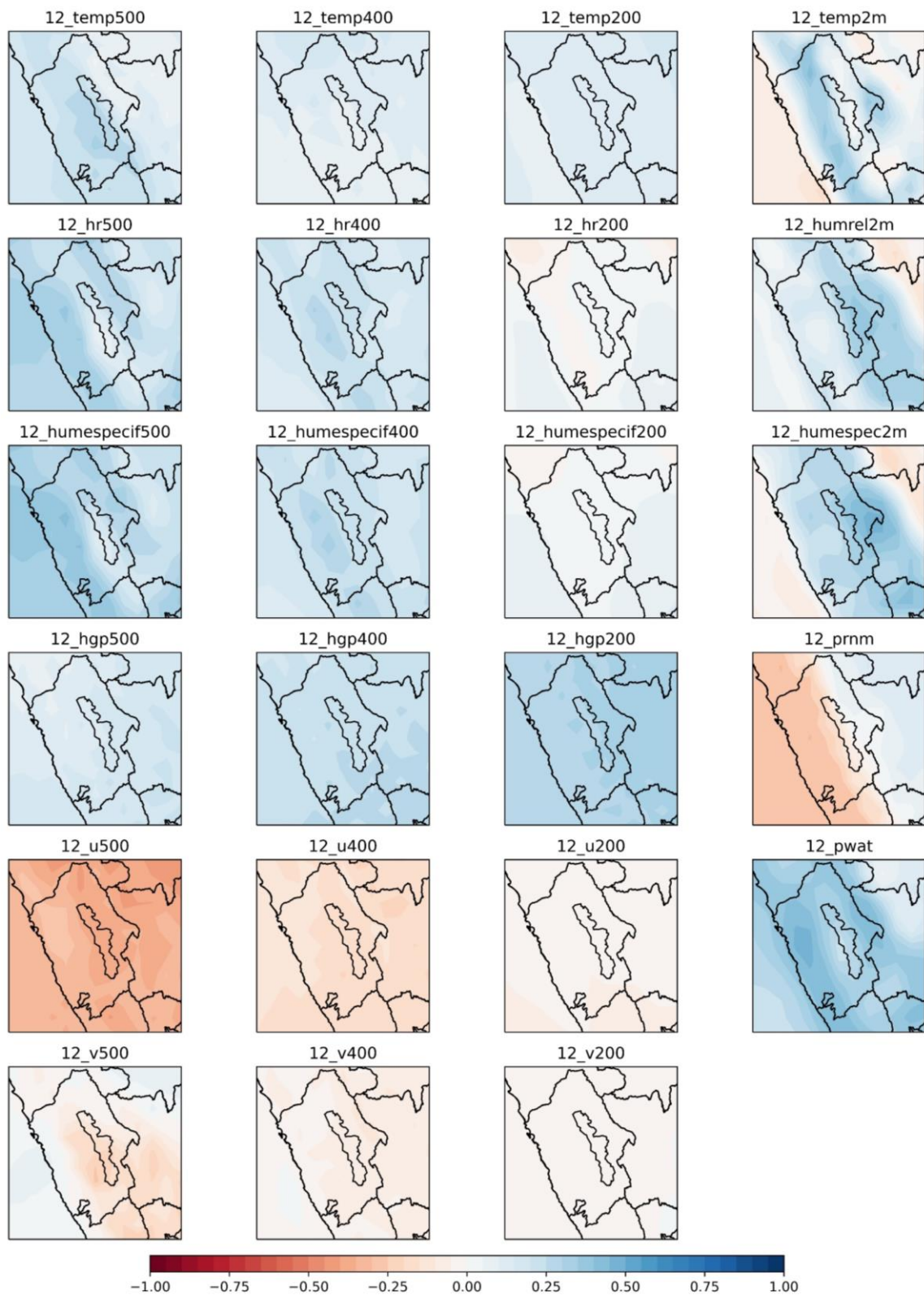
En el nivel de 200 hPa, la mayoría de las variables presentaron una correlación baja con la temperatura mínima, excepto la altura geopotencial que mostró una correlación positiva de hasta aproximadamente 0.45. Por esta razón, la altura geopotencial a 200 hPa se incluyó en el grupo de variables predictoras.

Las variables a 2 m del suelo presentaron correlaciones intensas con la alta frecuencia de la temperatura mínima, especialmente en la zona de Sierra central y áreas cercanas al área de estudio. Debido a esta alta correlación, las tres variables de modelo en superficie fueron consideradas como potenciales predictores para el desarrollo de los modelos de pronóstico.

Adicionalmente, estudios anteriores como el de Saavedra y Takahashi (2017) en los Andes centrales del Perú han destacado la correlación entre la humedad específica y la temperatura mínima, especialmente durante las temporadas secas. También se ha señalado la relevancia de la cobertura nubosa como un factor que afecta la temperatura mínima.

La variable Agua Precipitable (PWAT) mostró una correlación fuertemente positiva con la temperatura mínima, sugiriendo una relación entre el contenido de humedad en la atmósfera y las variaciones en la temperatura mínima en la región de estudio.

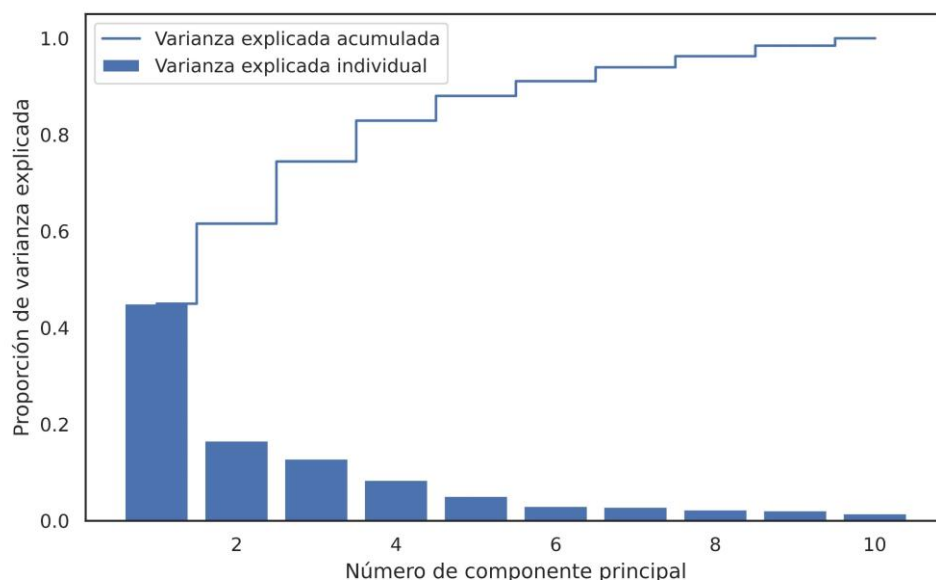
Finalmente, la presión reducida al nivel del mar presentó una influencia considerable con un patrón negativo en la zona occidental y ligeramente positivo en la zona oriental del área de estudio.



**Figura 14:** Mapas de correlación de Pearson de variabilidad de alta frecuencia de temperatura mínima y variables meteorológicas del modelo GFS correspondientes a las 12 UTC en la zona de estudio

La Figura 15 nos proporciona información sobre la varianza explicada por cada una de las componentes principales (PC) obtenidas al aplicar PCA a los datos de alta frecuencia del grupo de estaciones meteorológicas (predictando). Se observa que la primera PC representa aproximadamente el 45% de la varianza total explicada del conjunto de datos, la cual sería la componente más importante para predecir adecuadamente. La segunda y la tercera componente principal explican alrededor del 18% y el 15% de la varianza, respectivamente. Sin embargo, a partir de la sexta componente principal, la varianza explicada disminuye significativamente, siendo cada una de estas componentes menor al 5%. Estas componentes con un bajo porcentaje de varianza explicada se consideraron asociadas al ruido en los datos y, por lo tanto, se descartaron como predictores en los modelos de pronóstico.

Tras analizar la varianza explicada, se decidió conservar solo las cinco primeras PCs como predictores. Estas cinco componentes explican conjuntamente alrededor del 87% de la varianza total del conjunto original de datos. Esta elección se basó en que, en conjunto, estas componentes superaron el umbral del 80% de varianza explicada acumulada y en que el incremento porcentual de varianza acumulada a partir de la quinta componente principal fue bastante reducido. Cada una de estas componentes principales se utilizó como predictor para desarrollar las ecuaciones de regresión lineal múltiple necesarias para predecir las componentes principales a partir de las cuales se realizará la reconstrucción de la temperatura mínima a alta frecuencia.



**Figura 15:** Varianza explicada individual y acumulada de las componentes principales obtenidas de la variabilidad de alta frecuencia de temperatura mínima proveniente de estaciones meteorológicas

La Tabla 4 proporciona información sobre el número de componentes principales que se retuvieron como potenciales predictores después de realizar luego de aplicar PCA en cada una de las variables meteorológicas filtradas del modelo GFS en el dominio de estudio. Estas variables corresponden a las mediciones tomadas a las 12 UTC, ya que generalmente esta hora presenta cercanía con la ocurrencia de las temperaturas más bajas del día. El número de componentes representa el mínimo necesario para superar el umbral del 80% de varianza explicada acumulada.

**Tabla 4:** Número de componentes principales incluidas como potenciales predictores para cada variable meteorológica proveniente de modelo numérico para la predicción de temperatura mínima

Variable	Número de PCs
12_u500	1
12_hgp400	1
12_hgp200	1
12_v500	3
12_pwat	3
12_v400	1
12_prnm	2
12_humespecif500	4
12_u400	1
12_hr500	4
12_humespecif400	3
12_temp2m	8
12_humrel2m	8
12_temp200	1
12_hr400	4
12_humespec2m	6
12_temp500	2
12_temp400	1
12_hgp500	1

Inicialmente, cada una de las variables generó 169 componentes principales (derivados de los 169 puntos de grilla dentro del dominio) a partir del PCA. Sin embargo, se retuvo solo la información más relevante de cada variable en un número reducido de componentes principales. Por ejemplo, las variables meteorológicas medidas a 2 metros sobre el suelo requerían almacenar un número mayor de componentes para alcanzar el umbral de varianza

del 80%. Se necesitaron 6 componentes principales para la humedad específica y 8 para la temperatura y la humedad relativa.

Por otro lado, las variables del modelo GFS en niveles de presión por encima de 500 hPa mostraron un comportamiento opuesto, ya que requirieron de un número bastante menor de componentes principales para explicar al menos el 80% de la varianza de los conjuntos de datos originales. En algunos casos, solo se necesitaron 4 componentes principales. El anexo 1 proporciona gráficos de varianza explicada individual y acumulada para cada una de las variables meteorológicas trabajadas, lo que ayudó en la selección de las componentes principales para su inclusión como predictores en los modelos de pronóstico.

Las tablas 5, 6 y 7 resumen los resultados del proceso de validación de los modelos de pronóstico de temperaturas mínimas en tres horizontes de pronóstico diferentes. Se consideró la variabilidad de alta frecuencia tanto pronosticada como observada. El RMSE, utilizado como medida de error del pronóstico, arrojó los siguientes resultados:

- Para el primer día de pronóstico, los errores se situaron en el rango de 0.62 a 1.15 °C en todas las estaciones meteorológicas utilizadas.
- Para el segundo día de pronóstico, el rango de error aumentó ligeramente, oscilando entre 0.65 y 1.18 °C.
- En el tercer y último día de pronóstico, el rango de error fue de 0.66 a 1.24 °C.

La Figura 17 proporciona una representación visual de la evolución del RMSE para los tres días de pronóstico en cada una de las estaciones. En su mayoría, las estaciones mostraron un aumento en el error a medida que se extendía el horizonte temporal de pronóstico. No obstante, la estación Quebrada Antamina se destacó por un aumento extremadamente pequeño en el RMSE, pasando de 0.752 °C en el primer día a 0.755 °C en el tercer día. Esto sugiere que los modelos de pronóstico podrían tener un rendimiento similar en los tres horizontes de pronóstico solo en esta estación en particular.

En términos de RMSE, se clasificaron 4 estaciones meteorológicas con predicciones correctas (con errores en el rango de 0.5 a 0.8 °C), mientras que 6 estaciones se consideraron predicciones de calidad intermedia (con errores en el rango de 0.8 a 1.2 °C). Se observó que la altitud de las estaciones no pareció influir significativamente en el RMSE del modelo de pronóstico. Un ejemplo de esto es la comparación entre las estaciones Morizq y Glaciar, que están ubicadas en zonas cercanas, pero presentan una diferencia de aproximadamente 0.25 °C en sus valores de RMSE. Esto podría atribuirse a la complejidad de las condiciones



meteorológicas a alta elevación y a la orografía del terreno. De acuerdo a Ghielmi y Eccel (2006), quienes desarrollaron modelos para la predicción de heladas primaverales en una zona agrícola de montaña, los resultados de la aplicación de sus modelos presentan resultados variables entre las distintas estaciones, inclusive en aquellas estaciones que se encuentran relativamente cerca, debido a que la orografía y la morfología del terreno juegan un papel importante. Tomando en consideración que ambas estaciones se encuentran a 4,814 y 4,918 m s.n.m. respectivamente, es esperable que los resultados de pronóstico en dichas estaciones sean variables por lo complicado de las condiciones meteorológicas a esta elevación, a pesar de la relativa cercanía que presentaron en su ubicación.

Al comparar los resultados con una investigación similar realizada por Qona'ah et al. (2020), quienes utilizaron una metodología comparable para predecir temperaturas extremas diarias en Indonesia, se observa que los resultados de la presente investigación muestran un rango de error similar en alrededor de 4 estaciones, mientras que las otras 6 estaciones presentaron un mayor grado de error. Sin embargo, es importante tener en cuenta las diferencias en las condiciones meteorológicas y geográficas entre ambas investigaciones, ya que el área de estudio de la presente investigación se encuentra en una zona montañosa con una orografía compleja, a diferencia de las condiciones tropicales de archipiélago de dicha investigación.

**Tabla 5:** Resultados obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura mínima en el periodo de validación con un horizonte de pronóstico de 24 horas

Estación	MSE (°C)	RMSE (°C)	r
Shilla	0.842	0.918	0.571
Shancayan	1.307	1.143	0.477
Ocros	0.544	0.738	0.638
Chiquian	0.470	0.686	0.591
Purhuay	0.955	0.977	0.486
Quillcayhuanca	1.130	1.063	0.561
Chacas	1.334	1.155	0.648
Quebrada Antamina	0.566	0.752	0.435
Morizq	0.395	0.628	0.492
Glaciar	0.787	0.887	0.559

**Tabla 6:** Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura mínima en el periodo de validación con un horizonte de pronóstico de 48 horas

Estación	MSE (°C)	RMSE (°C)	r
Shilla	0.917	0.957	0.518
Shancayan	1.337	1.156	0.462
Ocros	0.577	0.759	0.610
Chiquian	0.491	0.701	0.571
Purhuay	0.953	0.976	0.487
Quillcayhuanca	1.211	1.100	0.522
Chacas	1.381	1.175	0.632
Quebrada Antamina	0.569	0.754	0.429
Morizq	0.423	0.650	0.460
Glaciar	0.837	0.915	0.525

**Tabla 7:** Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura mínima en el periodo de validación con un horizonte de pronóstico de 72 horas

Estación	MSE (°C)	RMSE (°C)	r
Shilla	0.966	0.983	0.478
Shancayan	1.395	1.181	0.423
Ocros	0.603	0.776	0.587
Chiquian	0.525	0.724	0.533
Purhuay	0.990	0.995	0.455
Quillcayhuanca	1.270	1.127	0.487
Chacas	1.530	1.237	0.579
Quebrada Antamina	0.570	0.755	0.427
Morizq	0.439	0.663	0.428
Glaciar	0.834	0.913	0.525

La Figura 16 presenta las series temporales obtenidas de la validación del modelo de pronóstico de alta frecuencia de temperatura mínima durante todo el año 2018. Las estaciones con menor grado de error con respecto a los valores observados (estaciones Morizq, Ocros, Chiquian y Quebrada Antamina) muestran series temporales con una gran similitud al comparar las predicciones con los valores observados.

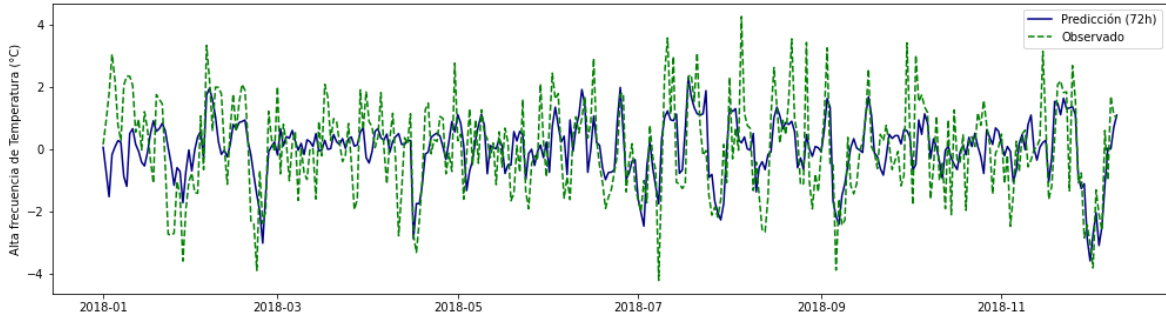
La estación Ocros, que arrojó la mejor correlación en el proceso de validación, presenta una gran similitud entre las series observada y pronosticada durante casi todo el periodo analizado. Sin embargo, se identificaron ciertas limitaciones en el modelo de pronóstico para predecir la magnitud de ciertos descensos intensos de temperatura. Aunque el modelo pudo identificar estos descensos de temperatura en la mayoría de los casos, en algunas estaciones no logró captar la magnitud de estos eventos fríos. Un ejemplo claro se observa a finales de febrero de 2018, donde hubo un evento de descenso de temperatura que afectó la mayoría de las estaciones. En la estación Chacas, por ejemplo, el modelo tuvo un error de pronóstico de hasta 2 °C, es decir, pronosticó una temperatura mínima 2 °C mayor de lo observado en este evento.

Otro caso identificable de descenso considerable de temperatura, que afectó a la mayoría de las estaciones, ocurrió a mediados del mes de diciembre. En este caso, a diferencia del evento anterior, el modelo fue capaz de identificar correctamente la magnitud del descenso de temperatura en prácticamente todas las estaciones afectadas.

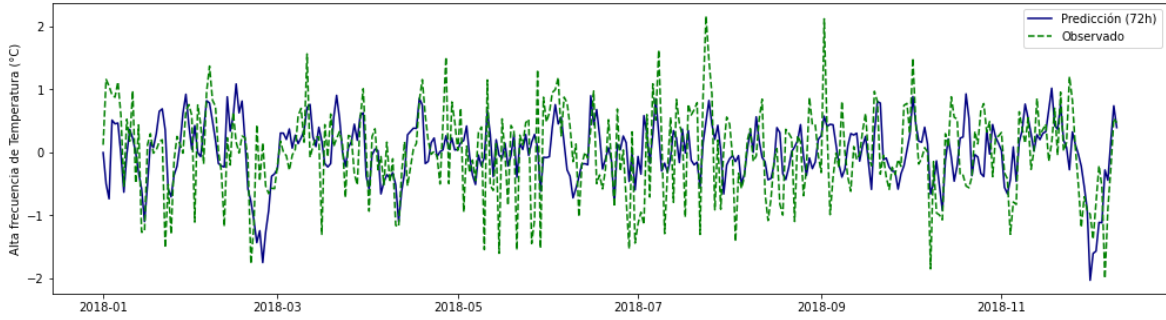
Es importante resaltar la importancia de analizar los eventos extremos, ya que, aunque el modelo para temperaturas mínimas mostró una buena capacidad predictiva durante casi todo el período de validación en todas las estaciones, son los eventos extremos, caracterizados por descensos más intensos de temperatura, los que tienen un mayor impacto y relevancia para los usuarios potenciales del modelo de pronóstico.

A pesar de que las estaciones Shancayan y Glaciar presentaron el mayor grado de error, el análisis de los valores pronosticados y observados en ambas estaciones indica que, aunque el modelo de pronóstico tiende a subestimar, lo cual concuerda con el mayor valor de RMSE en ambas estaciones, el modelo es capaz de identificar correctamente los incrementos y descensos de la variable pronosticada. Esto justifica la buena correlación lineal observada en ambas estaciones.

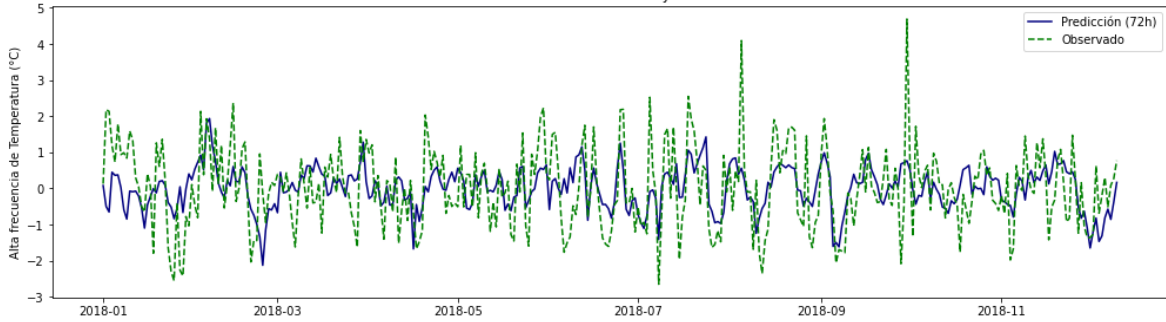
Pronóstico de variabilidad de alta frecuencia de Temperatura mínima  
Estación Chacas



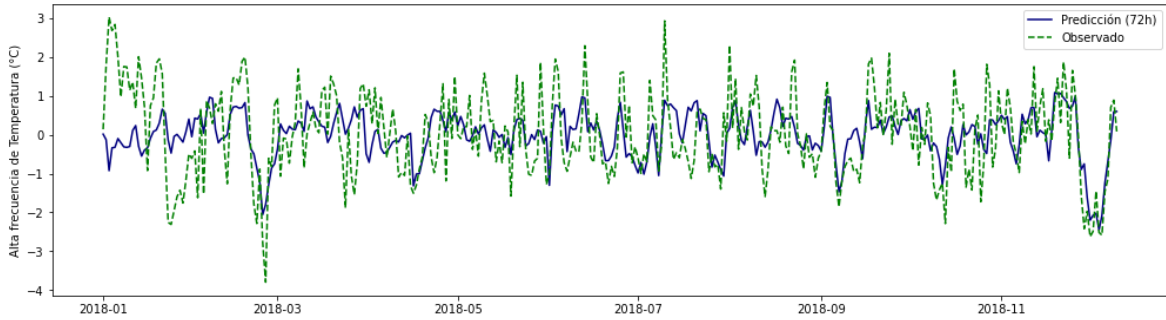
Estación Morizq



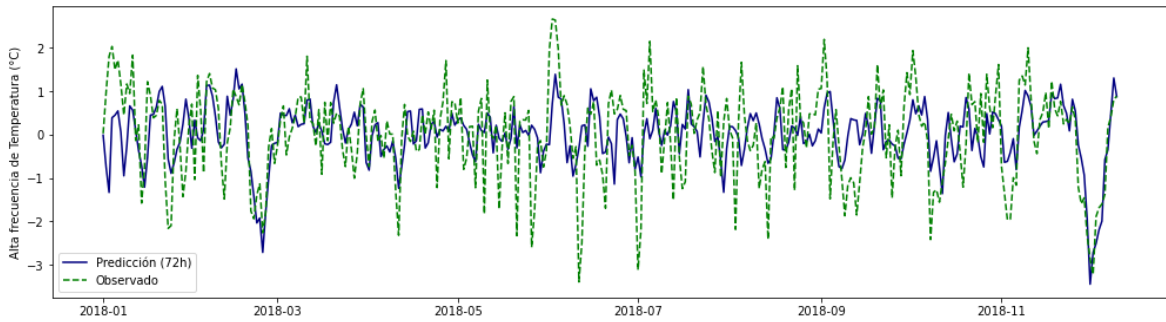
Estación Purhuay

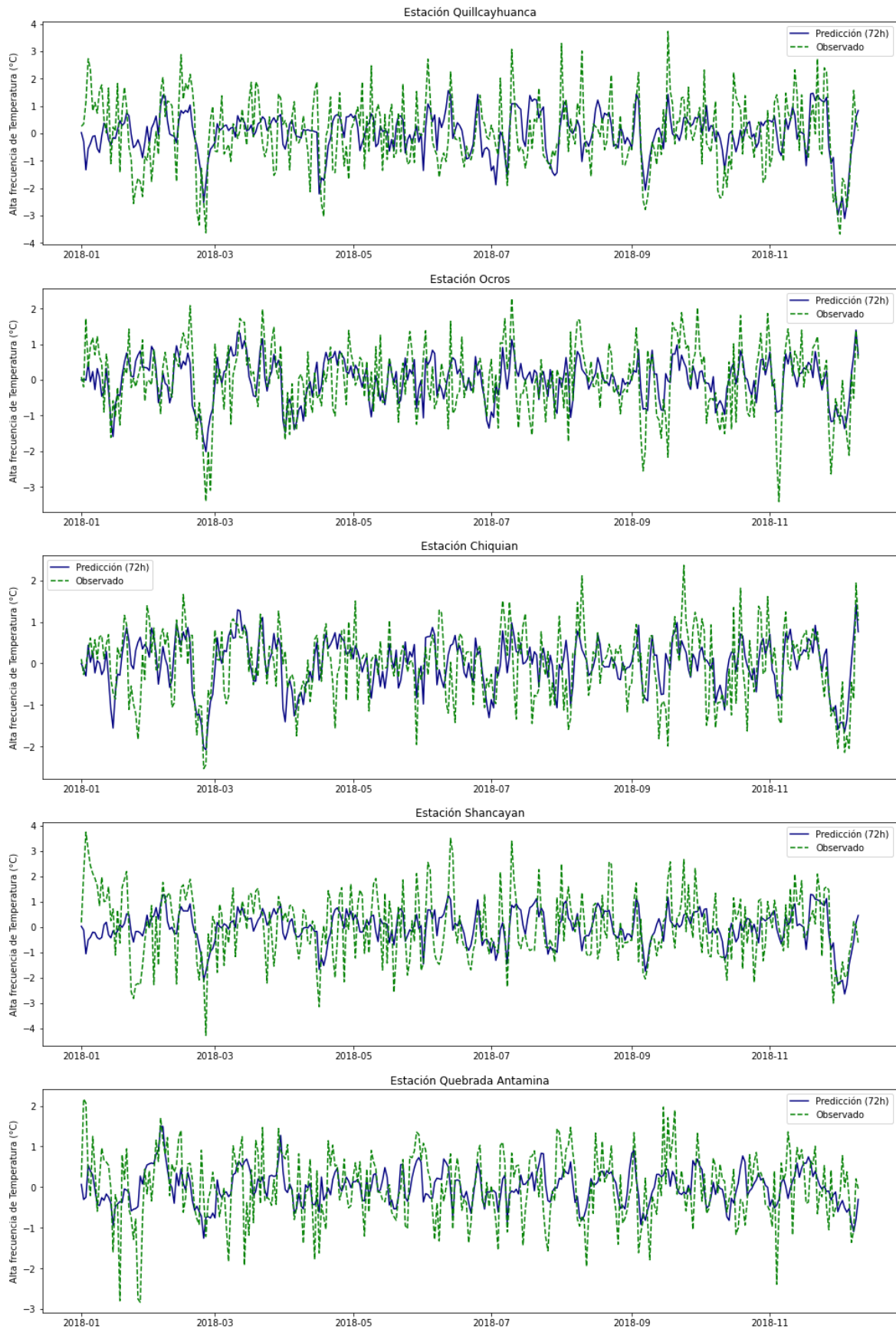


Estación Shilla

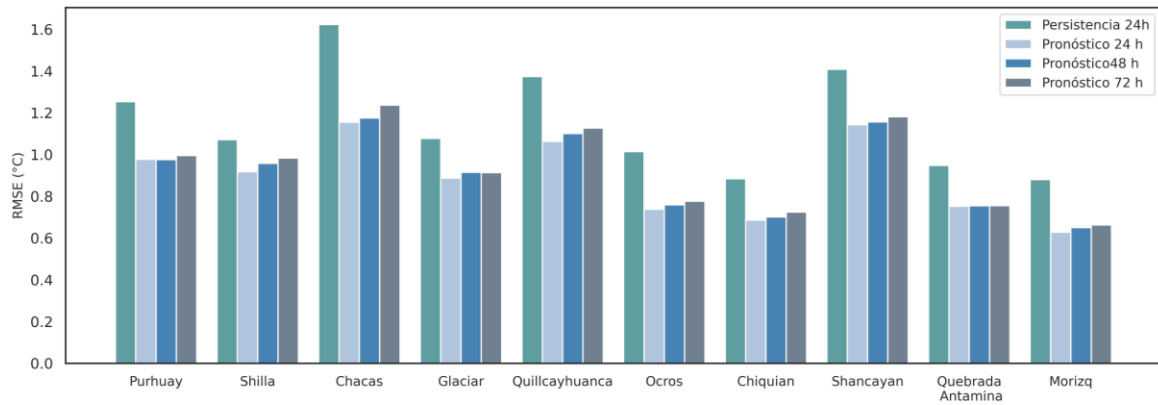


Estación Glaciar





**Figura 16:** Series temporales correspondientes a variabilidad de alta frecuencia de temperatura mínima pronosticada y observada para cada estación meteorológica



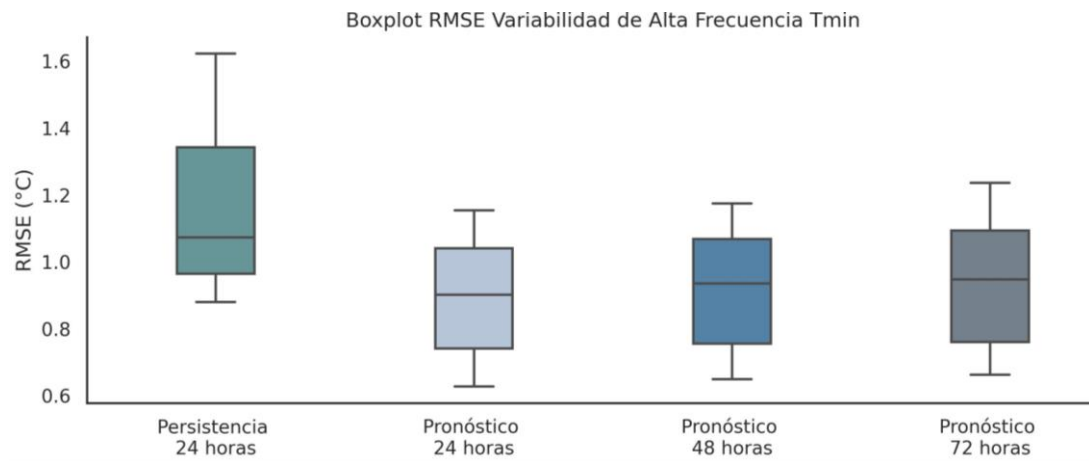
**Figura 17:** Gráfico de barras del RMSE para el pronóstico de temperatura mínima por persistencia de 24 horas y pronóstico con el modelo empírico con horizonte de pronóstico de 24h, 48h y 72h

La Figura 18 proporciona una visión general de la distribución del RMSE para cada día de pronóstico, teniendo en cuenta todas las estaciones meteorológicas analizadas, utilizando un diagrama de cajas. Además, se incluyó el error de pronóstico derivado de la persistencia de 24 horas, lo que significa asumir que la temperatura mínima de un día determinado será la misma que la del día anterior. Se observa que el pronóstico realizado con el modelo empírico presenta errores significativamente menores en todos los horizontes de pronóstico en comparación con el pronóstico por persistencia de 24 horas.

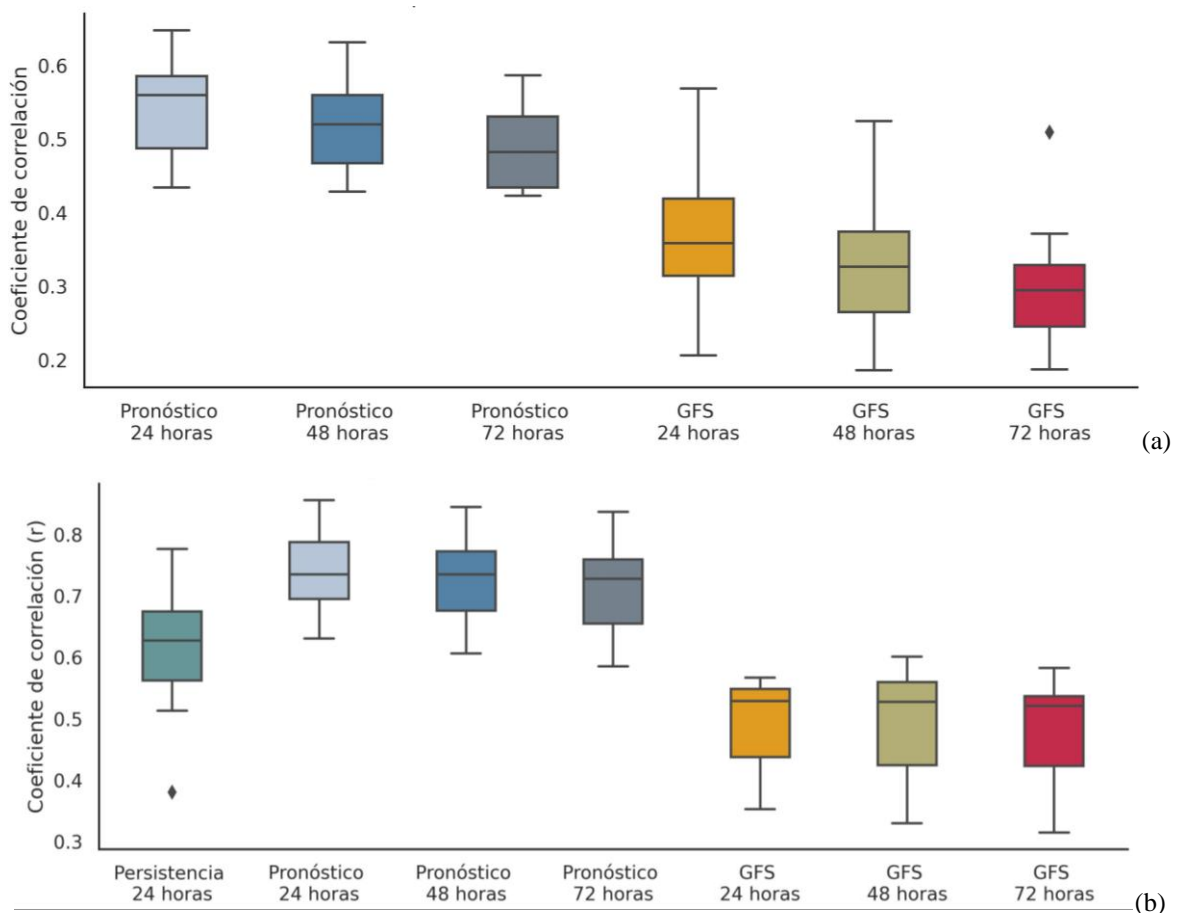
En la Figura 19, que muestra un diagrama de caja para los coeficientes de correlación lineal entre los valores observados y los valores predichos, se puede observar lo siguiente:

- En la Figura 19a, para el primer día de pronóstico, los coeficientes de correlación varían entre 0.44 y 0.65. La correlación tiende a disminuir a medida que se extiende el horizonte de pronóstico, llegando a un rango de 0.43-0.58 para el día 3. Es importante destacar que estos resultados se centran en la variabilidad de alta frecuencia, y los coeficientes de correlación en el rango de 0.4-0.6 se consideran una buena correlación. Se mostraron también los resultados de filtrar la variabilidad de alta frecuencia de los pronósticos directos del GFS indicando que el modelo empírico desarrollado presenta mejores resultados para captar la variabilidad diaria de esta variable en todos los horizontes de pronóstico analizados.
- En la Figura 19b, se presentan los resultados de correlación para los pronósticos finales de temperatura mínima, después de agregar la variabilidad de baja frecuencia para obtener los pronósticos finales de temperatura en cada estación. Se observa un aumento en la correlación, alcanzando un rango de 0.6-0.85, lo que indica una buena correlación entre los pronósticos y los valores observados en este escenario final.

Estos resultados sugieren que el modelo empírico utilizado en esta investigación tiene un desempeño sólido en la predicción de la temperatura mínima diaria, superando significativamente las predicciones basadas en la persistencia de 24 horas, así como con el pronóstico directo del GFS en los 3 días de pronóstico analizados.



**Figura 18:** Diagrama de cajas de RMSE para pronóstico de temperatura mínima a diferentes horizontes de tiempo y con el método por persistencia de 24 horas



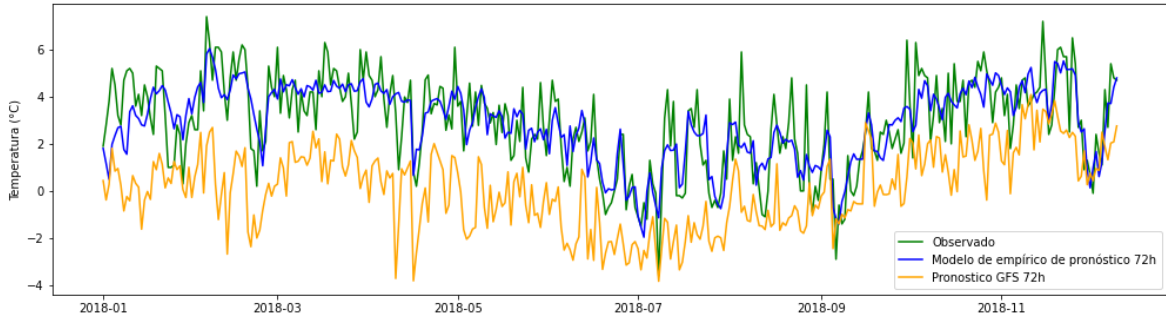
**Figura 19:** (a) Diagrama de cajas de correlación de Pearson para pronóstico alta frecuencia de temperatura mínima (b) Diagrama de cajas de correlación de Pearson para pronósticos finales de temperatura mínima

La Figura 20 muestra los meteogramas resultantes después de agregar la variabilidad de baja frecuencia para obtener los pronósticos finales de temperaturas mínimas diarias en las estaciones meteorológicas analizadas durante el período de validación. También se incluye el pronóstico directo de temperatura mínima obtenido del modelo GFS sin ningún tipo de procesamiento. En cuanto al pronóstico directo del modelo global GFS sobre el área de estudio, se observa que la mayoría de las estaciones meteorológicas presentaron pronósticos sesgados en comparación con los valores observados de temperatura mínima. Estos sesgos variaron entre las diferentes estaciones. En algunas estaciones, el modelo GFS tendió a subestimar la temperatura mínima, como en los casos de Shilla, Shancayan, Purhuay, Chacas y Quebrada Antamina. Por otro lado, hubo estaciones donde el modelo GFS sobrestimó sus pronósticos, como en las estaciones Morizq, Ocos y Glaciar. Solo las estaciones Chiquian y Quillcayhuanca presentaron pronósticos que se ajustaron correctamente a los valores observados. Los errores en el pronóstico del modelo global podrían estar relacionados principalmente con la representación inadecuada de la topografía en el modelo, que no refleja adecuadamente la elevación real de la zona de estudio. Además, estos errores podrían ser atribuidos a la incapacidad de los modelos numéricos para manejar procesos atmosféricos a una escala de subgrilla, lo que lleva a simplificaciones en la topografía y a la omisión de efectos locales a pequeña escala que influyen en el clima local (Durai y Bhradwaj, 2014).

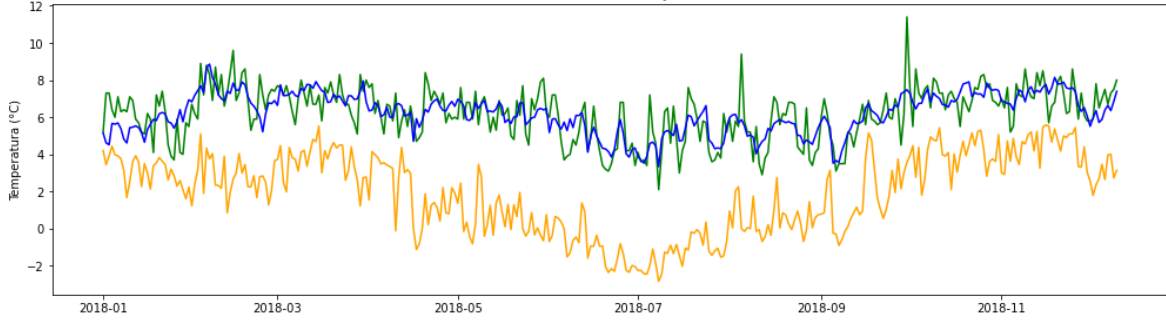
Comparando los pronósticos finales del modelo empírico con los valores observados, especialmente para el último horizonte de pronóstico de 72 horas, se observa que se lograron pronósticos correctos en la mayoría de las estaciones. Sin embargo, se notó una disminución en la exactitud durante los meses de verano, donde se identificaron las mayores diferencias entre los valores pronosticados y observados en muchas estaciones. A pesar de esto, resulta evidente que los pronósticos empíricos mejoran significativamente en comparación con el pronóstico directo obtenido del modelo GFS.



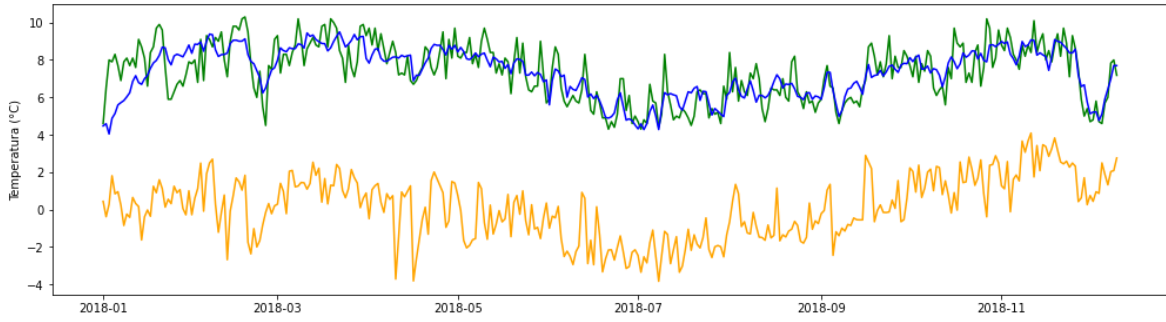
Pronóstico de temperatura mínima  
Estación Chacas



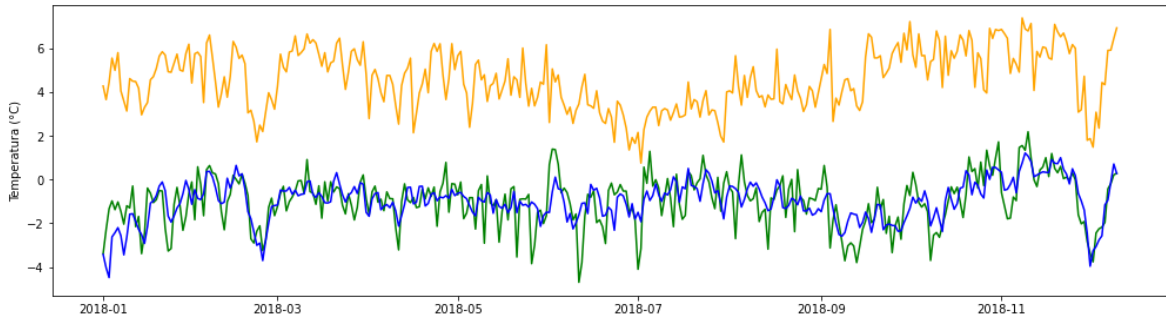
Estación Purhuay



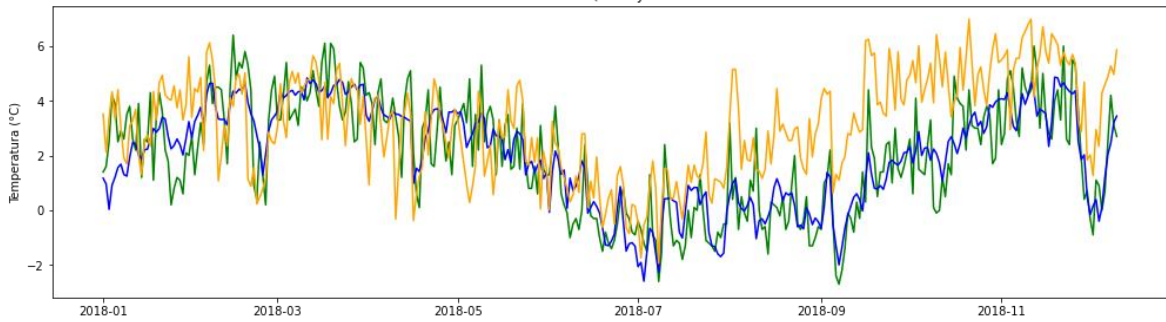
Estación Shilla



Estación Glaciar



Estación Quillcayhuanca





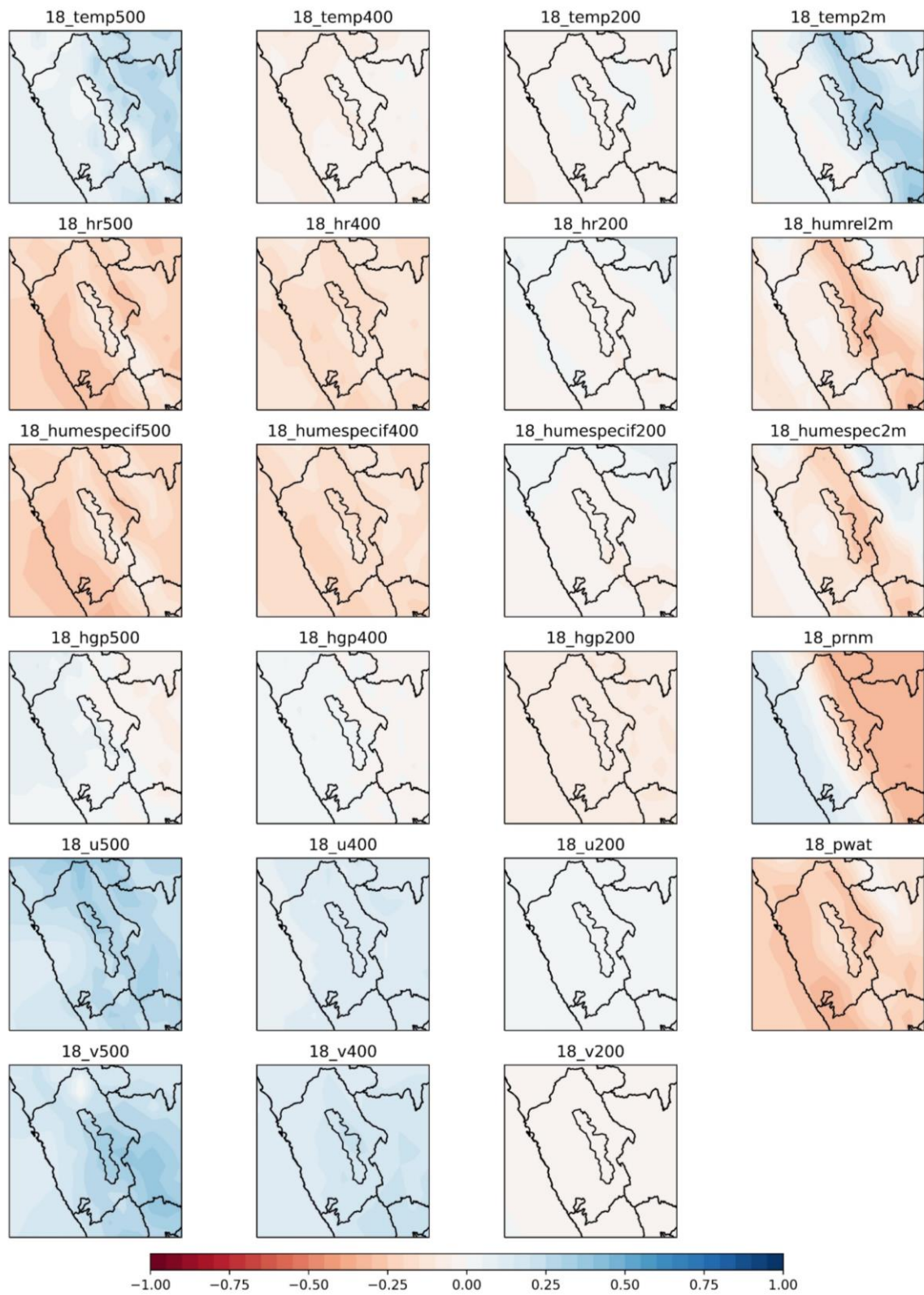
**Figura 20:** Pronóstico final de temperatura mínima en el periodo de validación

## 4.2. Modelo de pronóstico de temperatura máxima

En esta sección se analizó la relación entre la variabilidad de alta frecuencia de la temperatura máxima diaria y las variables meteorológicas del modelo GFS correspondientes a las 18 UTC. A continuación, se resumen los resultados derivados del análisis de la Figura 21, la cual muestra los mapas de correlación espacial de Pearson para temperatura máxima:

- Las variables en el nivel sinóptico de 200 hPa mostraron una correlación baja con la alta frecuencia de temperatura máxima, por lo que no se incluyeron en el modelo de pronóstico.
- En el nivel de 500 hPa, las variables de humedad específica y relativa presentaron correlaciones similares, en el rango de -0.45 a -0.25, lo que justificó su inclusión como potenciales predictores en el modelo de pronóstico.
- La altura geopotencial en el nivel de 500 hPa mostró un grado de correlación menor, mientras que la variable de viento zonal presentó una correlación negativa en todo el dominio espacial, especialmente cerca del área de estudio.
- En el nivel de 400 hPa, se observaron patrones de correlación similares a los del nivel de 500 hPa, aunque con una ligera disminución en el grado de correlación.
- Las variables a nivel de superficie, como la temperatura a 2 metros del suelo, la humedad relativa y específica, presentaron una mayor correlación con las variaciones de temperatura máxima. La temperatura a 2 metros del suelo mostró una correlación positiva, mientras que la humedad relativa y específica presentaron correlaciones negativas.
- La presión reducida al nivel del mar mostró un patrón de correlación inverso al observado en la temperatura mínima, con valores negativos en la zona occidental y valores fuertemente positivos en la zona oriental del área de estudio.
- El contenido de humedad en toda la columna atmosférica, representado por el PWAT, presentó una correlación considerable con las temperaturas máximas y mínimas en la zona de estudio.

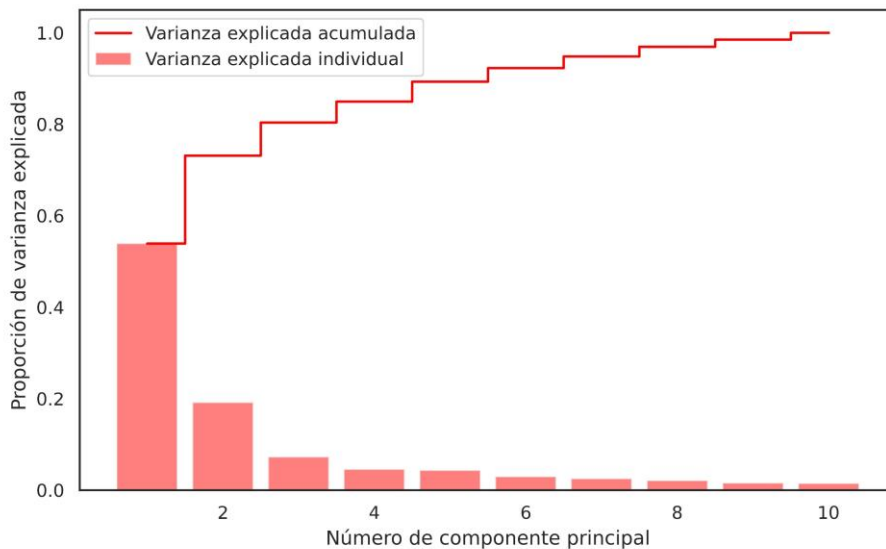
Estos resultados proporcionan información importante para la selección de predictores en el modelo de pronóstico de temperatura máxima diaria.



**Figura 21:** Mapas de correlación de Pearson de variabilidad de alta frecuencia de temperatura máxima y variables meteorológicas del modelo GFS correspondientes a las 18 UTC en la zona de estudio

La Figura 22 ilustra la varianza explicada por cada una de las componentes principales obtenidas al aplicar PCA sobre los datos de alta frecuencia de temperatura máxima en las estaciones meteorológicas del grupo de predicción. A diferencia de lo observado en el análisis de la temperatura mínima, la primera componente principal explica alrededor del 55% de la varianza en el conjunto de datos, lo que representa un aumento del 10% en comparación con lo que se obtuvo en temperatura mínima. Las componentes segunda y tercera explican aproximadamente el 19% y el 7% de la varianza, respectivamente. A partir de la cuarta componente principal, los porcentajes de varianza explicada disminuyen por debajo del 5%, lo que llevó a su eliminación, ya que estas componentes con un bajo porcentaje de varianza explicada no logran resumir adecuadamente el conjunto original de datos, siguiendo la recomendación de Bro y Smilde (2014).

Es importante destacar que las primeras componentes obtenidas para la temperatura máxima explicaron una mayor proporción de la varianza en comparación con el análisis de la temperatura mínima. Por lo tanto, utilizando solo las tres primeras componentes principales como predictores sería suficiente para superar el umbral del 80% de varianza explicada. Sin embargo, se optó por incluir las cinco primeras componentes principales, que acumulan cerca del 90% de la varianza, para igualar el número de componentes utilizadas como predictores en el pronóstico de la temperatura mínima. Es relevante señalar que la incorporación de estas dos componentes principales adicionales como predictores no generó cambios significativos en los resultados finales.



**Figura 22:** Varianza explicada individual y acumulada de las componentes principales obtenidas de la variabilidad de alta frecuencia de temperatura máxima proveniente de estaciones meteorológicas

La Tabla 8 presenta el número de componentes principales obtenidas a partir de las variables filtradas del modelo GFS a las 18 UTC, las cuales fueron retenidas como posibles predictores en el desarrollo de las ecuaciones de regresión lineal múltiple. Al igual que en el caso de la temperatura mínima, se estableció un umbral del 80% de varianza explicada para determinar cuántas componentes principales se conservarían. En el Anexo 2 se muestran los gráficos de varianza explicada individual y acumulada de cada variable meteorológica. Es relevante notar que las variables meteorológicas modeladas a 2 m sobre el suelo requirieron un mayor número de componentes principales para alcanzar el umbral de varianza establecido, lo cual concuerda con los resultados obtenidos para las 12 UTC. Específicamente, se necesitaron 8 componentes principales para la humedad relativa y 6 componentes principales para la temperatura y la humedad específica.

**Tabla 8:** Número de componentes principales incluidas como potenciales predictores para cada variable meteorológica proveniente de modelo numérico para la predicción de temperatura máxima

Variable	Número de PCs
18_humespec2m	6
18_prnm	2
18_u500	2
18_temp500	3
18_v200	1
18_humespecif500	5
18_u400	1
18_hgp200	1
18_humrel2m	8
18_pwat	3
18_hr400	4
18_hr500	5
18_hgp400	1
18_v500	3
18_humespecif200	4
18_temp400	1
18_temp200	1
18_humespecif400	4
18_temp2m	6
18_u200	1
18_hr200	4



Los resultados del proceso de validación del pronóstico de alta frecuencia de temperatura máxima para los tres horizontes de pronóstico se presentan en las tablas 9, 10 y 11. El error cuadrático medio de la raíz (RMSE) con respecto a los valores observados en el primer día de pronóstico varía en el rango de 0.98°C a 1.94°C. Este intervalo se amplía para el segundo día, con un RMSE entre 1°C y 2°C, y para el tercer día de pronóstico, donde se alcanza un RMSE entre 1.01°C y 2.05°C. Es importante destacar que el modelo de pronóstico para la temperatura máxima presenta un grado de error mayor en comparación con el pronóstico de temperatura mínima.

La Figura 24 ilustra el comportamiento del RMSE para cada una de las estaciones en cada horizonte de pronóstico. Se observa que el error de pronóstico tiende a aumentar a medida que se extiende el horizonte de pronóstico en la mayoría de las estaciones, con algunas excepciones, como Glaciar y Shancayan, donde el error en el segundo día es ligeramente menor que en los otros horizontes de pronóstico. Este aumento en el RMSE se produce en un rango de décimas de grado entre las 24 y 72 horas de pronóstico.

En función del intervalo de RMSE, los resultados del pronóstico de temperatura máxima se clasifican como predicciones de calidad intermedia para las estaciones Glaciar, Ocos y Shilla. El resto de las predicciones se consideran de mala calidad según los intervalos de RMSE utilizados por Qona'ah et al. (2020). A pesar del aumento en el error, las predicciones para la temperatura máxima muestran correlaciones más altas en el 70% de las estaciones meteorológicas en comparación con los resultados de temperatura mínima, lo que indica que, aunque las predicciones no son tan precisas en magnitud como las de la temperatura mínima, capturan mejor el comportamiento de la variable.

Al comparar estos resultados con los obtenidos por modelos empíricos de pronóstico desarrollados por Aliaga-Nestares et al. (2023), que obtuvieron valores de RMSE en el rango de 0.76°C –1.76°C y 0.57°C –1.69°C para temperatura máxima y mínima respectivamente, se observa una mejora significativa en los resultados de temperatura mínima con un RMSE de 0.62°C-1.15°C. Sin embargo, el modelo para el pronóstico de temperatura máxima muestra un rendimiento considerablemente inferior, con un RMSE que supera los 2°C en algunas estaciones. Además, los resultados obtenidos son en su mayoría inferiores a los obtenidos por Qona'ah et al. (2020), quienes lograron un RMSE para las 4 estaciones de Indonesia en un rango entre 0.94°C-1.08°C para el pronóstico de temperatura máxima.

**Tabla 9:** Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura máxima en el periodo de validación con un horizonte de pronóstico de 24 horas

Estación	MSE (°C)	RMSE (°C)	r
Shilla	1.401	1.184	0.619
Shancayan	1.811	1.346	0.557
Ocros	1.051	1.025	0.598
Chiquian	1.698	1.303	0.607
Purhuay	3.785	1.945	0.561
Quillcayhuanca	1.725	1.313	0.570
Chacas	2.887	1.699	0.639
Quebrada Antamina	2.793	1.671	0.641
Morizq	1.527	1.236	0.633
Glaciar	0.963	0.981	0.423

**Tabla 10:** Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura máxima en el periodo de validación con un horizonte de pronóstico de 48 horas

Estación	MSE (°C)	RMSE (°C)	r
Shilla	1.341	1.158	0.639
Shancayan	1.672	1.293	0.598
Ocros	1.175	1.084	0.542
Chiquian	1.764	1.328	0.587
Purhuay	3.984	1.996	0.533
Quillcayhuanca	1.795	1.340	0.550
Chacas	3.030	1.741	0.617
Quebrada Antamina	2.938	1.714	0.618
Morizq	1.638	1.280	0.600
Glaciar	0.985	0.993	0.405



**Tabla 11:** Resultados por estación obtenidos del pronóstico de variabilidad de alta frecuencia de temperatura máxima en el periodo de validación con un horizonte de pronóstico de 72 horas

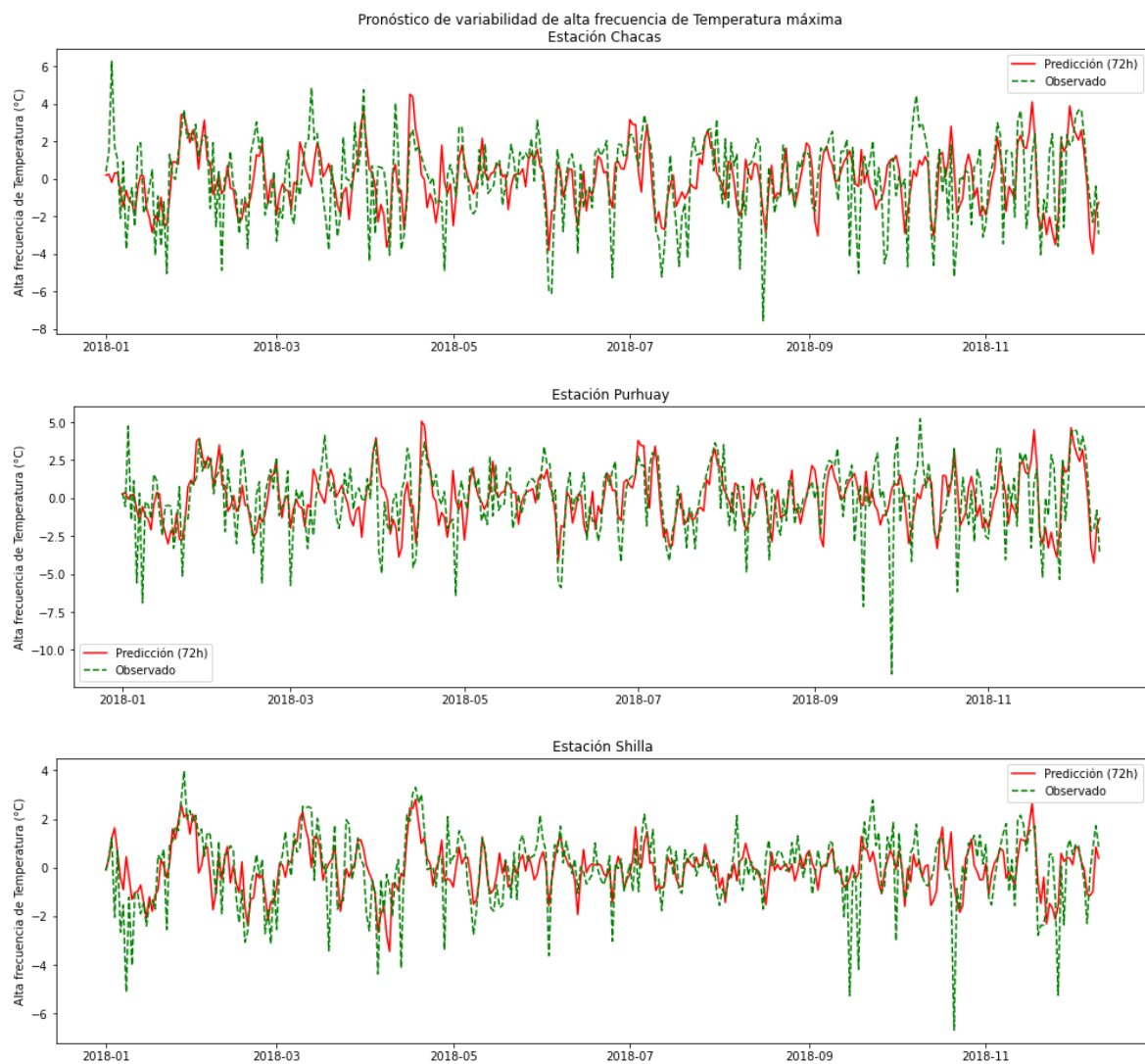
Estación	MSE (°C)	RMSE (°C)	r
Shilla	1.468	1.212	0.594
Shancayan	1.773	1.332	0.566
Ocros	1.213	1.101	0.519
Chiquian	1.790	1.338	0.578
Purhuay	4.218	2.054	0.510
Quillcayhuanca	1.974	1.405	0.499
Chacas	3.488	1.868	0.552
Quebrada Antamina	3.073	1.753	0.602
Morizq	1.731	1.316	0.575
Glaciar	1.028	1.014	0.373

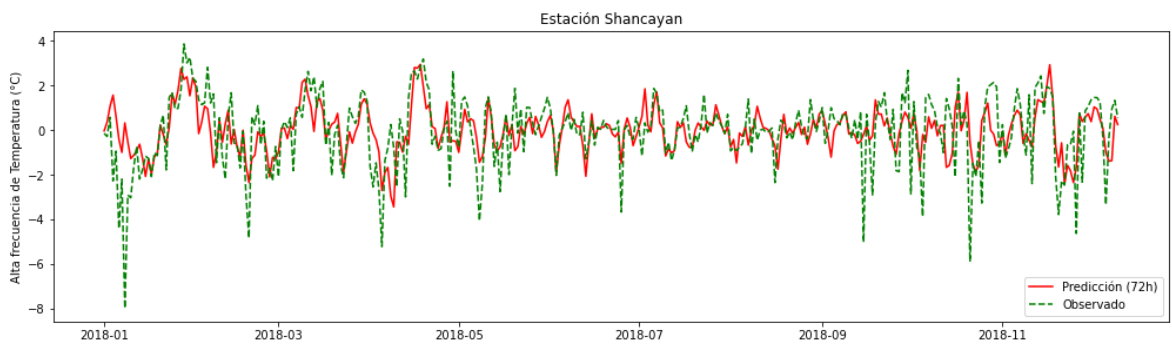
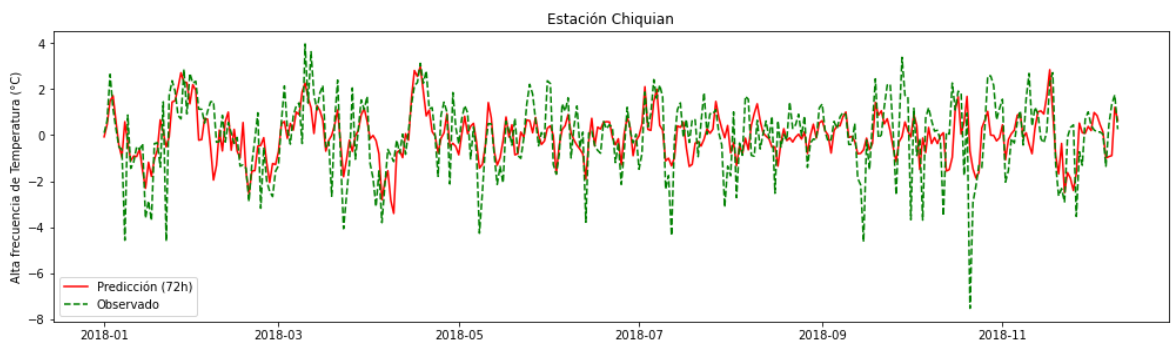
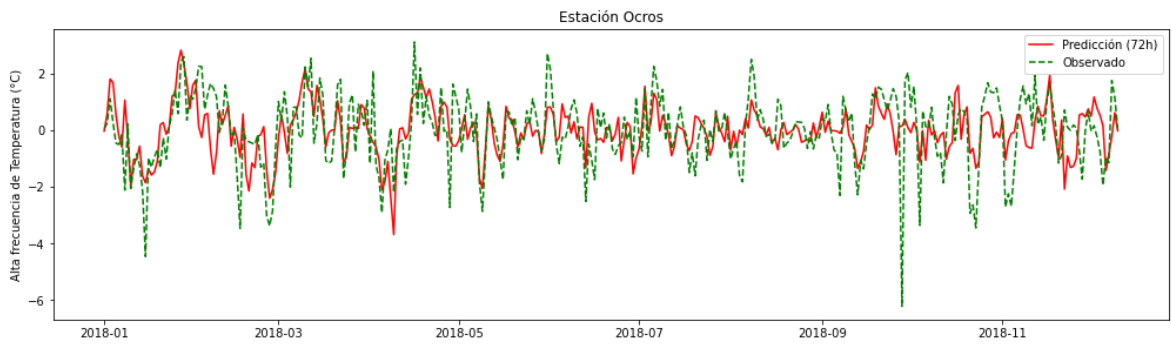
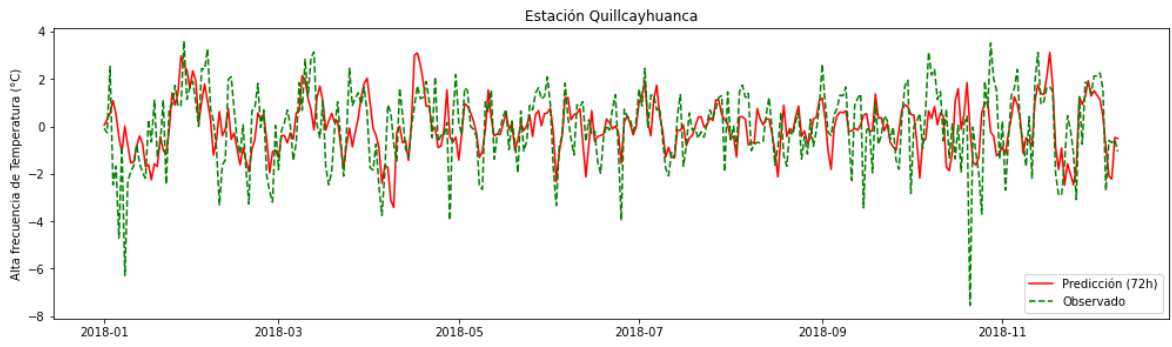
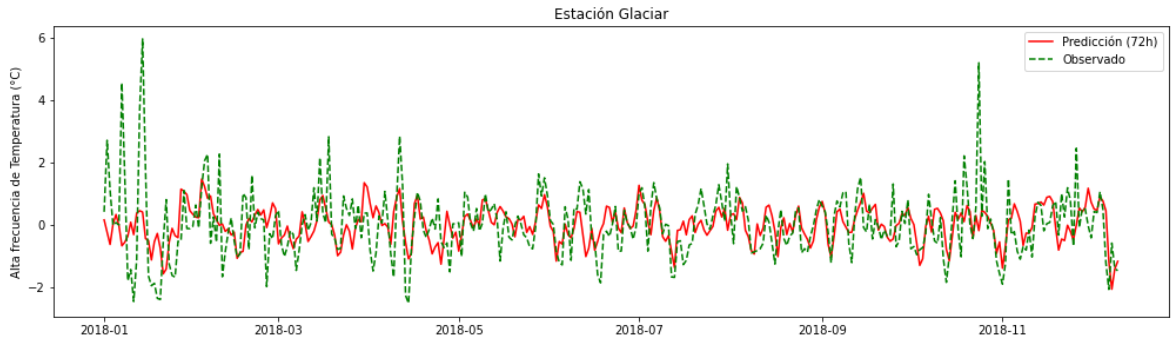
Las series de tiempo generadas durante el proceso de validación se presentan en la Figura 23, que muestra los pronósticos de alta frecuencia de temperatura máxima para el año 2018. En estas series de tiempo, se observa un mayor grado de variabilidad en comparación con la alta frecuencia de temperatura mínima. Este aumento en la variabilidad está relacionado con los mayores valores de RMSE obtenidos en los resultados. A pesar de este incremento en el error, se observa que los pronósticos logran identificar correctamente el comportamiento de la temperatura máxima en la mayoría de las estaciones meteorológicas.

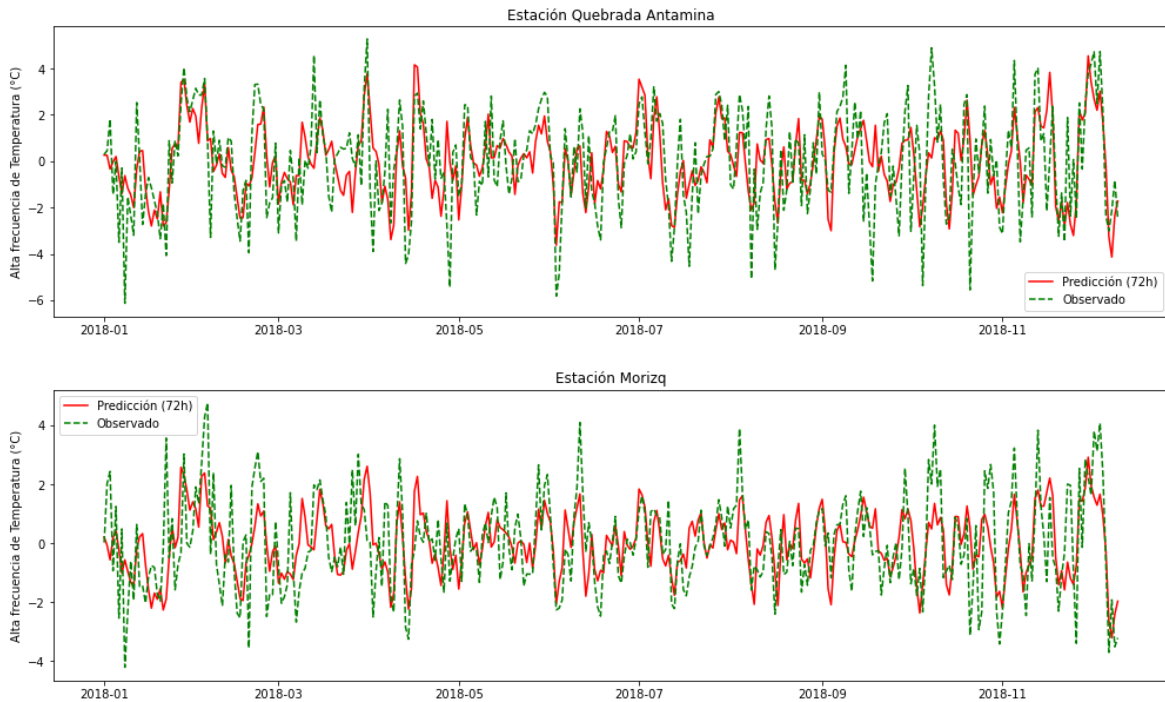
Este mejor desempeño se refleja en un aumento de la correlación, que se sitúa por encima de 0.5 en todas las estaciones meteorológicas, a excepción de la estación Glaciar. Esta estación, al ser de origen glaciar, presentó mayores desafíos en las predicciones por parte del modelo de pronóstico.

En relación con los eventos extremos, se observa una capacidad predictiva sólida durante la mayor parte del período de validación. Sin embargo, se presentan limitaciones en el pronóstico de los descensos más intensos de temperatura máxima. Estos descensos más intensos pueden estar asociados a fenómenos que no son captados adecuadamente por el modelo global y en consecuencia, el modelo empírico no puede preverlos de manera efectiva.

Según Dai et al. (1999), la nubosidad desempeña un papel importante en el comportamiento de las temperaturas máximas y mínimas. Las nubes pueden reducir la temperatura máxima al reflejar la luz solar y aumentar la temperatura mínima al incrementar la radiación de onda larga hacia la superficie. Además, la humedad atmosférica también influye en ambas variables. De manera similar, Aliaga-Nestares et al. (2023) mencionan la importancia de la nubosidad en el comportamiento de las temperaturas extremas diarias. Estas observaciones subrayan la relevancia de considerar la nubosidad y la humedad atmosférica como factores importantes en el desarrollo de modelos de pronóstico. La limitación de los modelos empíricos para identificar completamente los eventos más extremos podría estar relacionada con la falta de información directa sobre la cobertura nubosa entre los predictores del modelo global.







**Figura 23:** Series temporales correspondientes a variabilidad de alta frecuencia de temperatura mínima pronosticada y observada para cada estación meteorológica

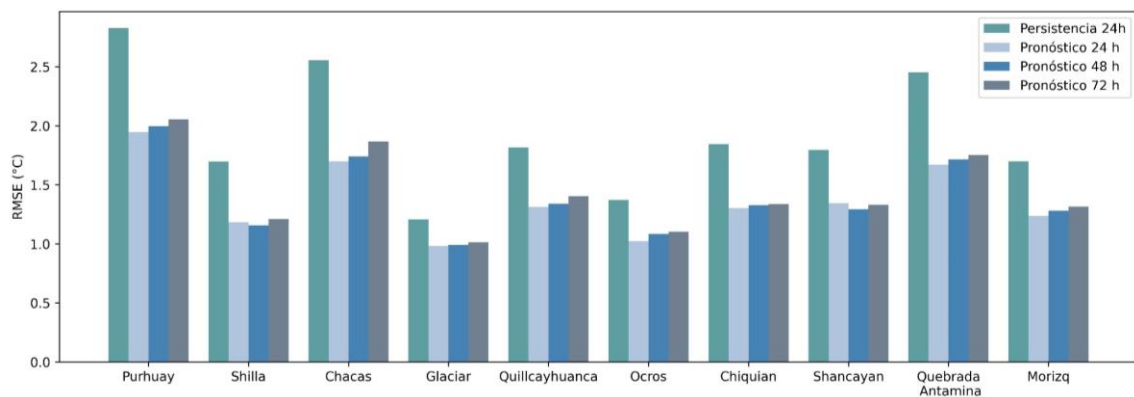
El diagrama de cajas representado en la Figura 25 muestra la distribución de RMSE para cada día de pronóstico. Se destaca que el error asociado al método de persistencia de 24 horas tiene un RMSE significativamente mayor en comparación con el rendimiento del modelo empírico en cada uno de los tres horizontes de pronóstico.

La Figura 26a presenta un diagrama de caja para los coeficientes de correlación lineal entre los valores observados y los pronosticados por el modelo empírico. En el primer día de pronóstico, los coeficientes se sitúan en un rango de 0.55 a 0.65 y tienden a disminuir a medida que se extiende el horizonte de pronóstico. En general, se observa que solo una estación presenta valores atípicos, que corresponde a la estación Glaciar. Como se explicó previamente, esta estación presentó desafíos en la modelación del comportamiento de la temperatura máxima debido a su naturaleza glaciar. Adicionalmente se añadió la correlación de temperatura máxima al filtrar la alta frecuencia del modelo GFS, dando como resultado que el modelo empírico de pronóstico también presenta correlaciones más altas con respecto al modelo global.

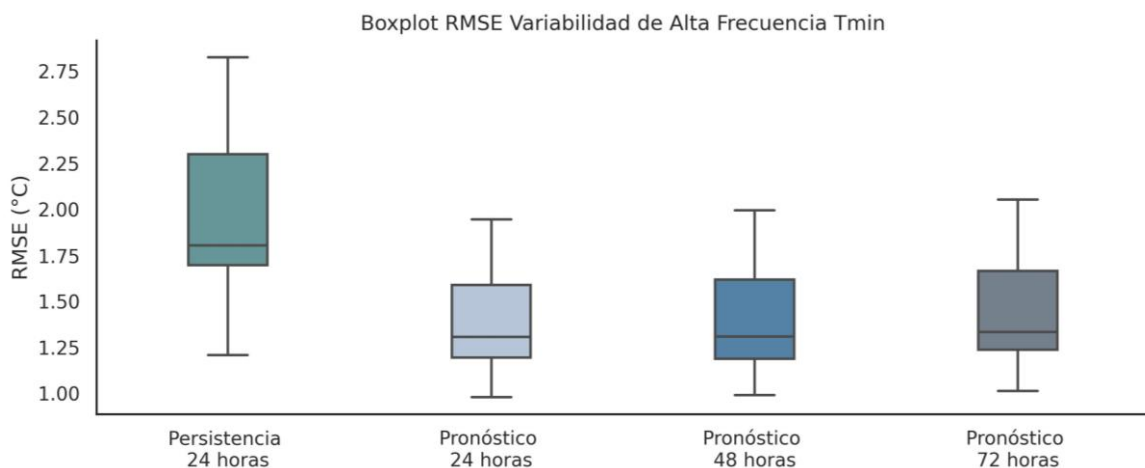
En la Figura 26b, se presentan los resultados de correlación para los pronósticos finales de temperatura máxima. Estos resultados muestran la capacidad predictiva sobresaliente del modelo empírico de pronóstico, con valores de correlación que superan 0.60 y llegan a

alcanzar hasta 0.78. En general, se logra una correlación fuerte entre los valores observados y los pronosticados. La correlación es considerablemente mayor a la que se obtuvo con el pronóstico por persistencia y con el pronóstico directo del modelo GFS, el cual presenta valores outliers negativos, lo cual indica que el pronóstico en algunas estaciones se encuentra completamente desviado de los valores observados.

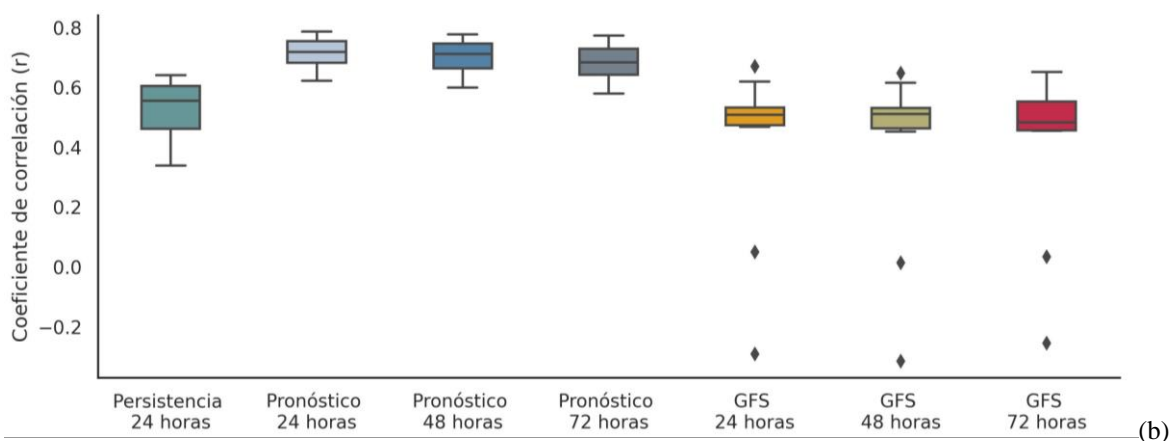
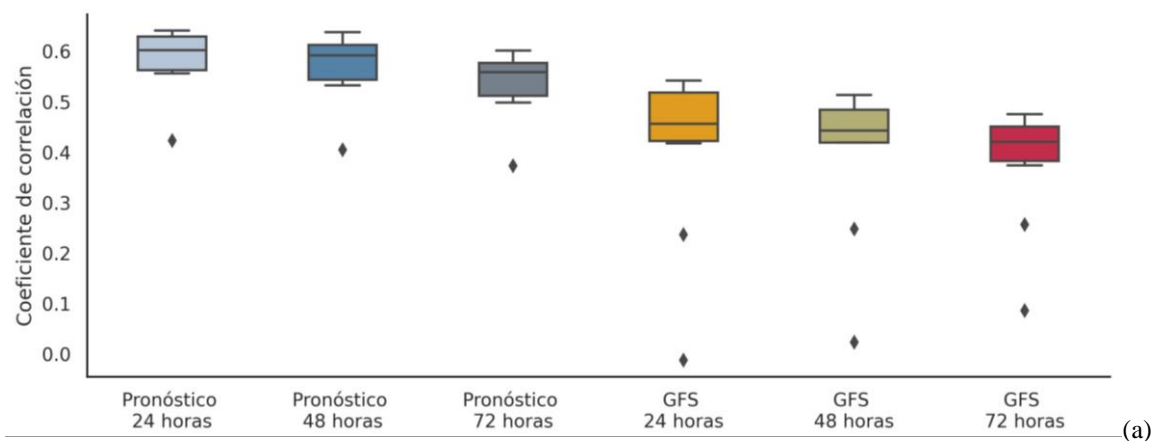
Estos resultados respaldan la eficacia del modelo empírico en la predicción de la temperatura máxima, a pesar de los desafíos asociados con eventos extremos y estaciones particulares como la Glaciar.



**Figura 24:** Gráfico de barras del RMSE para el pronóstico de temperatura máxima por persistencia de 24 horas y pronóstico con el modelo empírico con horizonte de pronóstico de 24h, 48h y 72h.



**Figura 25:** Box-Plot del RMSE para pronóstico de temperatura máxima a diferentes horizontes temporales y por el método por persistencia de 1 día

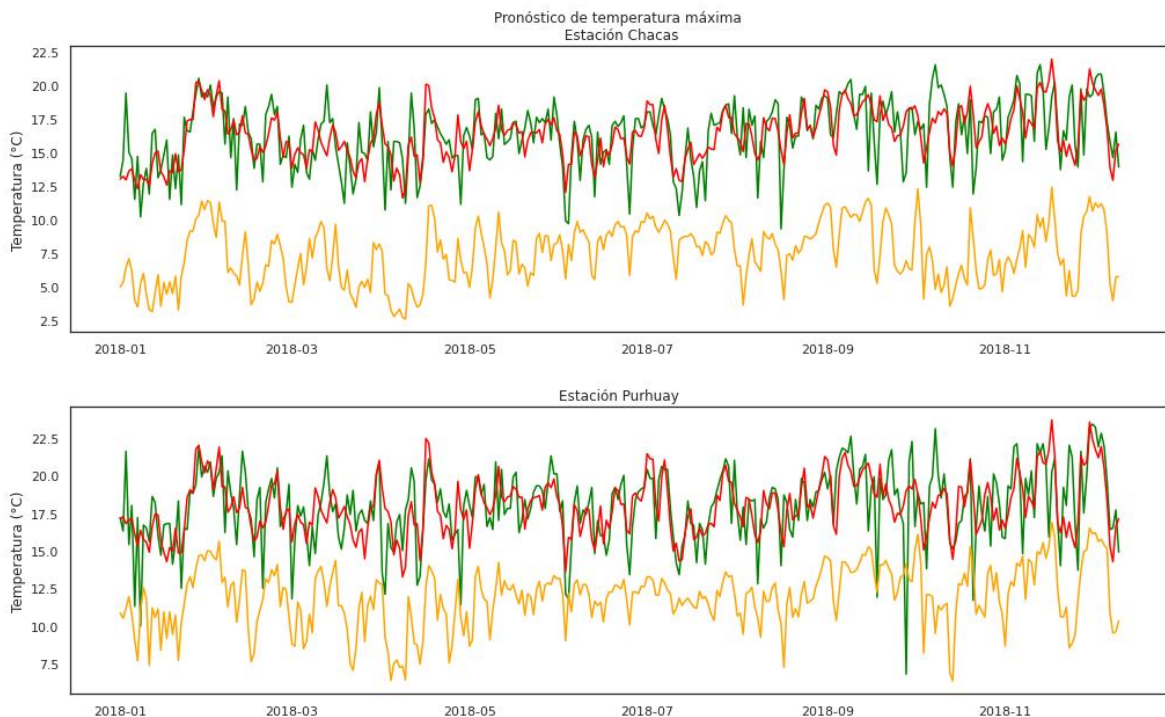


**Figura 26:** (a) Diagrama de cajas de correlación de Pearson para pronóstico alta frecuencia de temperatura máxima (b) Diagrama de cajas de correlación de Pearson para pronósticos finales de temperatura máxima

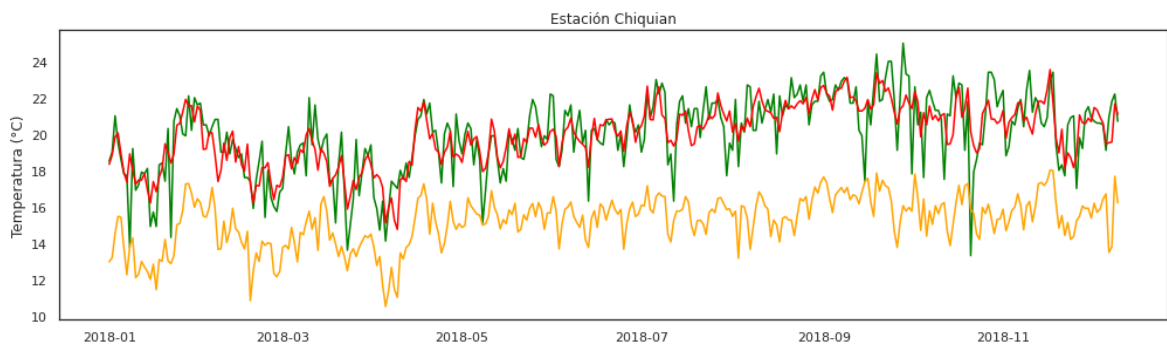
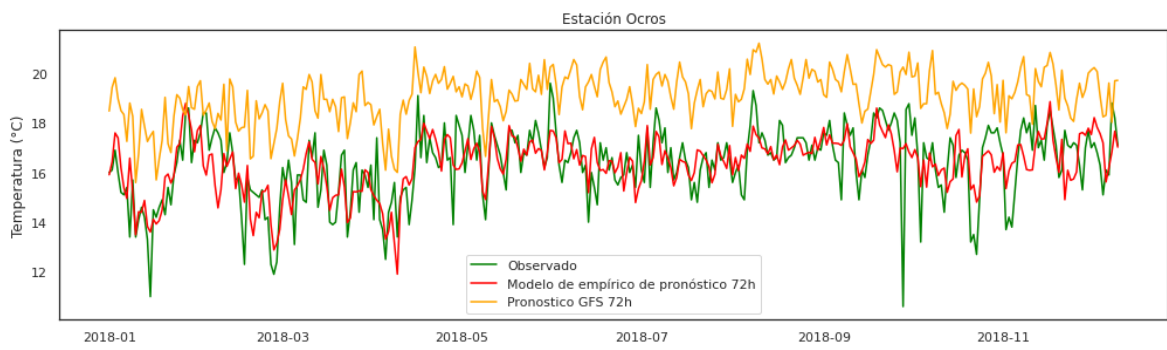
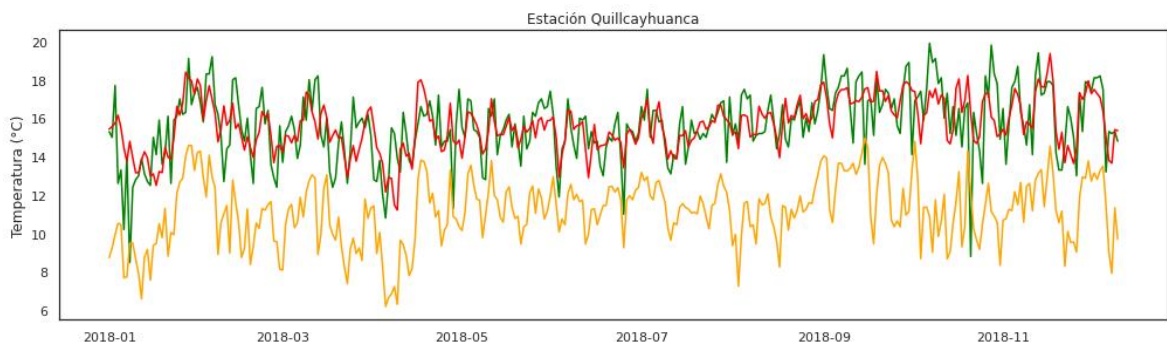
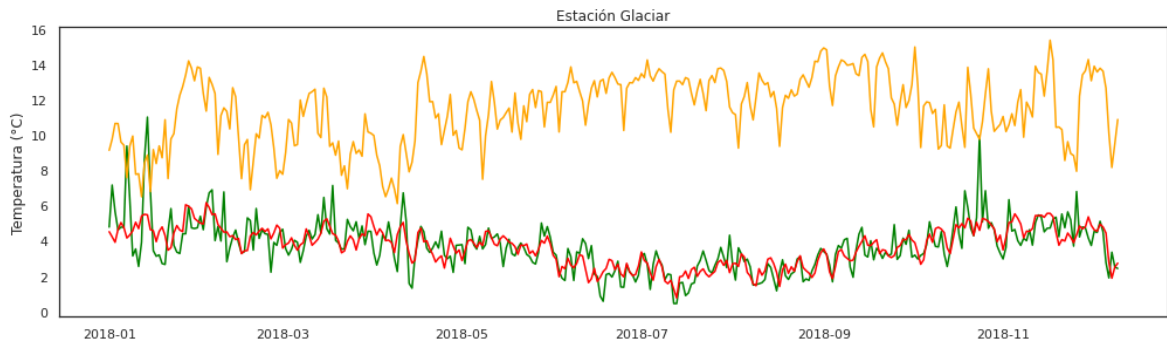
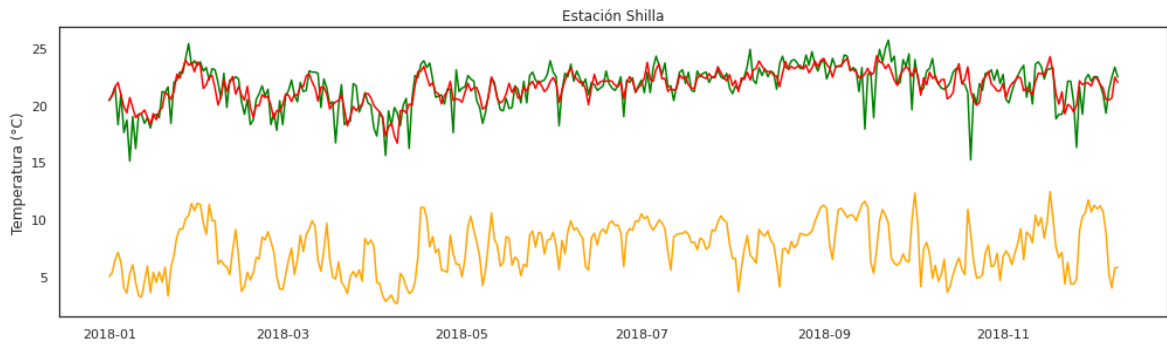
La Figura 27 exhibe las series temporales de la temperatura máxima diaria observada, los pronósticos finales del modelo empírico y los pronósticos directos sin posprocesamiento del GFS, todos con una anticipación de 72 horas. Se aprecia una notoria discrepancia en los pronósticos directos del GFS con respecto a la temperatura observada, con diferencias de varios grados Celsius. El GFS subestimó la temperatura máxima en las estaciones Shilla, Shancayan, Chiquian, Purhuay, Quillcayhuanca y Chacas. Por otro lado, sobrestimó la temperatura en las estaciones Ocros, Morizq y Glaciar. Únicamente en la estación Quebrada Antamina, los pronósticos del GFS se asemejaron a la temperatura máxima observada. Esta tendencia del modelo GFS a subestimar o sobrestimar los pronósticos fue similar a lo observado para la temperatura mínima en el 70% de las estaciones, lo que sugiere que gran parte de los errores de pronóstico del GFS podrían deberse a inexactitudes en la topografía

del modelo, que coloca estos puntos de pronóstico a altitudes que difieren considerablemente de la realidad.

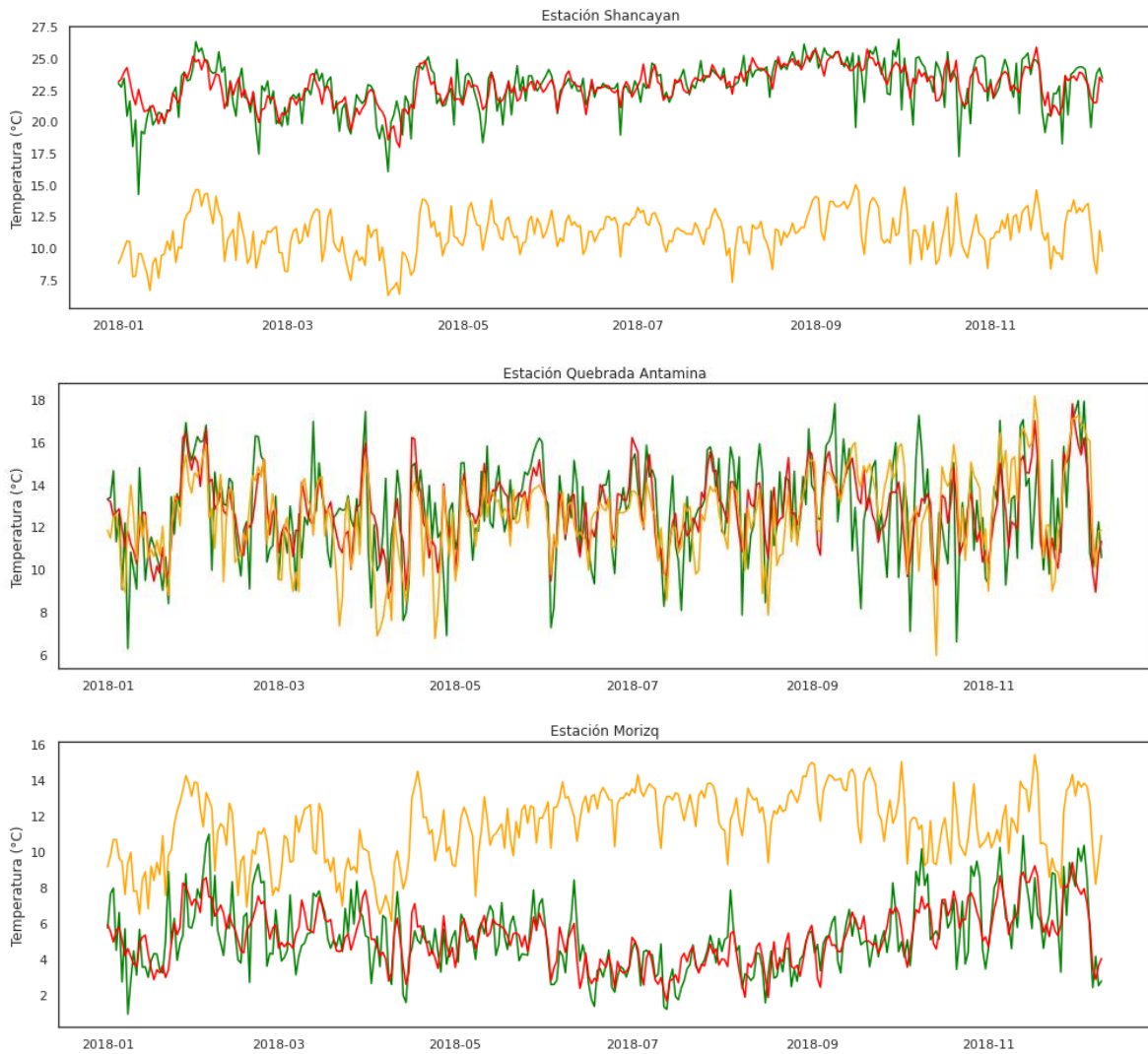
En contraste, los pronósticos finales del modelo empírico, en todos sus horizontes de pronóstico, superaron significativamente los pronósticos directos del GFS sin posprocesamiento. El modelo empírico logró modelar adecuadamente la temperatura máxima, una variable que exhibe menos estacionalidad que la temperatura mínima y está influenciada en gran medida por factores locales. Sin embargo, se observaron algunas limitaciones en la identificación de ciertos eventos de descenso de la temperatura máxima que el modelo no pudo prever. Un análisis adicional reveló que muchos de estos descensos de temperatura ocurrieron en días lluviosos, lo que sugiere que la nubosidad, relacionada con los eventos de precipitación, podría tener un gran impacto en la ocurrencia de estos eventos no identificados por el modelo empírico ni por el modelo GFS. En resumen, el modelo empírico superó al GFS en la predicción de la temperatura máxima, pero aún presenta desafíos en la identificación de ciertos eventos extremos relacionados con factores locales y condiciones climáticas específicas.











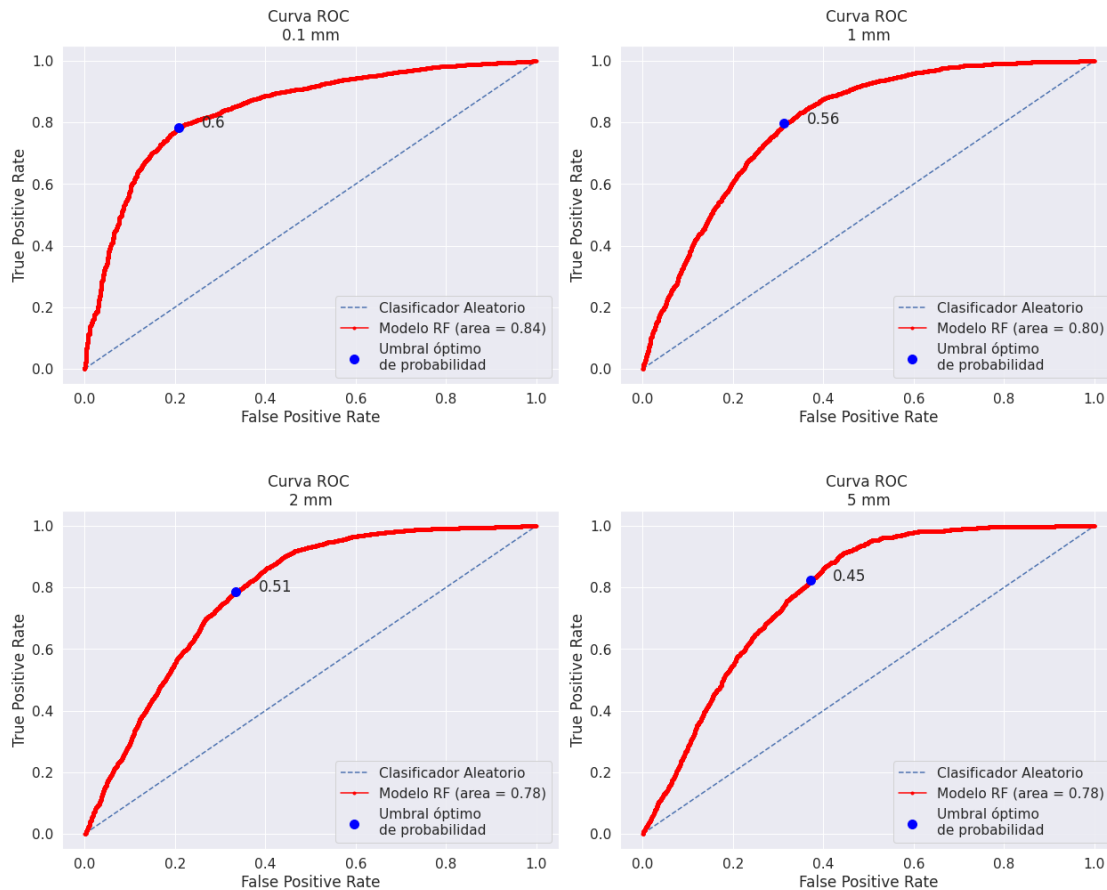
**Figura 27:** Pronóstico final de temperatura mínima en el periodo de validación

La aplicación de los modelos de pronóstico de temperaturas extremas diarias, se encuentra disponible en el enlace: <https://andy7864.github.io/AncashPrediccionTemperatura/>.

En el anexo 3 se muestra un ejemplo de pronóstico generado por el modelo de predicción de temperaturas extremas para los 10 puntos de pronóstico y 3 días de antelación, desarrollado en la presente tesis.

### 4.3. Modelo de pronóstico de probabilidad de precipitación

La eficacia de los pronósticos de probabilidad de precipitación se evaluó utilizando varias métricas asociadas a tareas de clasificación. Los resultados se presentaron para diferentes umbrales de precipitación y se tomó en cuenta todos los puntos de grilla dentro del área de pronóstico utilizados para entrenar el modelo Random Forests.



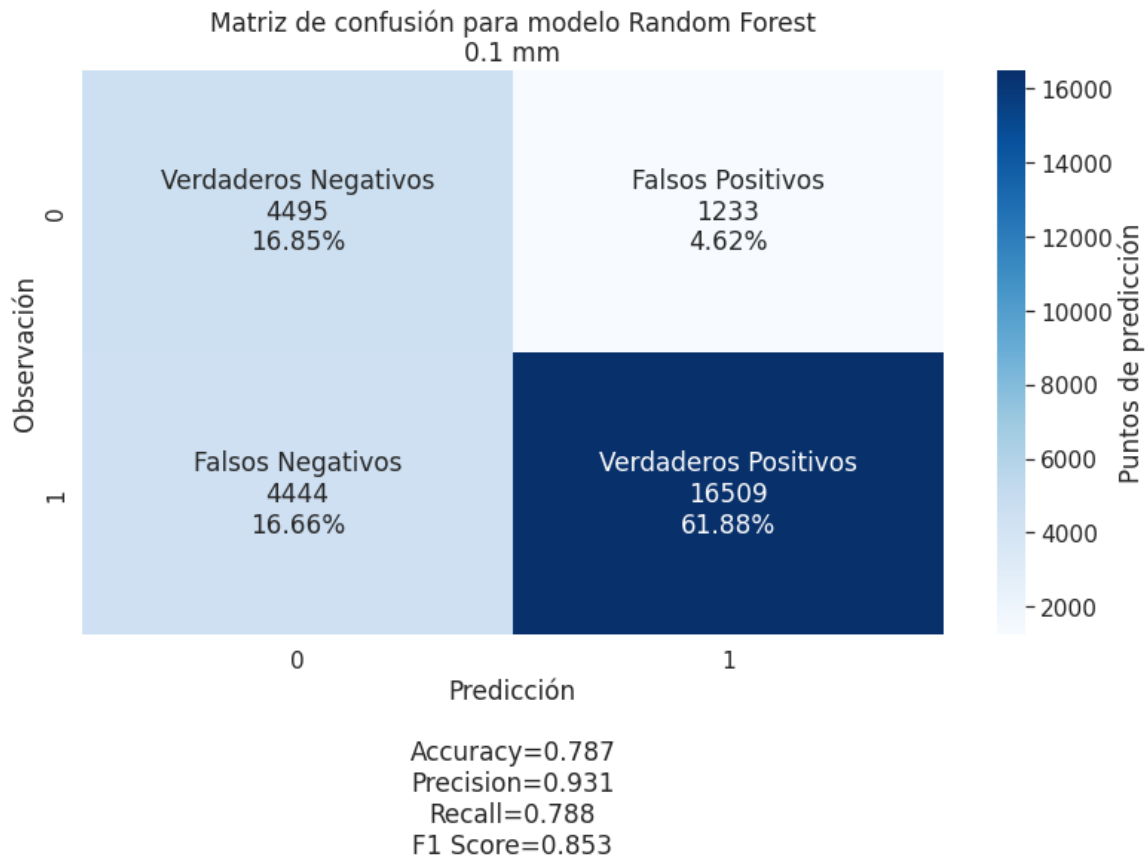
**Figura 28:** Curvas ROC para distintos umbrales de precipitación

La Figura 28 muestra las curvas ROC para el período de testeo en diferentes umbrales de precipitación. Se observó que a medida que se aumenta el umbral de precipitación, la capacidad de discriminación de los eventos disminuye de manera constante, lo que se reflejó en la reducción del área bajo la curva ROC (AUC). El AUC disminuyó de 0.84 para el umbral de 0.1 mm a 0.78 para los umbrales de 2 mm y 5 mm. Estos resultados contrastan con los obtenidos por Jha et al. (2018), quienes encontraron que la capacidad de discriminación del pronóstico era mayor para eventos de lluvia más intensa ( $> 5$  mm) con un tiempo de anticipación de 1 día. Observaron que las curvas ROC para este umbral estaban más cerca de la esquina izquierda que para eventos de precipitación más ligera ( $< 0.2$  mm).

La curva ROC permitió identificar el umbral óptimo de probabilidad en cada uno de los modelos. Este umbral indica la probabilidad a la cual se distinguen los eventos positivos de los negativos y donde se logra la mayor tasa de verdaderos positivos y la menor tasa de falsos positivos.

Se encontró que el umbral óptimo de probabilidad disminuyó a medida que se incrementó el umbral de precipitación. Para el umbral de 0.1 mm, el umbral óptimo fue de 0.6, para el umbral de 1 mm fue de 0.56, para el umbral de 2 mm fue de 0.51 y, finalmente, para el umbral de 5 mm de precipitación, fue de 0.45. Estos resultados son ligeramente diferentes a los obtenidos por De la Cruz (2016), quien encontró umbrales óptimos para estaciones en la Sierra del Perú en el rango de 0.2 a 0.4, variando según el umbral de precipitación del modelo. Sin embargo, se observa una tendencia similar en cuanto a que el umbral de probabilidad óptimo tiende a ser menor para categorías mayores de precipitación.

La capacidad de discriminación de eventos de lluvia varió según el umbral de precipitación, y el umbral óptimo de probabilidad disminuyó a medida que se aumentó el umbral de precipitación.



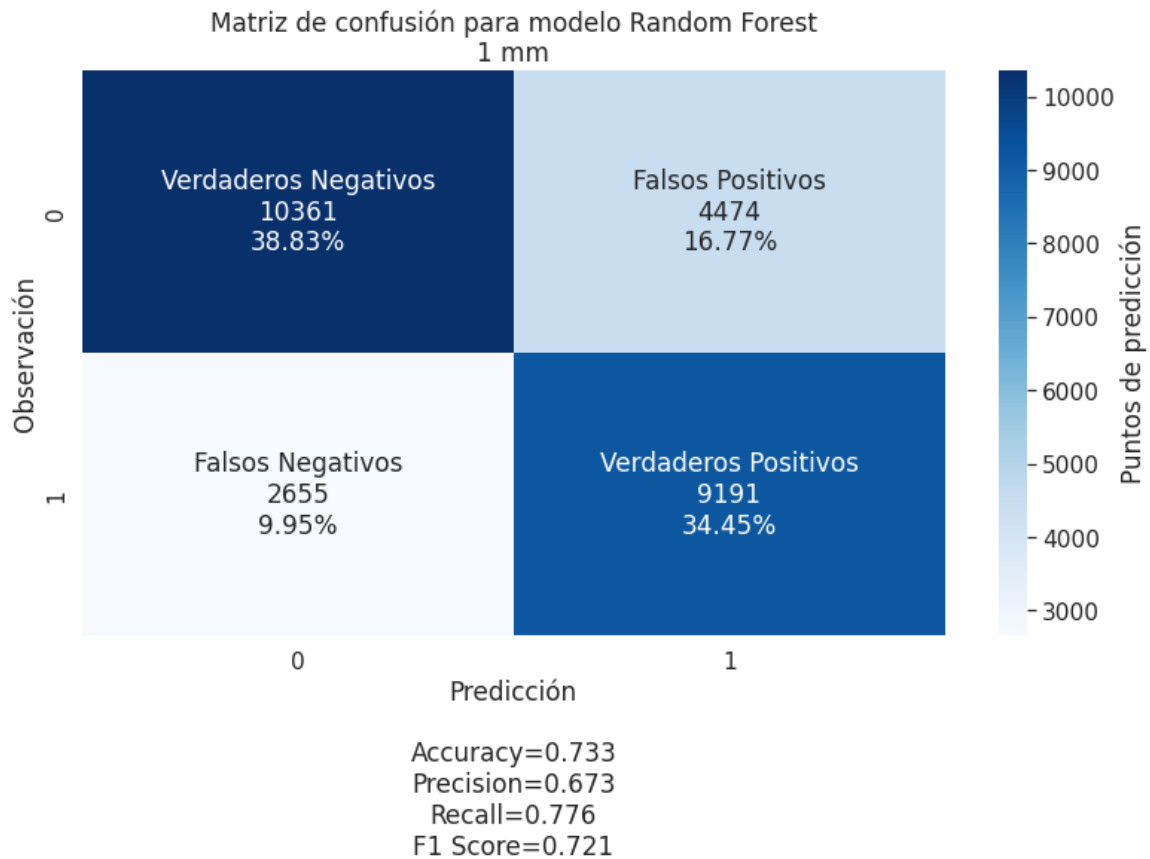
**Figura 29:** Matriz de confusión para modelo de pronóstico de precipitación con umbral de 0.1 mm

La Figura 29 muestra la matriz de confusión para el modelo de pronóstico con un umbral de 0.1 mm. En esta matriz, se observa lo siguiente:

- Los verdaderos positivos representan el 61.88% de los casos. Esto significa que el modelo acierta al predecir eventos de precipitación por encima de 0.1 mm en la mayoría de las ocasiones.
- Los falsos positivos son solo el 4.62% de los casos. Este indicador es positivo, ya que muestra que el modelo rara vez predice eventos de precipitación por encima de 0.1 mm cuando en realidad no ocurren. En otras palabras, las falsas alarmas son poco frecuentes.
- Los verdaderos negativos representan el 16.85% de las observaciones totales. Estos son casos en los que el modelo predice correctamente que no ocurrirá precipitación por encima de 0.1 mm.
- Los falsos negativos son el 16.66% de los casos. Esto indica que el modelo tiene dificultades para diferenciar eventos de precipitación por debajo de 0.1 mm, lo que

resulta en una cantidad significativa de falsos negativos. Esta limitación se debe en parte a que la zona de estudio tiene una alta proporción de casos con precipitación por encima de 0.1 mm, lo que hace que sea más difícil para el modelo identificar eventos más ligeros.

En resumen, el modelo de pronóstico con un umbral de 0.1 mm muestra una capacidad predictiva bastante buena para identificar eventos de precipitación por encima de este umbral, con un bajo porcentaje de falsas alarmas. Sin embargo, tiene dificultades para diferenciar eventos de precipitación más ligeros, lo que resulta en un número significativo de falsos negativos en estas situaciones.

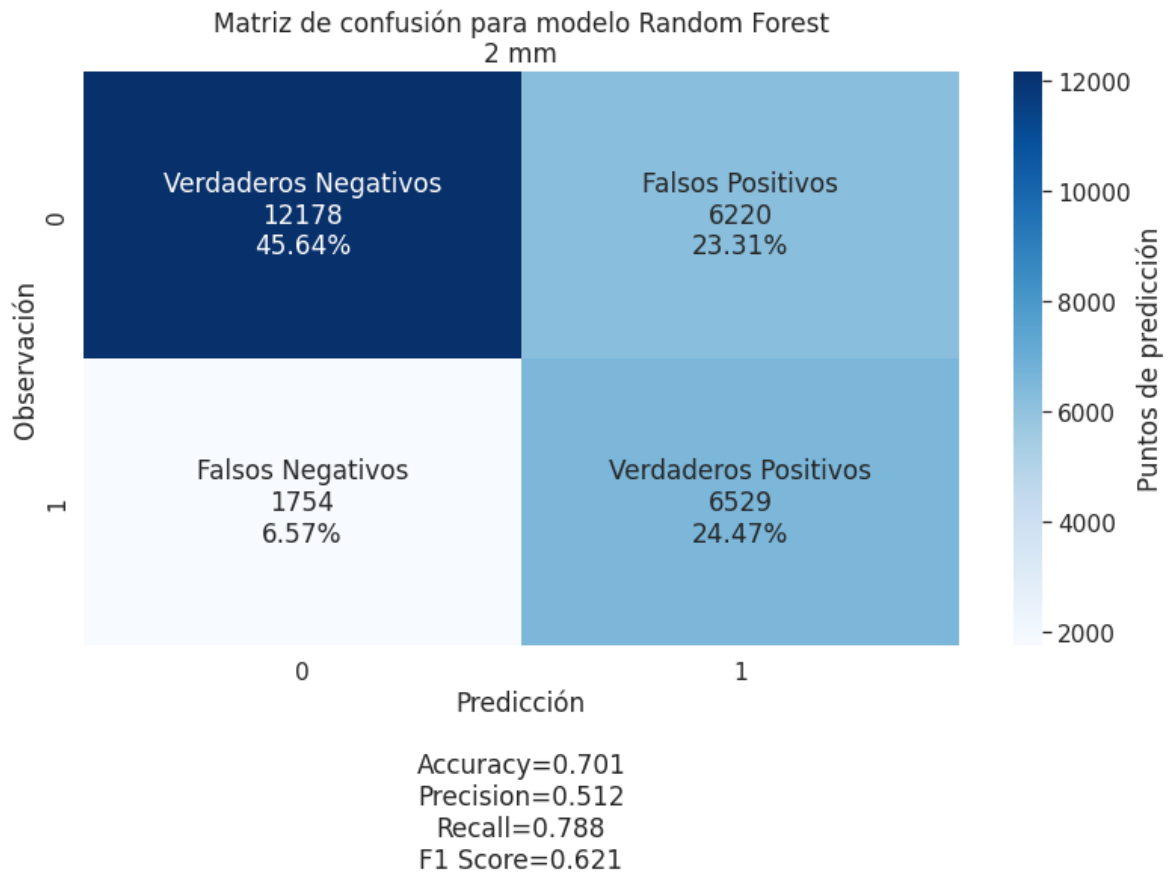


**Figura 30:** Matriz de confusión para modelo de pronóstico de precipitación con umbral de 1 mm

En la Figura 30, se muestra la matriz de confusión para el modelo de pronóstico de precipitación con un umbral de 1 mm. Esta matriz ofrece información valiosa sobre el rendimiento del modelo en la predicción de eventos de precipitación por encima de este umbral:

- Verdaderos Positivos: Representan el 34.45% de los casos. Estos son casos en los que el modelo predice correctamente eventos de precipitación por encima de 1 mm.
- Falsos Positivos: Constituyen el 16.77% de los casos. Estos casos son falsas alarmas, en los que el modelo predice precipitación por encima de 1 mm cuando en realidad no ocurre.
- Verdaderos Negativos: Representan el 38.83% de las observaciones totales. Estos son casos en los que el modelo predice correctamente que no habrá precipitación por encima de 1 mm.
- Falsos Negativos: Constituyen aproximadamente el 9.95% de los casos. Estos son eventos de precipitación por encima de 1 mm que el modelo no logra predecir.

En general, este modelo de pronóstico con un umbral de 1 mm muestra un buen rendimiento, con un porcentaje de acierto total de alrededor del 73%. Se destaca su capacidad para predecir eventos por debajo del umbral de precipitación (verdaderos negativos) y también para detectar eventos por encima del umbral (verdaderos positivos). Aunque se generan algunas falsas alarmas (falsos positivos), el modelo demuestra ser efectivo en la mayoría de los casos. Esto se debe en parte a que la distribución de eventos positivos y negativos está más equilibrada en este umbral, lo que facilita el entrenamiento del modelo Random Forests.



**Figura 31:** Matriz de confusión para modelo de pronóstico de precipitación con umbral de 2 mm

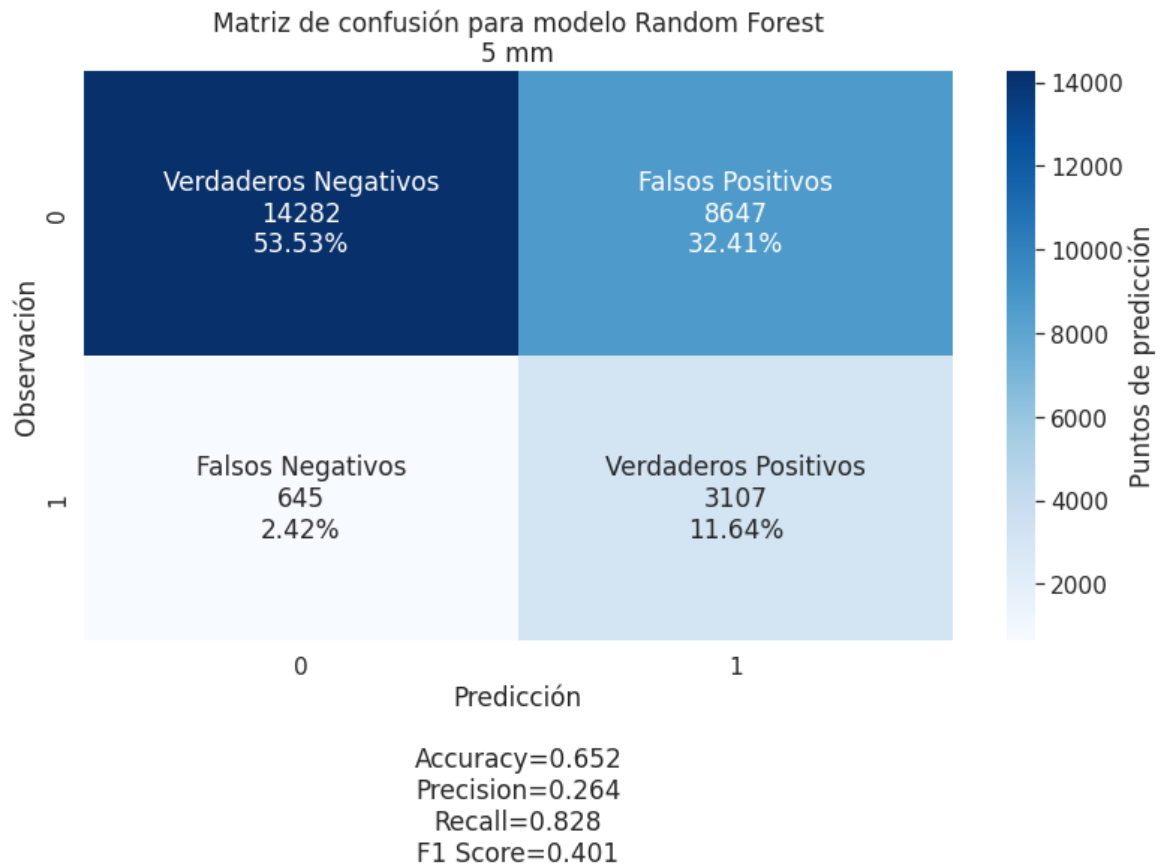
La Figura 31 muestra la matriz de confusión para el modelo de pronóstico de precipitación con un umbral de 2 mm. Esta matriz proporciona información importante sobre cómo el modelo se desempeña en la predicción de eventos de precipitación por encima de este umbral:

- Verdaderos Positivos: Representan el 24.47% de los casos. Estos son eventos en los que el modelo predice correctamente la ocurrencia de precipitación por encima de 2 mm.
- Falsos Positivos: Constituyen aproximadamente el 23.31% de los casos. Estos son casos en los que el modelo predice la precipitación por encima del umbral de 2 mm cuando en realidad no ocurre.
- Verdaderos Negativos: Representan el 45.64% de las observaciones totales. Estos son casos en los que el modelo predice correctamente que no habrá precipitación por encima de 2 mm.



- Falsos Negativos: Constituyen un porcentaje bajo, alrededor del 6.57% de los casos. Estos son eventos de precipitación por encima de 2 mm que el modelo no logra predecir.

Este modelo de pronóstico con umbral de 2 mm muestra un buen rendimiento en la identificación de eventos por debajo del umbral (verdaderos negativos), con un porcentaje alto de aciertos en esta categoría. Sin embargo, su capacidad para identificar eventos por encima del umbral (verdaderos positivos) es limitada, y genera un número significativo de falsas alarmas (falsos positivos). La limitación principal parece estar relacionada con la cantidad de eventos positivos en los datos de entrenamiento, que es considerablemente menor que los eventos negativos, lo que resulta en un mayor grado de imprecisión en las predicciones positivas.



**Figura 32:** Matriz de confusión para modelo de pronóstico de precipitación con umbral de 5 mm

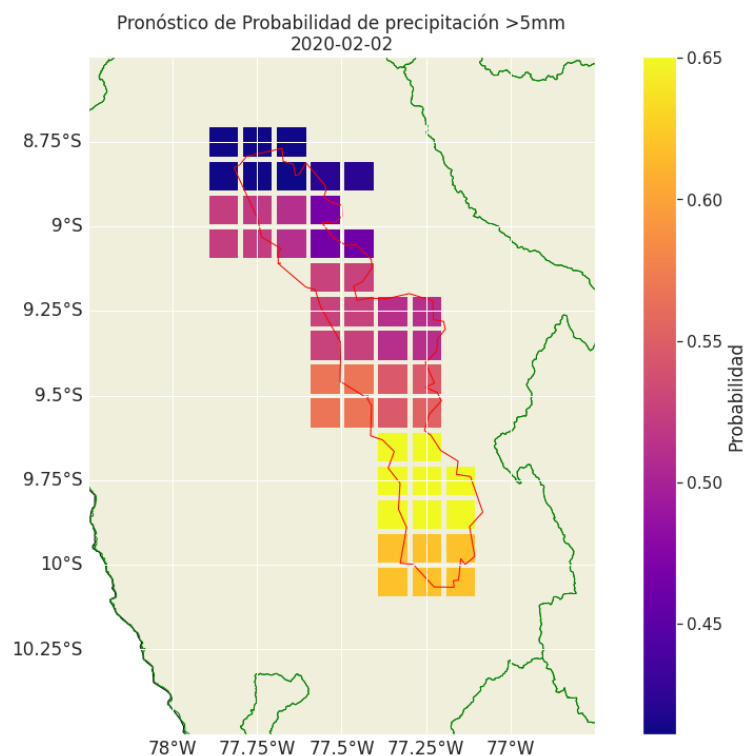
La Figura 32 muestra la matriz de confusión para el modelo de pronóstico de precipitación con un umbral de 5 mm. Esta matriz ofrece una visión detallada del desempeño del modelo en la predicción de eventos de precipitación por encima de este umbral:

- Verdaderos Positivos: Representan el 11.64% de los casos. Estos son eventos en los que el modelo realiza una predicción acertada al anticipar la ocurrencia de precipitación por encima de 5 mm.
- Falsos Positivos: Constituyen aproximadamente el 32.41% de los casos. Estos son casos en los que el modelo predice la precipitación por encima del umbral de 5 mm cuando en realidad no ocurre.
- Verdaderos Negativos: Representan el 53.53% de las observaciones totales. Estos son casos en los que el modelo predice correctamente que no habrá precipitación por encima de 5 mm.
- Falsos Negativos: Constituyen un porcentaje bajo, alrededor del 2.42% de los casos. Estos son eventos de precipitación por encima de 5 mm que el modelo no logra predecir.

El modelo de pronóstico con umbral de 5 mm muestra un bajo rendimiento en la identificación de eventos de precipitación por encima de este umbral (verdaderos positivos), con un porcentaje bastante reducido de aciertos en esta categoría. También genera un número considerable de falsas alarmas (falsos positivos), lo que indica que el modelo tiende a sobrestimar la ocurrencia de eventos de alta precipitación. Por otro lado, el modelo presenta un buen rendimiento en la identificación de eventos sin precipitación significativa (verdaderos negativos), con un alto porcentaje de aciertos en esta categoría. Los falsos negativos, que representan eventos de alta precipitación que no se pronostican, son limitados en número.

Este rendimiento sugiere que el modelo de pronóstico con umbral de 5 mm es más conservador en la predicción de eventos de alta precipitación y tiende a ser cauteloso al realizar predicciones positivas. Esta cautela puede deberse a la falta de suficientes eventos de alta precipitación en los datos de entrenamiento para mejorar la precisión en esta categoría. La baja precisión de este modelo de pronóstico se refleja en un F1-score de 0.401 siendo, este el modelo de predicción con los resultados menos alentadores.

Las predicciones de los modelos de pronóstico se realizan en puntos de grilla en términos de probabilidad como se muestra en la Figura 33.



**Figura 33:** Mapa de probabilidad de precipitación para el modelo de pronóstico con umbral de 5 mm/día

## V. CONCLUSIONES

- La selección de los predictores desempeñó un papel fundamental en la precisión de los modelos de pronóstico. Para las temperaturas mínimas, se destacaron la temperatura y el flujo zonal a nivel de 500 hPa, así como la temperatura y la humedad en superficie (expresada como humedad específica y relativa). Estos predictores se identificaron como los más importantes para los modelos de predicción de temperaturas mínimas. En el caso de las temperaturas máximas, la variable meteorológica más relevante resultó ser la temperatura en superficie, seguida de las variables de humedad en la superficie. Esto sugiere que las condiciones en la capa más cercana a la superficie, a 2 metros sobre el suelo, tienen un impacto significativo en las temperaturas máximas, en comparación con las variables a niveles medios y altos de la atmósfera. El uso de componentes principales contribuyó a captar información meteorológica de toda la zona de estudio, lo cual ayudó significativamente a la precisión en las predicciones.
- Se concluye que la aplicación de PCA sobre la matriz de predicción que agrupa las distintas variables meteorológicas que rodean a cada punto de pronóstico, fue un paso vital en la construcción de modelos de probabilidad de precipitación. El desarrollo de los modelos Random Forests para probabilidad de precipitación presentaron diferencias considerables entre sí tanto en los parámetros de entrenamiento como en los resultados obtenidos, debido a los distintos umbrales de precipitación acumulada diaria que se emplearon.
- Los resultados de validación de los modelos de predicción de temperaturas superan con creces las predicciones directas del modelo global en la zona de estudio, así como las predicciones basadas en la persistencia, lo que demuestra su capacidad para proporcionar pronósticos más precisos y útiles. La capacidad de pronóstico de las temperaturas mínimas fue considerablemente superior en comparación con las temperaturas máximas. Esto sugiere que el modelo es más efectivo en la predicción

de eventos de temperatura mínima. A pesar de los resultados positivos, se identificaron limitaciones en la capacidad de los modelos para identificar la magnitud de ciertos eventos climáticos extremos. Estos eventos extremos pueden estar relacionados con factores locales que no son completamente capturados por el modelo global. Con respecto a la validación de los modelos de probabilidad de precipitación, los resultados variaron según el umbral de precipitación considerado. Los modelos de 0.1 mm/día y 1 mm/día demostraron buena eficiencia en las predicciones, mientras que el modelo de 5 mm/día generó un elevado número de falsas alarmas. El modelo de 2 mm/día fue eficiente para eventos por debajo del umbral, pero no logró predecir de manera precisa eventos por encima de este. La variabilidad en los resultados de los modelos de precipitación puede atribuirse en parte a la cantidad desbalanceada de eventos de lluvia en los datos de entrenamiento. Esto dificulta la capacidad del modelo Random Forests para diferenciar adecuadamente las clases con menos eventos.

- Como conclusión general del presente trabajo, los modelos de predicción desarrollados muestran un buen rendimiento general en la predicción de temperaturas mínimas y máximas, superando las predicciones directas y las basadas en la persistencia. Sin embargo, existen limitaciones en la identificación de eventos extremos y en la predicción de precipitación, especialmente para umbrales elevados, que podrían abordarse mejorando la calidad y cantidad de datos de entrenamiento y considerando factores locales en futuras investigaciones.

## VI. RECOMENDACIONES

A partir de los resultados obtenidos en la presente investigación, se hacen las siguientes recomendaciones:

- Replicar la metodología en otras zonas del Perú, empleando una mayor cantidad de datos observados y extendiendo el horizonte de pronóstico, teniendo especial consideración en la selección de las variables meteorológicas de modelo numérico que se utilizarán para la generación de estos modelos de pronóstico además de la posibilidad de emplear técnicas más modernas de machine learning que permitan superar los resultados obtenidos en la presente investigación.
- Debido a la amplia disponibilidad y fácil acceso a los datos del modelo numérico GFS 0.25°, así como el bajo costo computacional que requieren los modelos de pronóstico desarrollados en el presente estudio, se recomienda emplearlas como herramientas de apoyo para el pronóstico operativo evaluando constantemente su capacidad predictiva, así como las limitaciones que se presenten.
- Reentrenar los modelos de pronóstico de probabilidad de precipitación estableciendo diferentes umbrales de precipitación acumulada diaria obtenidos de estaciones meteorológicas en la zona de estudio, en lugar del entrenamiento realizado con información grillada de precipitación acumulada diaria con el fin de contrastar resultados y obtener información más fiable en el punto de pronóstico calibrado.
- Entrenar modelos de pronóstico bajo la misma metodología, empleando información del GFS 0.5° o de otros modelos numéricos, sean estos modelos globales o regionales. Se debe tener en cuenta que el dominio utilizado debe ser similar para que los resultados de estos distintos modelos empíricos de pronóstico desarrollados sean comparables entre sí y de esta forma elegir los que mejor representen las variables pronosticadas en cada zona de estudio trabajada.
- Considerar el uso de otros estadísticos, como sesgo o Error Absoluto Medio, en futuros análisis y evaluaciones del desempeño de los modelos de pronóstico. Estos estadísticos son cruciales en el pronóstico operativo, ya que ofrecen una comprensión más completa de la exactitud de los pronósticos y permiten ajustes o mejoras específicas.

## VII. BIBLIOGRAFÍA

- Acuña, D., Ávalos, G., Cubas, F., Llacza, A., Jácome, G., & Barreto, C. (2013). *Evaluación de los modelos CMIP5 del IPCC en el Perú: Proyecciones al año 2030 en la Región Ancash*.  
[https://repositorio.senamhi.gob.pe/bitstream/handle/20.500.12542/215/Evaluacion-modelos-CMIP5-IPCC-Peru-proyecciones-a%C3%B1o-2030-region-Ancash\\_2013.pdf?sequence=1&isAllowed=y](https://repositorio.senamhi.gob.pe/bitstream/handle/20.500.12542/215/Evaluacion-modelos-CMIP5-IPCC-Peru-proyecciones-a%C3%B1o-2030-region-Ancash_2013.pdf?sequence=1&isAllowed=y)
- Aliaga-Nestares, V., De La Cruz, G., & Takahashi, K. (2023). Comparison between the Operational and Statistical Daily Maximum and Minimum Temperature Forecasts on the Central Coast of Peru. *Weather and Forecasting*, 38(4), 555–570.  
<https://doi.org/10.1175/waf-d-21-0094.1>
- Al-Matarneh, L., Sheta, A., Bani-Ahmad, S., Alshaer, J., & Al-Oqily, I. (2014). Development of temperature-based weather forecasting models using neural networks and fuzzy logic. *International Journal of Multimedia and Ubiquitous Engineering*, 9(12), 343–366. <https://doi.org/10.14257/ijmue.2014.9.12.31>
- Alvarez, M. S., Vera, C. S., Kiladis, G. N., & Liebmann, B. (2016). Influence of the Madden Julian Oscillation on precipitation and surface air temperature in South America. *Climate Dynamics*, 46(1–2), 245–262. <https://doi.org/10.1007/s00382-015-2581-6>
- Anthes, R. A., Kuo, Y. -H, Hsie, E. -Y, Low-Nam, S., & Bettge, T. W. (1989). Estimation of skill and uncertainty in regional numerical models. *Quarterly Journal of the Royal Meteorological Society*, 115(488), 763–806. <https://doi.org/10.1002/qj.49711548803>
- Aybar, C., Fernández, C., Huerta, A., Lavado, W., Vega, F., & Felipe-Obando, O. (2020). Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day. *Hydrological Sciences Journal*, 65(5), 770–785.  
<https://doi.org/10.1080/02626667.2019.1649411>

- Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. En *Nature* (Vol. 525, Número 7567, pp. 47–55). Nature Publishing Group. <https://doi.org/10.1038/nature14956>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Anal. Methods*, 6(9), 2812–2831. <https://doi.org/10.1039/C3AY41907J>
- Calvo Sánchez, F. J. (2018). Modelos atmosféricos. En *Física del caos en la predicción meteorológica* (pp. 129–143). Agencia Estatal de Meteorología. <https://doi.org/10.31978/014-18-009-x.10>
- Casimiro, W. L., Labat, D., Ronchail, J., Espinoza, J. C., & Guyot, J. L. (2012). Trends in rainfall and temperature in the Peruvian Amazon-Andes basin over the last 40 years (1965-2007). *Hydrological Processes*, n/a-n/a. <https://doi.org/10.1002/hyp.9418>
- Castro, A., Davila, C., Laura, W., Cubas, F., Avalos, G., López Ocaña, C., Villena, D., Valdez, M., Urbiola, J., Trebejo, I., Menis, L., & Marín, D. (2021). *CLIMAS DEL PERÚ - Mapa de Clasificación Climática Nacional*.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chatfield, C. (1975). *The Analysis of Time Series: Theory and Practice*. Springer US. <https://doi.org/10.1007/978-1-4899-2925-9>
- Collins, S. N., James, R. S., Ray, P., Chen, K., Lassman, A., & Brownlee, J. (2013). Grids in Numerical Weather and Climate Models. *Climate Change and Regional/Local Responses*, 256. <https://doi.org/10.5772/55922>
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random Forests. En *Ensemble Machine Learning* (pp. 157–175). Springer US. [https://doi.org/10.1007/978-1-4419-9326-7\\_5](https://doi.org/10.1007/978-1-4419-9326-7_5)
- Dai, A., Trenberth, K. E., & Karl, T. R. (1999). Effects of Clouds, Soil Moisture, Precipitation, and Water Vapor on Diurnal Temperature Range. *Journal of Climate*, 12(8), 2451–2473. [https://doi.org/10.1175/1520-0442\(1999\)012<2451:EOCSMP>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2451:EOCSMP>2.0.CO;2)



- Dalianis, H. (2018). Evaluation Metrics and Evaluation. En *Clinical Text Mining* (pp. 45–53). Springer International Publishing. [https://doi.org/10.1007/978-3-319-78503-5\\_6](https://doi.org/10.1007/978-3-319-78503-5_6)
- De la Cruz, G. A. (2016). *Pronóstico cuantitativo de precipitación en términos de probabilidades* [Universidad Nacional Agraria La Molina]. <https://hdl.handle.net/20.500.12996/3126>
- Deduy, I. (2019). *Regresión sobre componentes principales*. Universidad de Sevilla.
- Deng, K. (1998). *OMEGA: ON-LINE MEMORY-BASED GENERAL PURPOSE SYSTEM CLASSIFIER*. Carnegie Mellon's School of Computer Science.
- Durai, V. R., & Bhradwaj, R. (2014). Evaluation of statistical bias correction methods for numerical weather prediction model forecasts of maximum and minimum temperatures. *Natural Hazards*, 73(3), 1229–1254. <https://doi.org/10.1007/s11069-014-1136-1>
- Eccel, E., Ghielmi, L., Granitto, P., Barbiero, R., Grazzini, F., & Cesari, D. (2007). Prediction of minimum temperatures in an alpine region by linear and non-linear post-processing of meteorological models. *Nonlinear Processes in Geophysics*, 14(3), 211–222. <https://doi.org/10.5194/npg-14-211-2007>
- Feng, W., Sui, H., Tu, J., Huang, W., & Sun, K. (2018). A novel change detection approach based on visual saliency and random forest from multi-temporal high-resolution remote-sensing images. *International Journal of Remote Sensing*, 39(22), 7998–8021. <https://doi.org/10.1080/01431161.2018.1479794>
- Gajowniczek, K., Ząbkowski, T., & Szupiluk, R. (2014). ESTIMATING THE ROC CURVE AND ITS SIGNIFICANCE FOR CLASSIFICATION MODELS' ASSESSMENT. *QUANTITATIVE METHODS IN ECONOMICS*, XV, 382–391.
- Garreaud, R. D., & Aceituno, P. (2007). Atmospheric Circulation and Climatic Variability. En *The Physical Geography of South America*. Oxford University Press. <https://doi.org/10.1093/oso/9780195313413.003.0010>
- Ghielmi, L., & Eccel, E. (2006). Descriptive models and artificial neural networks for spring frost prediction in an agricultural mountain area. *Computers and Electronics in Agriculture*, 54(2), 101–114. <https://doi.org/10.1016/j.compag.2006.09.001>

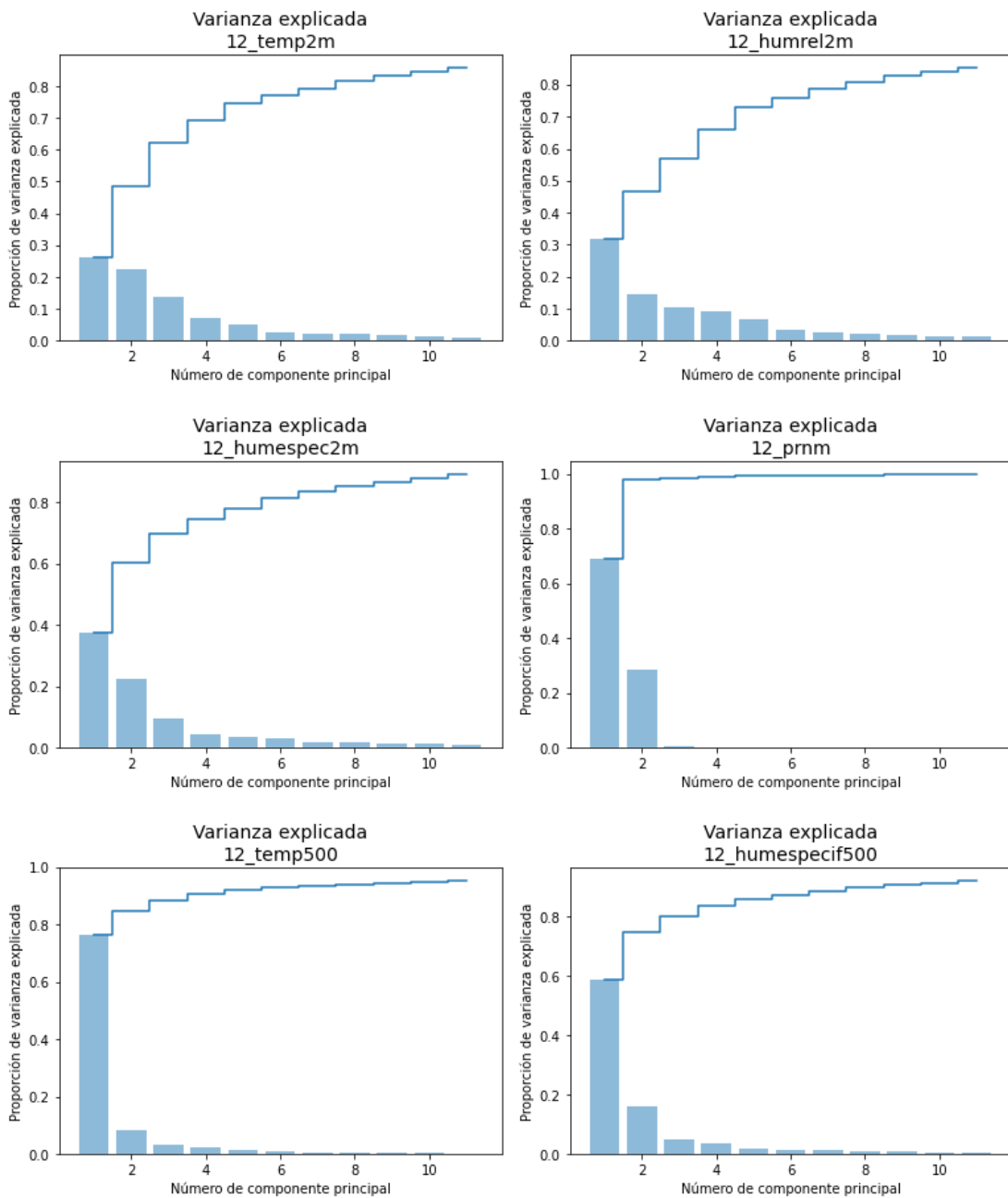
- Glahn, H. R., & Lowry, D. A. (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology*, *11*(8), 1203–1211. [https://doi.org/10.1175/1520-0450\(1972\)011<1203:TUOMOS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2)
- Golding, B. W. (2015). Numerical Models: Regional Prediction Models. En *Encyclopedia of Atmospheric Sciences: Second Edition* (Second Edi, Vol. 4). Elsevier. <https://doi.org/10.1016/B978-0-12-382225-3.00253-X>
- Herman, G. R., & Schumacher, R. S. (2018). Money Doesn't Grow on Trees, but Forecasts Do: Forecasting Extreme Precipitation with Random Forests. *Monthly Weather Review*, *146*(5), 1571–1600. <https://doi.org/10.1175/MWR-D-17-0250.1>
- Ho, T. K. (1995). Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, 278–282.
- Imfeld, N., Sedlmeier, K., Gubler, S., Marrou, K. C., Davila, C. P., Huerta, A., Lavado-Casimiro, W., Rohrer, M., Scherrer, S. C., & Schwierz, C. (2021). A combined view on precipitation and temperature climatology and trends in the southern Andes of Peru. *International Journal of Climatology*, *41*(1), 679–698. <https://doi.org/10.1002/joc.6645>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer New York. <https://doi.org/10.1007/978-1-4614-7138-7>
- Jha, S. K., Shrestha, D. L., Stadnyk, T. A., & Coulibaly, P. (2018). Evaluation of ensemble precipitation forecasts generated through post-processing in a Canadian catchment. *Hydrology and Earth System Sciences*, *22*(3), 1957–1969. <https://doi.org/10.5194/hess-22-1957-2018>
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6a ed.). Pearson.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. En *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Vol. 374, Número 2065). Royal Society of London. <https://doi.org/10.1098/rsta.2015.0202>
- Jolliffe, I. T. (1982). A Note on the Use of Principal Components in Regression. *Applied Statistics*, *31*(3), 300. <https://doi.org/10.2307/2348005>

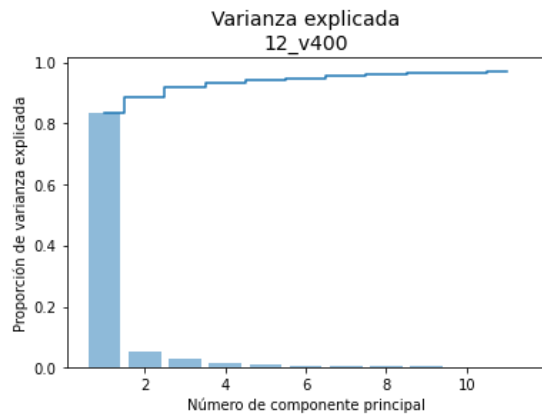
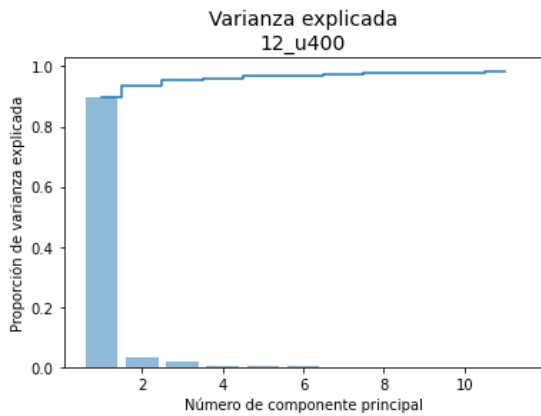
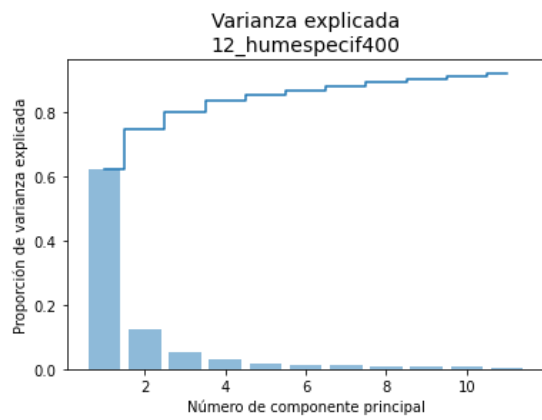
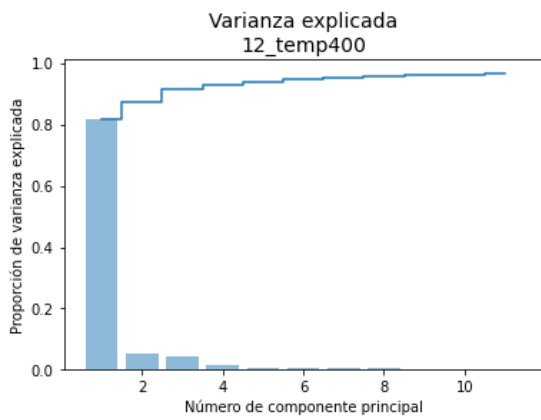
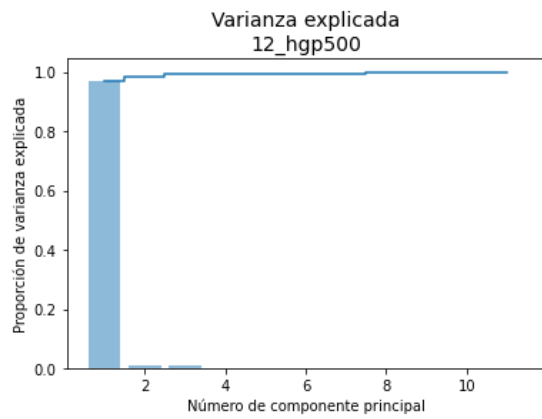
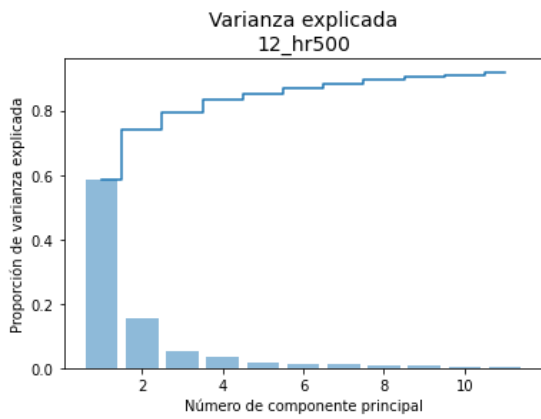
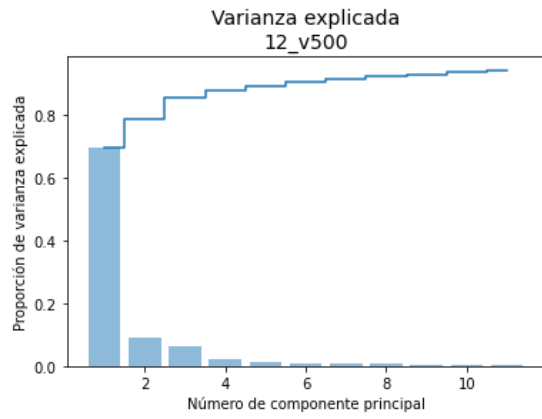
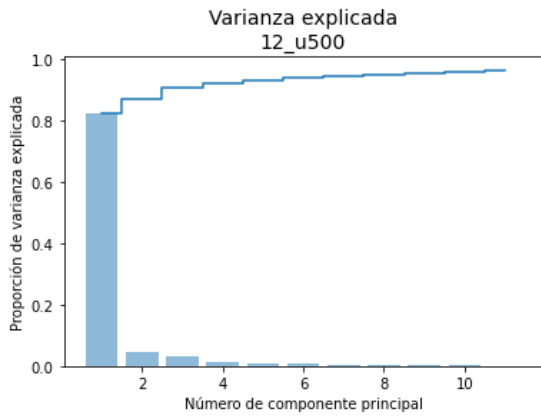
- Kimura, R. (2002). Numerical weather prediction. *Journal of Wind Engineering and Industrial Aerodynamics*, 90(12–15), 1403–1414. [https://doi.org/10.1016/S0167-6105\(02\)00261-1](https://doi.org/10.1016/S0167-6105(02)00261-1)
- Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning*, 30(2/3), 271–274. <https://doi.org/10.1023/A:1017181826899>
- Martínez Pérez, J. A., & Pérez Martín, P. S. (2023). La curva ROC. *Medicina de Familia. SEMERGEN*, 49(1), 101821. <https://doi.org/10.1016/j.semerg.2022.101821>
- Marzban, C., Sandgathe, S., & Kalnay, E. (2006). MOS, perfect prog, and reanalysis. *Monthly Weather Review*, 134(2), 657–663. <https://doi.org/10.1175/MWR3088.1>
- Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal: the journal of Medical Association of Malawi*, 24(3), 69–71.
- Navarro, O. (2009). *Selección de Variables en Regresión Componentes Principales*.
- NCEP. (2015). *NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive*. UCAR/NCAR - Research Data Archive. <https://doi.org/10.5065/D65D8PWK>
- NOAA. (s/f). *Filters: High Pass*. Recuperado el 16 de abril de 2023, de [https://training.weather.gov/pds/climate/pcu2/statistics/Stats/part2/Filter\\_HP.htm](https://training.weather.gov/pds/climate/pcu2/statistics/Stats/part2/Filter_HP.htm)
- Qona'ah, N., Sutikno, Ferawati, K., & Nirwana, M. B. (2020). Temperature Forecast Using Ridge Regression as Model Output Statistics. *Proceeding International Conference on Science and Engineering*, 3(April), 383–388. <https://doi.org/10.14421/icse.v3.533>
- Rokach, L., & Maimon, O. (2006). Decision Trees. En *Data Mining and Knowledge Discovery Handbook* (pp. 165–192). Springer-Verlag. [https://doi.org/10.1007/0-387-25465-X\\_9](https://doi.org/10.1007/0-387-25465-X_9)
- Ruddiman, W. (2013). *Earth's Climate: Past and Future* (3a ed.).
- Saavedra, M., & Takahashi, K. (2017). Physical controls on frost events in the central Andes of Peru using in situ observations and energy flux models. *Agricultural and Forest Meteorology*, 239, 58–70. <https://doi.org/10.1016/j.agrformet.2017.02.019>
- SENAMHI. (2020). *CLIMAS DEL PERÚ – Mapa de Clasificación Climática Nacional*. <https://www.senamhi.gob.pe/load/file/01404SENA-4.pdf>

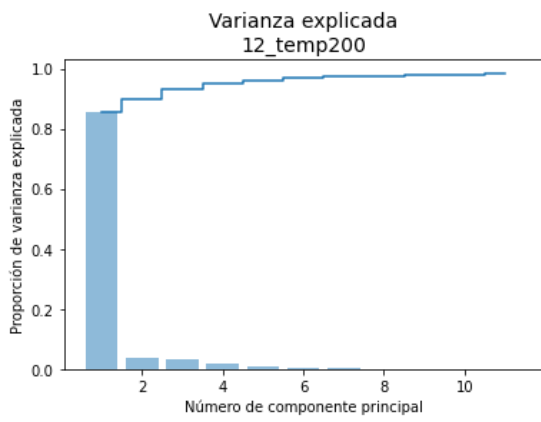
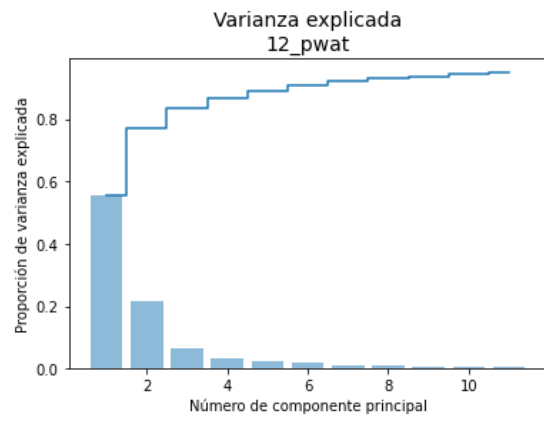
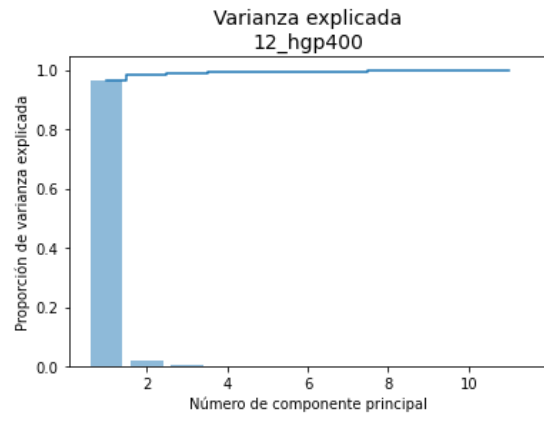
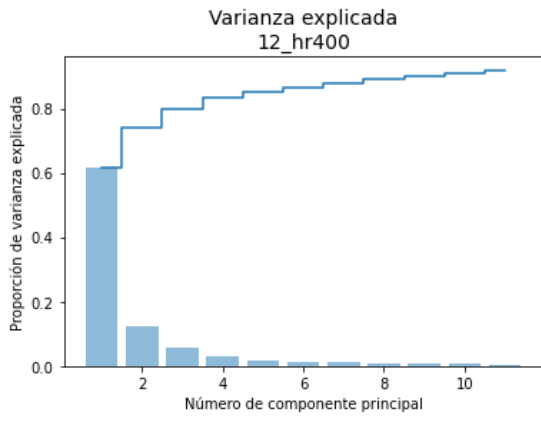
- SERNANP. (2011). *Parque Nacional Huascarán plan maestro 2010 - 2015*.
- SERNANP. (2019, enero 1). *Parque Nacional Huascarán*.  
<https://www.gob.pe/institucion/sernanp/informes-publicaciones/1949215-parque-nacional-huascarán>
- Sevink, J. (2009). *La Cordillera Blanca: un paisaje explicado* (1a ed.).
- Sokolova, M., Japkowicz, N., & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. En A. Sattar & B. Kang (Eds.), *AI 2006: Advances in Artificial Intelligence* (pp. 1015–1021). Springer Berlin Heidelberg.
- Sulca, J., Vuille, M., Roundy, P., Takahashi, K., Espinoza, J., Silva, Y., Trasmonte, G., & Zubieta, R. (2018). Climatology of extreme cold events in the central Peruvian Andes during austral summer: origin, types and teleconnections. *Quarterly Journal of the Royal Meteorological Society*, *144*(717), 2693–2714. <https://doi.org/10.1002/qj.3398>
- Tan, P.-N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Pearson.
- Valderá-Figueroa, N., Ballester-Pérez, M., & Sánchez-Pérez, Á. L. (2018). La persistencia como referencia en la estimación de la habilidad de las predicciones del Tiempo a corto plazo. *Revista Cubana De Meteorología*, *24*(1), 29–43. <http://rcm.insmet.cu/index.php/rcm/article/view/258>
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Academic Press., Ed.; 3rd ed.).
- Woodcock, F. (1984). Australian Experimental Model Output Statistics Forecasts of Daily Maximum and Minimum Temperature. *Monthly Weather Review*, *112*(10), 2112–2121. [https://doi.org/10.1175/1520-0493\(1984\)112<2112:AEMOSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<2112:AEMOSF>2.0.CO;2)
- Yang, F., Pan, H. L., Krueger, S. K., Moorthi, S., & Lord, S. J. (2006). Evaluation of the NCEP global forecast system at the ARM SGP site. *Monthly Weather Review*, *134*(12), 3668–3690. <https://doi.org/10.1175/MWR3264.1>

## VIII. ANEXOS

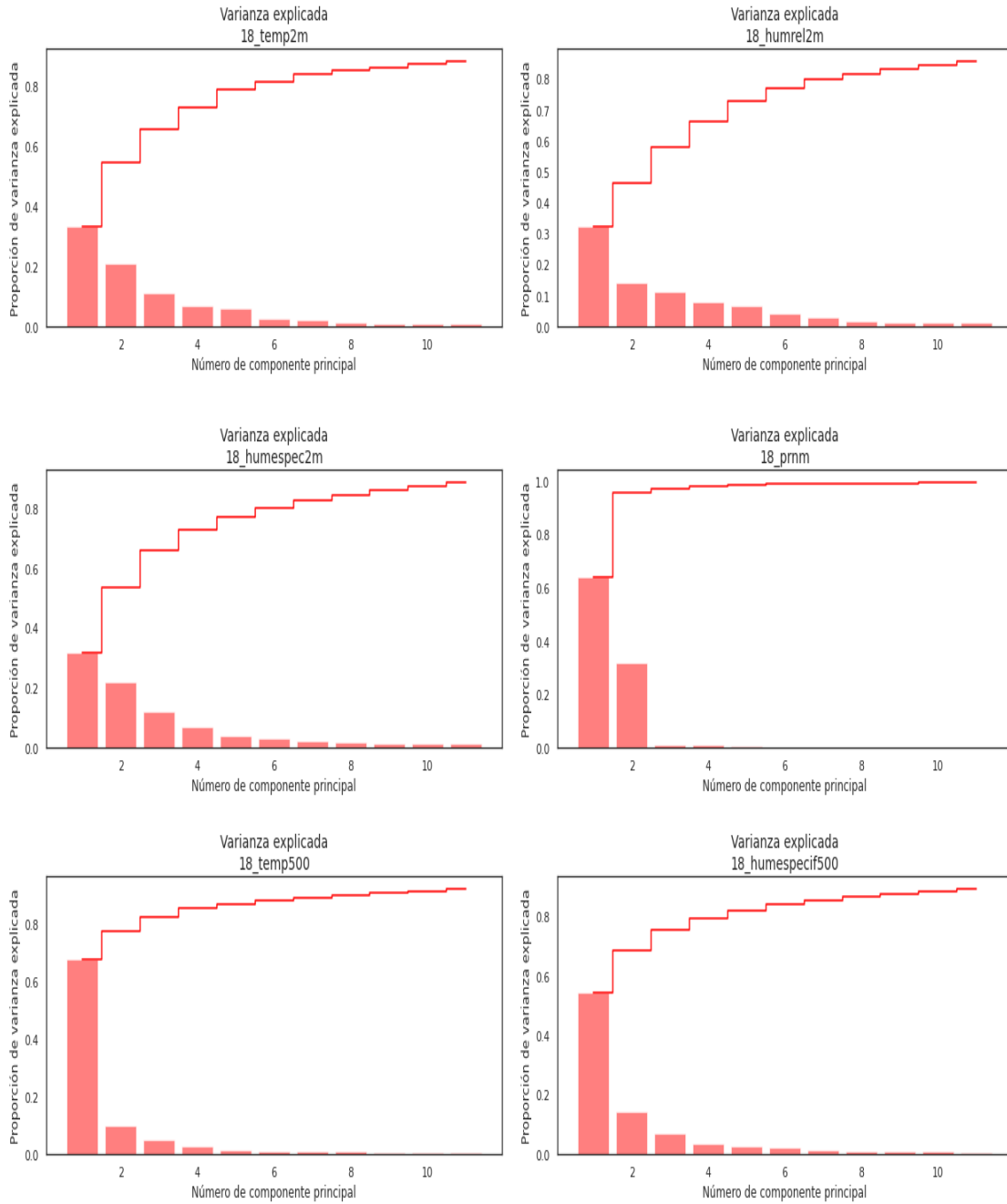
**Anexo 1:** Gráficas de varianza explicada para variables meteorológicas filtradas del modelo numérico GFS empleadas para el modelo de pronóstico de temperatura mínima



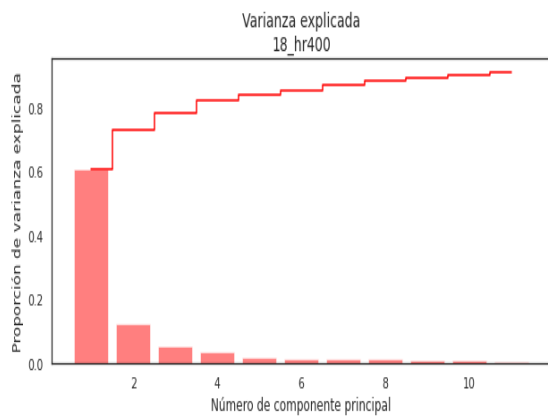
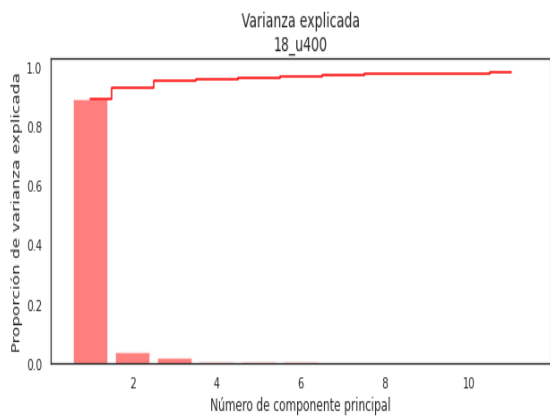
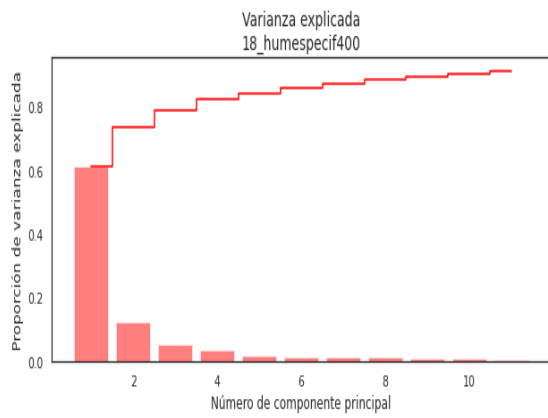
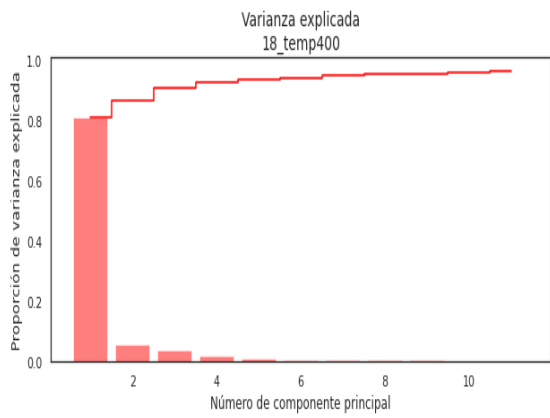
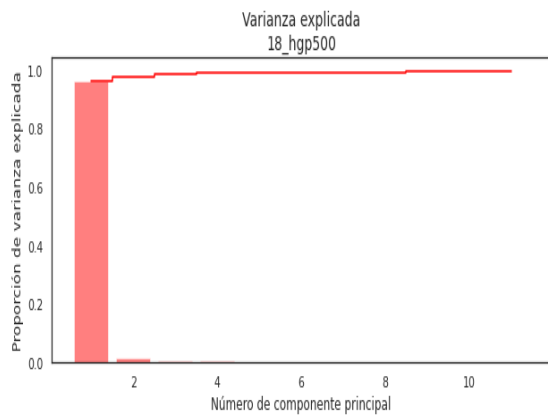
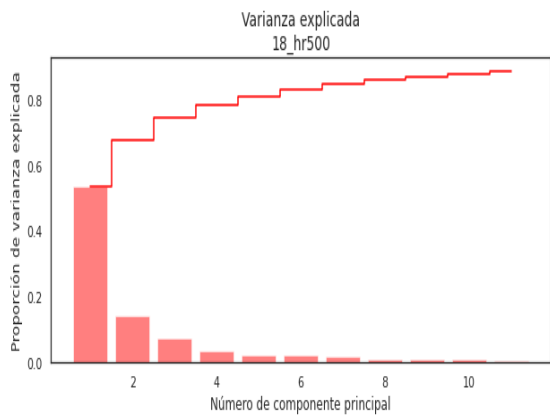
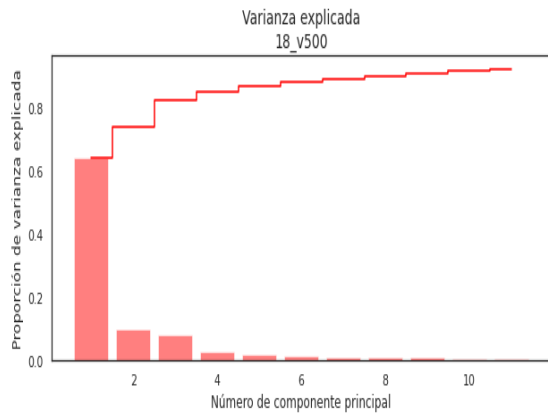
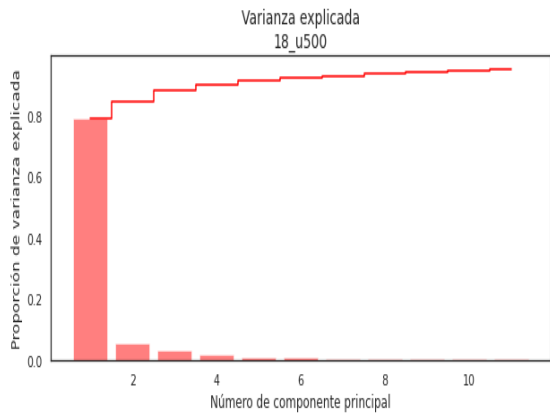


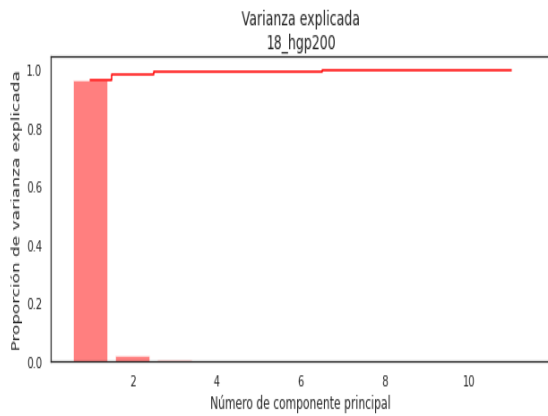
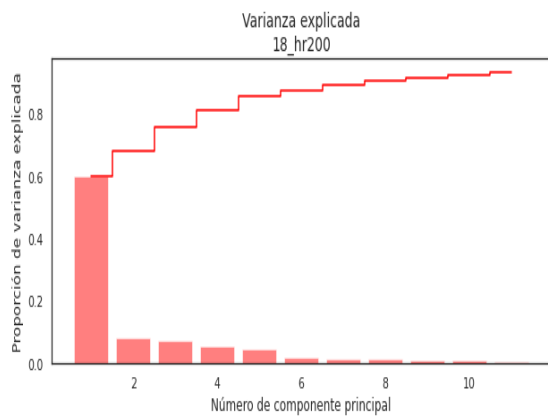
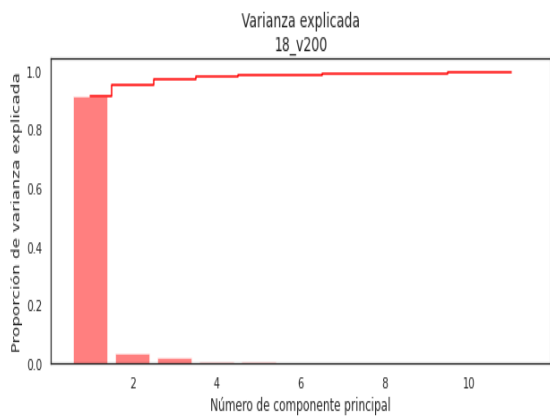
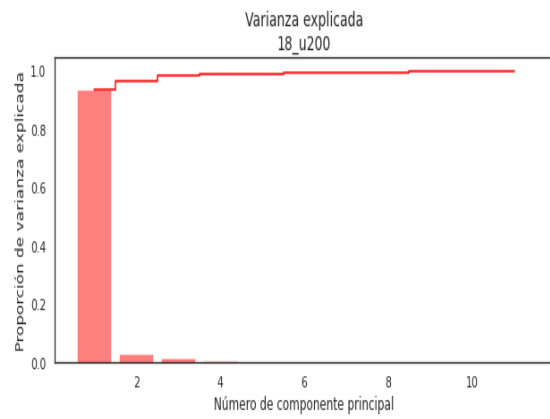
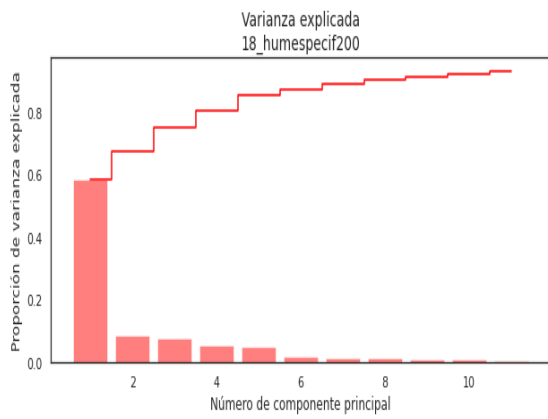
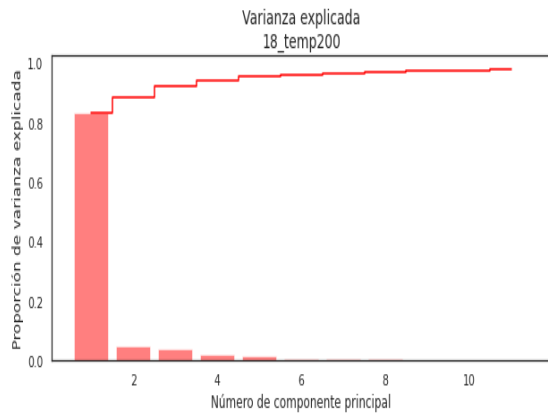
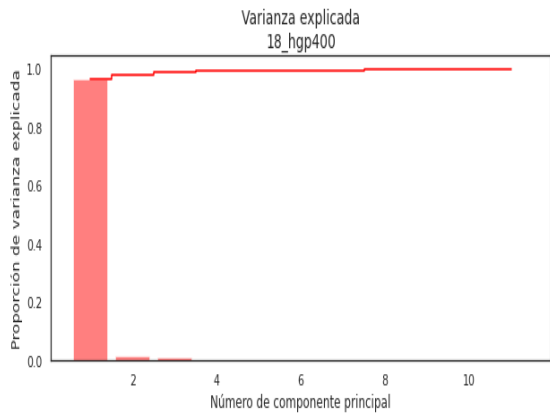


**Anexo 2:** Gráficas de varianza explicada para variables meteorológicas filtradas del modelo numérico GFS empleadas para el modelo de pronóstico de temperatura máxima









### Anexo 3: Mapa de pronóstico de temperaturas extremas diarias en la región Áncash

