

Predicción de heladas meteorológicas mediante modelos ensemble de aprendizaje automático con un enfoque multi-horizonte para la alerta agroclimática en el Altiplano de Puno, 2015-2024

Meteorological Frost Prediction Using Ensemble Machine Learning Models With a Multi-Horizon Approach for Agroclimatic Early Warning in the Puno Altiplano, 2015–2024

 Romel Gonzalo Quispe-Valero^{1*} y  Jorge Luis Flores-Quispe²

Resumen

Se evaluó la capacidad predictiva de nueve algoritmos de aprendizaje automático para predecir heladas meteorológicas en seis estaciones de la región de Puno (Perú), a lo largo de tres horizontes temporales diferentes (24, 48 y 72 h). En el segundo paso, los datos de reanálisis climático (ERA5-Land mediante la API de Open-Meteo) del período 2015–2024 se consolidaron en 526 032 registros horarios correspondientes a 21 918 observaciones diarias de hasta 57 predictores agroclimáticos. Las heladas presentaron una prevalencia media regional del 17,4 %, mostrando una marcada heterogeneidad espacial que osciló desde el 31,3 % en Ilave hasta el 7,7 % en Puno, evidenciando el papel regulador de la temperatura del lago Titicaca. Se realizó una validación cruzada estratificada en cinco pliegues para estimar el AUC-ROC, la F1-Score, la precisión, la sensibilidad y el kappa de Cohen del rendimiento del modelo. El mejor desempeño global se obtuvo mediante el ensemble por votación, con un AUC-ROC de 0,957 y una F1-Score de 0,744 a las 24 h, lo cual correspondió a valores de 0,913 y 0,656 en un tiempo de corte de 72 h, respectivamente. La optimización del umbral condujo a una mejora en la F1-Score regional de +0,025 frente al umbral convencional ($t = 0,50$). El análisis de valores SHAP reveló que el impulsor con mayor impacto fue la temperatura mínima de promedio móvil de 3 días (media |SHAP| = 1,274). El modelo desarrollado representa una herramienta objetiva de alerta agroclimática temprana para cultivos en las tierras altas andinas.

Palabras clave: aprendizaje automático, predicción meteorológica, helada, Perú, base de datos.

Abstract

The predictive capacity of nine machine learning algorithms was evaluated for forecasting meteorological frosts at six weather stations in the Puno region (Peru) across three different time horizons (24, 48, and 72 h). In the second step, climate reanalysis data (ERA5-Land using the Open-Meteo API) from the period 2015–2024 were consolidated into 526,032 hourly records corresponding to 21,918 daily observations from up to 57 agroclimatic predictors. Frosts showed a regional average prevalence of 17.4%, exhibiting marked spatial heterogeneity ranging from 31.3% in Ilave to 7.7% in Puno, highlighting the regulatory role of Lake Titicaca's temperature. A five-fold stratified cross-validation was performed to estimate the AUC-ROC, F1-score, accuracy, sensitivity, and Cohen's kappa for model performance. The best overall performance was obtained using the voting ensemble, with an AUC-ROC of 0.957 and an F1-score of 0.744 at 24 h, corresponding to values of 0.913 and 0.656 at a 72 h cutoff time, respectively. Threshold optimization led to a regional F1-score improvement of +0.025 compared to the conventional threshold ($t = 0.50$). SHAP value analysis revealed that the driver with the greatest impact was the 3-day moving average minimum temperature (mean |SHAP| = 1.274). The developed model represents an objective early agroclimatic warning tool for crops in the Andean highlands.

Keywords: machine learning, weather forecasting, frost, Peru, database.

Recibido: 25/03/2026

Aceptado: 20/08/2026

Publicado: 28/08/2026

Sección: Artículo Original

***Autor correspondiente:** logonza64@gmail.com

Introducción

Las heladas meteorológicas son los principales factores de riesgo agroclimático en el altiplano peruano-boliviano, que anualmente afectan a grandes áreas de cultivos estratégicos y específicamente a la quinua (*Chenopodium quinoa* Willd.). Además, su presencia se ha detectado en cultivos de tubérculos, como las papas (*Solanum tuberosum* L.) y las habas (*Vicia faba* L.) (Benique-Olivera & Ojeda-Tito, 2024; Quispe-Mamani et al., 2023). El presente estudio se enfoca en la región de Puno, con altitudes de 3800 a 4500 msnm, y cuyo régimen climático está determinado por el lago Titicaca y las corrientes andinas (Arias et al., 2021), donde las temperaturas mínimas pueden descender a -15°C entre

los meses de mayo a agosto, con impactos económicos estimados de millones de soles/año (MINAGRI, 2021; SENAMHI, 2023).

¹Departamento de Estadística e Informática, Facultad de Ingeniería Estadística e Informática, Universidad Nacional del Altiplano, Puno, Perú. ORCID: [0009-0007-7960-3275](https://orcid.org/0009-0007-7960-3275)

²Universidad Nacional del Altiplano, Puno, Perú. ORCID: [0000-0002-3796-4167](https://orcid.org/0000-0002-3796-4167)

Como citar: Quispe-Valero, R. G. & Flores-Quispe, J. L. (2026). Predicción de heladas meteorológicas mediante modelos ensemble de aprendizaje automático con un enfoque multi-horizonte para la alerta agroclimática en el Altiplano de Puno, 2015-2024. *Revista de Investigaciones Altoandinas - Journal of High Andean Research*, 28, e28849. <https://doi.org/10.18271/ria.2026.849>



Attribution 4.0 Internacional (CC BY 4.0) Share-Adapt

Los enfoques de pronóstico de heladas basados en NWP se ven desafiados por las limitaciones en la resolución espacial y temporal a escala local de los Andes altos (Junquas et al., 2022; Schultz et al., 2021). Los modelos de aprendizaje automático, incluidos los sistemas de pronóstico meteorológico global de última generación (Lam et al., 2023; Bi et al., 2023; Nguyen et al., 2023), han demostrado tener ventajas para capturar dinámicas no lineales complejas a partir de datos observacionales y mejorar la predicción de eventos extremos en los últimos años (Rasp et al., 2024; Dewitte et al., 2021).

Sin embargo, la mayoría de estos estudios se han realizado en latitudes medias (Europa, Norteamérica), y solo unos pocos cubren la región altoandina de Sudamérica, que presenta características climáticas distintivas que combinan alta irradiación solar, baja presión atmosférica y el efecto termorregulador del lago Titicaca (Arias et al., 2021). Esta brecha proporciona la justificación para desarrollar aquí modelos predictivos calibrados regionalmente.

El objetivo del presente estudio fue comparar el rendimiento de predicción de nueve algoritmos de aprendizaje automático, centrándose en métodos de aumento por ensamble, para eventos de heladas previstos en seis estaciones de la región de Puno en horizontes de 24, 48 y 72 horas, utilizando datos de reanálisis ERA5-Land de acceso abierto del periodo 2015–2024.

Materiales y métodos

Área de estudio y estaciones

El estudio comprendió seis estaciones meteorológicas representativas de la región Puno: Puno (3827 m, ribera del lago Titicaca), Juliaca (3824 m), Ayaviri (3923 m), Ilave (3866 m, zona continental), Lampa (3892 m) y Santa Lucía (3986 m). La selección cubrió gradientes altitudinales, de continentalidad y de exposición al cuerpo de agua lacustre.

Fuente de datos

Se utilizaron los datos observacionales climáticos disponibles, incluyendo ERA5-Land (Muñoz-Sabater et al., 2021), a escala horaria del modelo de aplicación agrometeorológica para el intervalo de tiempo enero de 2015 – diciembre de 2024, proporcionado mediante la API de Open-Meteo (Zippenfenig, 2023) de forma gratuita y sin restricciones de uso académico. Las variables descargadas fueron la temperatura del aire a 2 m (máx., mín., media), temperatura de punto de rocío (fórmula del rocío), precipitación, velocidad y dirección del viento a 10 m, presión superficial, nubosidad, humedad relativa a 2 m, radiación solar de onda corta y temperatura del suelo a 0 cm. Esto dio como resultado un conjunto de datos total de 526 032 registros horarios (60 archivos csv: por estación-año).

Preprocesamiento y construcción de características

Los datos horarios se agregaron a una escala diaria (media, mínimo, máximo y desviación estándar). Se definió el horizonte de predicción (H24h, H48h o H72h), y la variable objetivo (helada futura) como la aparición de una temperatura mínima ≤ 0 °C en la zona analizada. En total, se han construido 57 predictores: (a) estadísticas diarias de todas las variables meteorológicas; (b) métricas de medias móviles de 3 y 7 días de temperatura y precipitación; (c) características de estacionalidad, incluidas el mes y el día del año codificados en formas seno/coseno para variables cíclicas; (d) términos de interacción en la forma [viento $\times$ nubosidad]. Para los valores faltantes (menos del 5 % de temperatura del suelo), los rellenamos con la mediana por columna.

El problema de clases desbalanceadas en la clasificación 0-1, con un porcentaje medio de la clase positiva del 17,4 % (todas las heladas), se trató siguiendo a Kaur et al. (2022) mediante ponderación de clases (`class_weight='balanced'`) para los modelos aplicables y mediante `scale_pos_weight` para XGBoost. El procedimiento se implementó en Python con la librería `scikit-learn v1.4.0` (Grisel et al., 2024).

Los horizontes de 24, 48 y 72 horas se seleccionaron en función de la ventana temporal efectiva para la mitigación táctica agrometeorológica (activación de sistemas de aspersión, acondicionamiento de cubiertas térmicas, quemas controladas) en cultivos altoandinos durante el periodo en que las variables locales de superficie conservan una importante capacidad predictiva discriminante.

Modelos de aprendizaje automático

Se evaluaron nueve clasificadores binarios: regresión logística (línea base), Random Forest, AdaBoost, Gradient Boosting, XGBoost, LightGBM y CatBoost (los tres últimos basados en gradient boosting; Bentéjac et al., 2021; Rozante et al., 2023), además de un VotingClassifier (votación suave con RF, XGB, LGB y CatB) y un StackingClassifier (metaclasificador: regresión logística con imputaciones).

Evaluación y validación

La evaluación se realizó mediante validación cruzada estratificada de cinco pliegues ($k=5$) con una semilla aleatoria fija (`random_state=42`), reportando la media $\pm$ la desviación estándar de: AUC-ROC, F1-Score, Precisión, Sensibilidad, Exactitud y el Kappa de Cohen. Se llevó a cabo una optimización adicional del umbral de clasificación, entre el F1-Score y el Índice de Youden, en un conjunto de validación separado (15% de los datos), validación espacial Leave-One-Station-Out (LOSO) con LightGBM para cuantificar qué tan variable es el rendimiento en el espacio, así como un análisis de la calibración de probabilidades con Platt Scaling y Regresión Isotónica (métrica Brier Score).

Interpretabilidad: Valores SHAP

Para analizar el mejor modelo, se calcularon valores SHAP (SHapley Additive exPlanations) (que es un método post hoc basado en la teoría de juegos cooperativos; Chen et al., 2023) para una selección aleatoria de 1000 observaciones del conjunto de entrenamiento para el horizonte H24h. Se utilizó un diagrama de enjambre (beeswarm), un diagrama de dependencia para la variable más importante y el ranking de importancia global según la media |SHAP|

Resultados

Análisis descriptivo del conjunto de datos

Un resumen de las estadísticas descriptivas para las seis estaciones se muestra en la tabla 1. La prevalencia de heladas fue altamente heterogénea espacialmente, oscilando desde el 7,7 % en Puno (influencia del lago) hasta el 31,3 % en Ilave (zona continental alejada del lago). Esta variación espacial de cuatro veces en la prevalencia crea contextos de desbalance de clases completamente diferentes entre estaciones.

Tabla 1

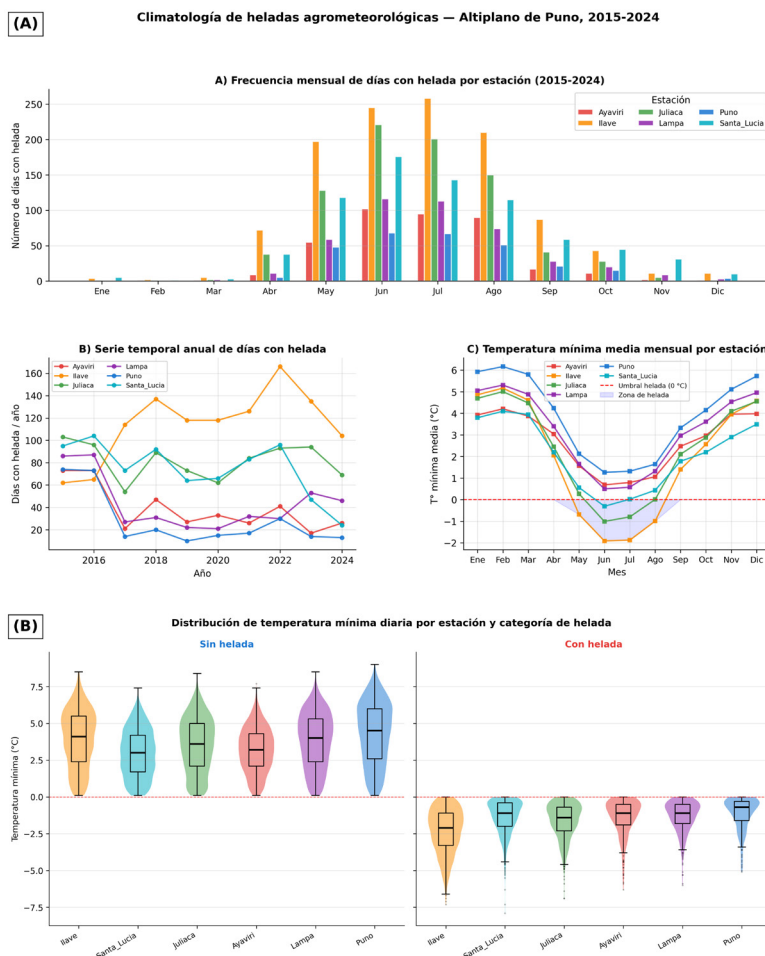
Resumen de variables meteorológicas y prevalencia de heladas por estación (2015-2024)

Estación	N (días)	Heladas (%)	T° mín (°C)	T° media (°C)	HR (%)	Viento (km/h)	Precip. (mm/d)
Ayaviri	3653	384 (10,5 %)	$2,7 \pm 2,0$	$8,3 \pm 1,4$	$58,6 \pm 18,7$	$6,9 \pm 2,1$	$0,13 \pm 0,19$
Ilave	3653	1145 (31,3 %)	$2,0 \pm 3,4$	$8,1 \pm 1,8$	$62,9 \pm 16,0$	$10,1 \pm 2,2$	$0,12 \pm 0,19$
Juliaca	3653	817 (22,4 %)	$2,4 \pm 2,7$	$8,7 \pm 1,6$	$59,1 \pm 18,1$	$8,4 \pm 2,6$	$0,11 \pm 0,17$
Lampa	3653	435 (11,9 %)	$3,2 \pm 2,4$	$9,1 \pm 1,5$	$55,5 \pm 19,8$	$6,4 \pm 1,6$	$0,11 \pm 0,17$
Puno	3653	280 (7,7 %)	$3,9 \pm 2,5$	$9,3 \pm 1,5$	$58,9 \pm 18,1$	$9,1 \pm 2,0$	$0,13 \pm 0,19$
Santa Lucía	3653	744 (20,4 %)	$2,1 \pm 2,3$	$7,5 \pm 1,6$	$59,5 \pm 19,6$	$7,4 \pm 2,3$	$0,12 \pm 0,18$

Nota: Los valores corresponden a Media $\pm$ Desviación Estándar de datos diarios. Heladas (%) indica el porcentaje de días con $T_{\min} \leq 0^\circ\text{C}$.

Figura 1

Climatología estacional y distribución térmica de heladas en la región Puno: (A) Frecuencia mensual y ciclo anual de heladas; (B) Distribución empírica de la temperatura mínima diaria según condición de helada por estación meteorológica



Comparación de modelos por horizonte de predicción

Las métricas de desempeño de los nueve modelos evaluados en los horizontes de 24, 48 y 72 horas se muestran en la Tabla 2 y una comparación de muestra se ilustra en la Figura 2. En los tres horizontes, el

método de Votación mostró el mayor AUC-ROC medio: 0,957 (H24h), 0,926 (H48h) y 0,913 (H72h). Para el F1-Score, Votación y LightGBM finalizaron con un F1 = 0,744 a 24 h.

Tabla 2

Métricas de evaluación predictiva (media $\pm$ desv. est.) de los 9 modelos evaluados bajo validación cruzada estratificada ($k=5$).

Modelo	Horizonte	Exactitud	Precisión	Sensibilidad	F1-Score	AUC-ROC	Kappa
Voting	H24h	0,905 $\pm$ 0,004	0,698 $\pm$ 0,012	0,797 $\pm$ 0,007	0,744 $\pm$ 0,009	0,957 $\pm$ 0,001	0,687 $\pm$ 0,010
LightGBM	H24h	0,906 $\pm$ 0,004	0,706 $\pm$ 0,012	0,787 $\pm$ 0,004	0,744 $\pm$ 0,008	0,955 $\pm$ 0,001	0,687 $\pm$ 0,010
XGBoost	H24h	0,904 $\pm$ 0,004	0,695 $\pm$ 0,014	0,792 $\pm$ 0,011	0,740 $\pm$ 0,010	0,956 $\pm$ 0,001	0,683 $\pm$ 0,015
GradientBoosting	H24h	0,911 $\pm$ 0,005	0,779 $\pm$ 0,020	0,685 $\pm$ 0,010	0,729 $\pm$ 0,012	0,953 $\pm$ 0,001	0,676 $\pm$ 0,015
Stacking	H24h	0,891 $\pm$ 0,003	0,643 $\pm$ 0,010	0,841 $\pm$ 0,008	0,729 $\pm$ 0,006	0,953 $\pm$ 0,002	0,663 $\pm$ 0,008
CatBoost	H24h	0,881 $\pm$ 0,001	0,608 $\pm$ 0,003	0,884 $\pm$ 0,009	0,720 $\pm$ 0,003	0,955 $\pm$ 0,001	0,647 $\pm$ 0,004
RandomForest	H24h	0,908 $\pm$ 0,003	0,781 $\pm$ 0,020	0,652 $\pm$ 0,006	0,711 $\pm$ 0,006	0,952 $\pm$ 0,001	0,656 $\pm$ 0,008
Logística	H24h	0,868 $\pm$ 0,004	0,577 $\pm$ 0,008	0,895 $\pm$ 0,006	0,702 $\pm$ 0,005	0,952 $\pm$ 0,002	0,621 $\pm$ 0,007
AdaBoost	H24h	0,898 $\pm$ 0,005	0,738 $\pm$ 0,019	0,641 $\pm$ 0,010	0,686 $\pm$ 0,013	0,938 $\pm$ 0,002	0,626 $\pm$ 0,016
XGBoost	H48h	0,876 $\pm$ 0,004	0,622 $\pm$ 0,011	0,735 $\pm$ 0,011	0,674 $\pm$ 0,011	0,925 $\pm$ 0,003	0,595 $\pm$ 0,011
Voting	H48h	0,876 $\pm$ 0,004	0,621 $\pm$ 0,011	0,734 $\pm$ 0,012	0,673 $\pm$ 0,011	0,926 $\pm$ 0,003	0,599 $\pm$ 0,014
LightGBM	H48h	0,876 $\pm$ 0,004	0,622 $\pm$ 0,010	0,732 $\pm$ 0,012	0,672 $\pm$ 0,009	0,924 $\pm$ 0,004	0,596 $\pm$ 0,011
Stacking	H48h	0,856 $\pm$ 0,003	0,561 $\pm$ 0,007	0,790 $\pm$ 0,011	0,656 $\pm$ 0,008	0,920 $\pm$ 0,005	0,569 $\pm$ 0,010
CatBoost	H48h	0,840 $\pm$ 0,005	0,525 $\pm$ 0,009	0,846 $\pm$ 0,009	0,648 $\pm$ 0,009	0,924 $\pm$ 0,004	0,552 $\pm$ 0,011
GradientBoosting	H48h	0,890 $\pm$ 0,003	0,727 $\pm$ 0,011	0,583 $\pm$ 0,014	0,647 $\pm$ 0,009	0,923 $\pm$ 0,004	0,583 $\pm$ 0,011
Logística	H48h	0,827 $\pm$ 0,005	0,502 $\pm$ 0,009	0,854 $\pm$ 0,007	0,632 $\pm$ 0,009	0,919 $\pm$ 0,005	0,529 $\pm$ 0,012
RandomForest	H48h	0,885 $\pm$ 0,004	0,726 $\pm$ 0,019	0,546 $\pm$ 0,018	0,623 $\pm$ 0,015	0,921 $\pm$ 0,004	0,557 $\pm$ 0,017
AdaBoost	H48h	0,879 $\pm$ 0,006	0,715 $\pm$ 0,027	0,501 $\pm$ 0,019	0,589 $\pm$ 0,022	0,907 $\pm$ 0,006	0,521 $\pm$ 0,025
Voting	H72h	0,870 $\pm$ 0,003	0,608 $\pm$ 0,009	0,712 $\pm$ 0,004	0,656 $\pm$ 0,006	0,913 $\pm$ 0,006	0,578 $\pm$ 0,011
XGBoost	H72h	0,869 $\pm$ 0,004	0,603 $\pm$ 0,009	0,714 $\pm$ 0,010	0,654 $\pm$ 0,009	0,911 $\pm$ 0,007	0,576 $\pm$ 0,010
LightGBM	H72h	0,868 $\pm$ 0,005	0,602 $\pm$ 0,014	0,707 $\pm$ 0,007	0,650 $\pm$ 0,010	0,911 $\pm$ 0,006	0,569 $\pm$ 0,013
Stacking	H72h	0,850 $\pm$ 0,005	0,548 $\pm$ 0,012	0,780 $\pm$ 0,005	0,644 $\pm$ 0,008	0,909 $\pm$ 0,005	0,552 $\pm$ 0,011
CatBoost	H72h	0,830 $\pm$ 0,005	0,508 $\pm$ 0,009	0,827 $\pm$ 0,005	0,629 $\pm$ 0,007	0,910 $\pm$ 0,006	0,527 $\pm$ 0,010
GradientBoosting	H72h	0,882 $\pm$ 0,004	0,710 $\pm$ 0,015	0,546 $\pm$ 0,018	0,617 $\pm$ 0,014	0,908 $\pm$ 0,005	0,549 $\pm$ 0,016
Logística	H72h	0,812 $\pm$ 0,006	0,476 $\pm$ 0,009	0,843 $\pm$ 0,006	0,609 $\pm$ 0,007	0,902 $\pm$ 0,005	0,497 $\pm$ 0,010
RandomForest	H72h	0,880 $\pm$ 0,003	0,714 $\pm$ 0,018	0,515 $\pm$ 0,014	0,598 $\pm$ 0,011	0,910 $\pm$ 0,005	0,530 $\pm$ 0,012
AdaBoost	H72h	0,871 $\pm$ 0,004	0,702 $\pm$ 0,019	0,452 $\pm$ 0,014	0,550 $\pm$ 0,016	0,892 $\pm$ 0,005	0,479 $\pm$ 0,018

Nota: Los valores reportados corresponden al promedio de 5 pliegues de validación cruzada.

Figura 2

Comparación del desempeño predictivo (AUC-ROC y F1-Score) para los 9 modelos evaluados a través de los tres horizontes de pronóstico (24, 48 y 72 h).

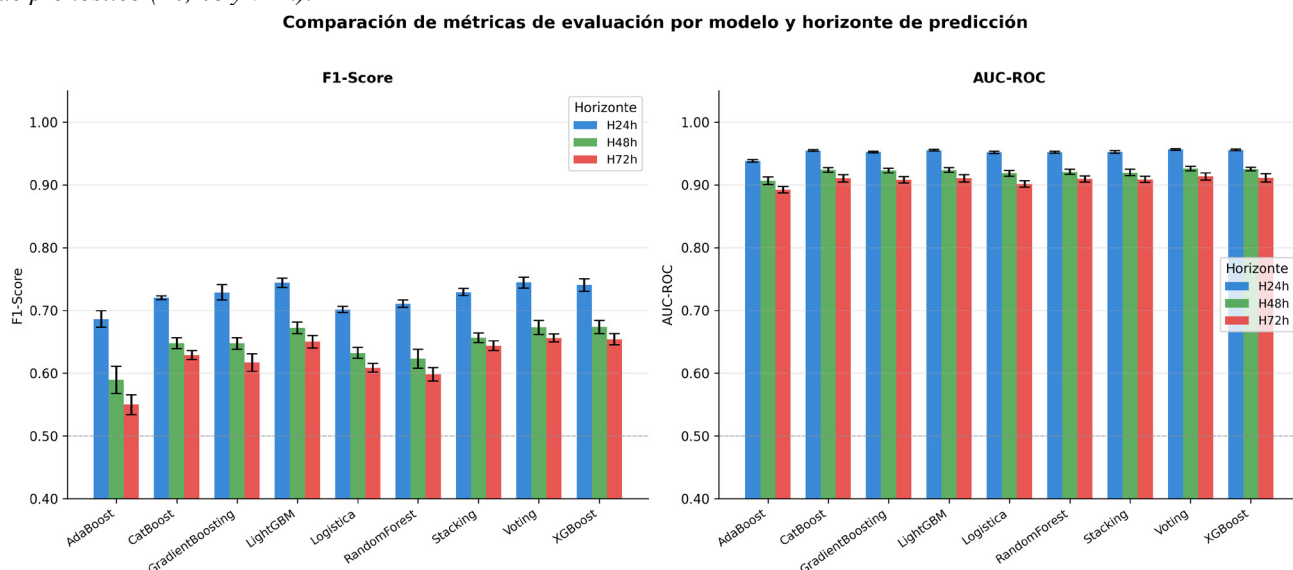


Figura 3

Curvas ROC para los clasificadores evaluados en el horizonte de predicción de 24 horas (H24h)

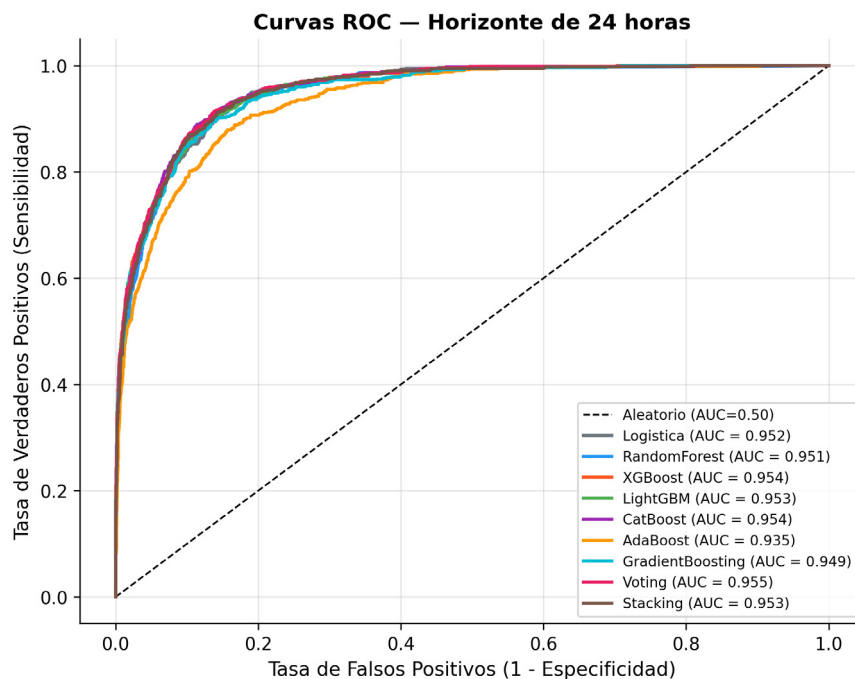


Figura 4

Ganancia en F1-Score alcanzada mediante el ajuste de umbral de decisión por modelo y horizonte

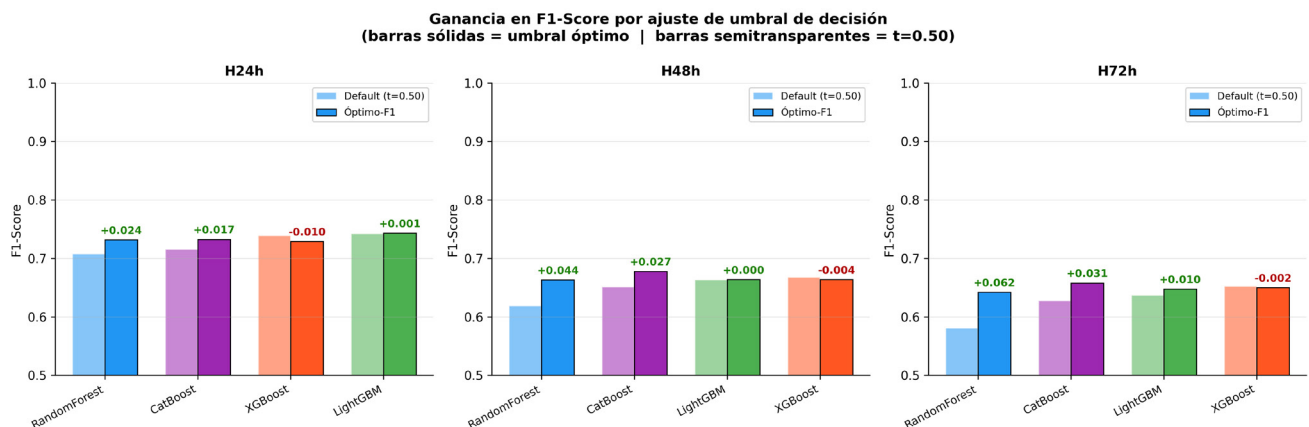
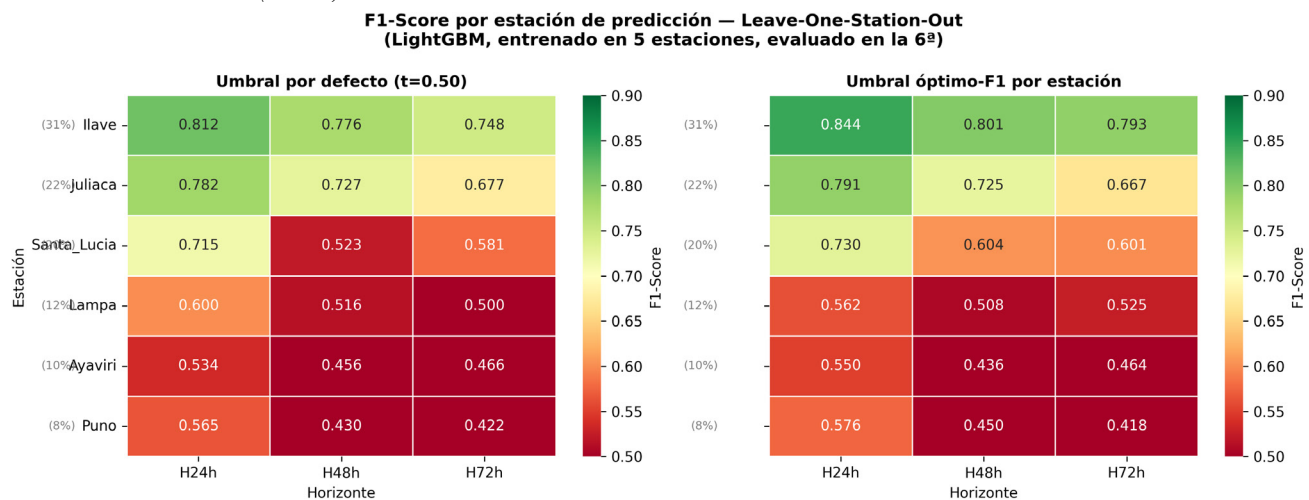


Figura 5

Mapa de calor y desempeño espacial del F1-Score por estación meteorológica evaluada independientemente bajo el enfoque Leave-One-Station-Out (LOSO)



Optimización del umbral de clasificación

La Figura 4 representa la ganancia en F1-Score al reemplazar el umbral por defecto ($t=0,50$) por el umbral óptimo-F1. Las ganancias medias fueron: +0,008 (H24h), +0,017 (H48h) y +0,025 (H72h).

Análisis espacial Leave-One-Station-Out

El análisis LOSO de la F1-Score (Figura 5) indicó que la prevalencia local de las heladas está fuertemente correlacionada con las predicciones. Con LightGBM, el umbral óptimo osciló entre $t^* = 0,18$ en Ilave (31,3 % de heladas) y $t^* = 0,82$ en Puno (7,7 % de heladas), alcanzando una $F1 = 0,844$ en H24h.

Interpretabilidad mediante valores SHAP

El análisis SHAP para H24h se muestra en la Tabla 3 y en las figuras 6 y 7. La variable globalmente más importante fue la temperatura mínima media móvil de 3 días (media $|SHAP| = 1,274$), seguida por la temperatura media diaria (0,819).

Calibración de probabilidades

La Figura 7 ilustra los diagramas de confiabilidad (Reliability Diagrams). El escalado de Platt redujo el Brier Score de 0,0822 a 0,0631 en CatBoost (-23,2 %).

Tabla 3
Ranking Top-15 de importancia global de predictores según el valor medio absoluto $|SHAP|$ (H24h).

Rank	Predictor	$ SHAP $ medio	Contribución (%)
1	T° mín. media móvil 3d (°C)	1,2743	14,16 %
2	Temperatura media (°C)	0,8186	9,10 %
3	Temperatura mín. (°C)	0,7275	8,08 %
4	Precipitación media (mm)	0,4419	4,91 %
5	winddirection_10m_mean	0,3561	3,96 %
6	T° media móvil 3d (°C)	0,3364	3,74 %
7	Presión superficial media (hPa)	0,2808	3,12 %
8	Punto de rocío medio (°C)	0,2659	2,96 %
9	dewpoint_2m_max	0,2544	2,83 %
10	Velocidad viento media (km/h)	0,2370	2,63 %
11	Punto de rocío mín. (°C)	0,2312	2,57 %
12	surface_pressure_max	0,2272	2,52 %
13	Viento × Nubosidad	0,2144	2,38 %
14	surface_pressure_min	0,1844	2,05 %
15	winddirection_10m_min	0,1721	1,91 %

Nota: Calculado sobre una muestra representativa del modelo Voting Ensemble.

Figura 6
Interpretabilidad del modelo Voting mediante análisis SHAP (H24h): (A) Diagrama beeswarm de atribución de características para los 20 predictores principales; (B) Gráfico de dependencia SHAP para la temperatura mínima promedio móvil de 3 días (°C)

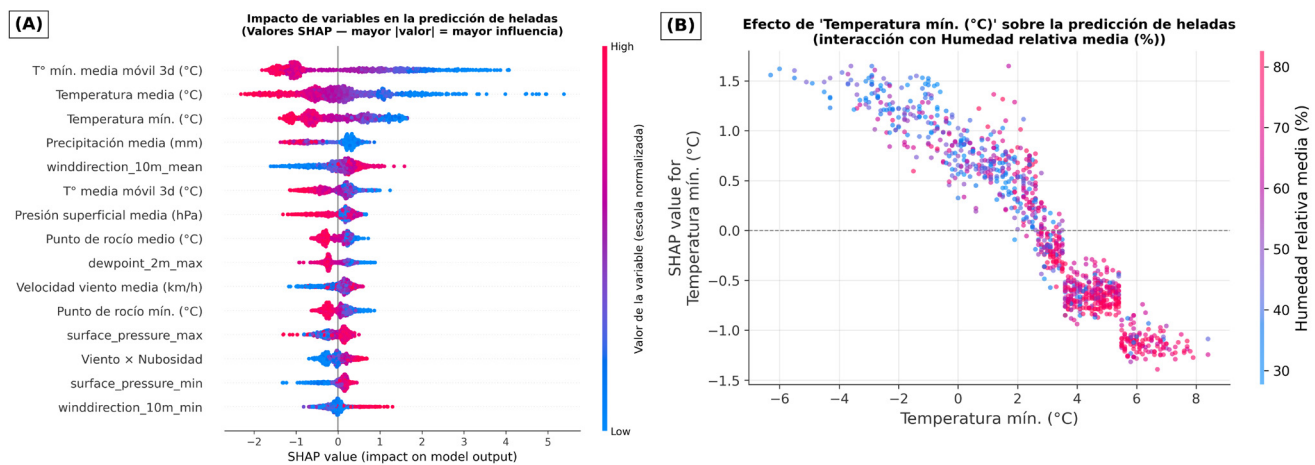
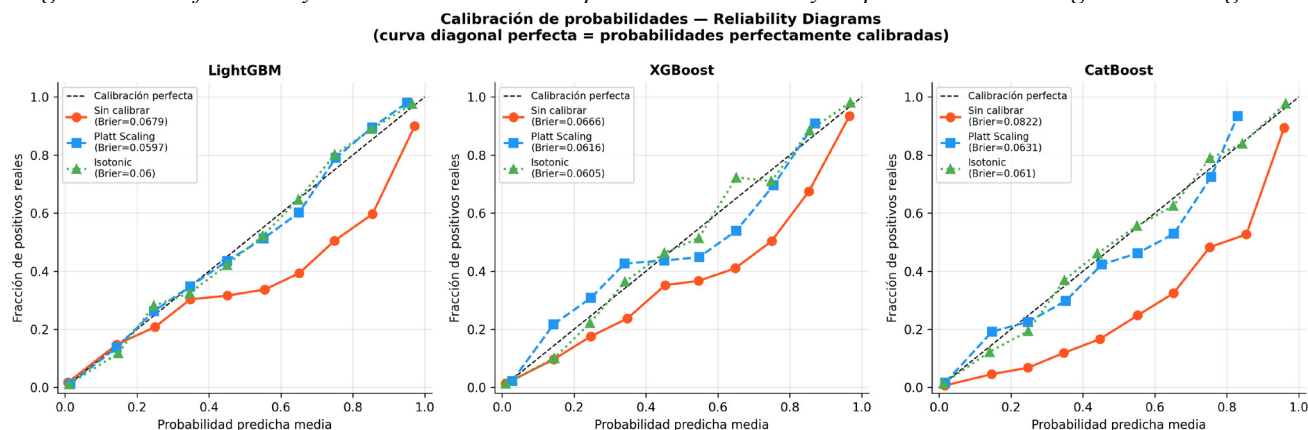


Figura 7

Diagramas de confiabilidad y curvas de calibración de probabilidades antes y después de Platt Scaling e Isotonic Regression



Discusión

Los resultados sugieren que los métodos de conjunto potenciados son competitivos y robustos para la predicción de heladas en el altiplano peruano. El AUC-ROC obtenido (0,913 – 0,957) es, por tanto, más alto que los reportados por Rozante et al. (2023) para Random Forest en la predicción de heladas en Brasil (AUC=0,85), además de ser similar a los de Lira et al. (2022), que extrajeron características de una variable heterogénea mediante atención espaciotemporal con redes neuronales (AUC = 0,91- 0,94) para la predicción de heladas. Los estudios de pronóstico agrometeorológico basados en ML (Diedrichsen et al., 2023) demostraron recientemente que los algoritmos de boosting superan de manera consistente a los métodos estadísticos clásicos en contextos desbalanceados. Además, el enfoque propuesto se beneficia de (a) datos de reanálisis con alta densidad temporal, (b) características diseñadas relevantes para el altiplano y (c) la validación espacial LOSO.

La identificación del umbral óptimo en función de la prevalencia local ($t^* = 0,18$ en Ilave versus $t^* = 0,82$ en Puno) tiene consecuencias directas para la implementación en campo de sistemas de alerta: no existe un único umbral óptimo en toda la región, con diferencias sustanciales entre subregiones. Este hallazgo corrobora lo que Baño-Medina et al. (2021) han venido mostrando, al señalar la necesidad de calibración a escala local de los modelos de predicción climática en Sudamérica.

La preponderancia de la temperatura mínima de la media móvil de 3 días en el análisis de SHAP se ajusta bien a nuestra comprensión de la física que subyace a este fenómeno: el enfriamiento radiativo inducido por heladas sobre terreno plano se modula mediante condiciones persistentes de masa de aire frío acopladas de días antes (Díaz & Vera, 2024). La distinción de la dirección del viento entre las cinco variables principales también reafirma el papel que desempeñan los sistemas anticiclónicos del Pacífico Sur en la advección de aire continental frío (Arias et al., 2021).

Sobre la adecuación de los datos de entrada, Huerta et al. (2023) analizaron la temperatura distribuida

especialmente en el Perú utilizando reanálisis frente a los registros de estaciones del SENAMHI en la región del altiplano y encontraron una buena concordancia entre ambos conjuntos, lo que también respalda la confiabilidad de los predictores generados para este estudio.

Las limitaciones del estudio incluyeron las siguientes: (a) no se utilizaron observaciones directas de las estaciones del SENAMHI, lo que generó incertidumbre en la representación local, aunque la resolución de 9 km obtenida para ERA5-Land es adecuada para la escala regional (Muñoz-Sabater et al., 2021); y b) aunque este estudio ofrece una herramienta para la alerta temprana a corto plazo (1 a 3 días), extenderla a pronósticos de subestacional a mensual (15 a 30 días) es una línea prioritaria de extensión para la gestión estratégica y la planificación meteorológica de la siembra agrícola. Esto requerirá vincular modelos de aprendizaje automático con forzamientos climáticos a gran escala (por ejemplo, las teleconexiones de ENSO, MJO y SAM) e información de reforecast acoplados océano-atmósfera.

Conclusiones

Los modelos de ensamble de aprendizaje automático, en particular el modelo Voting (que integra Random Forest, XGBoost, LightGBM y CatBoost), mostraron una capacidad predictiva superior para la clasificación de heladas meteorológicas en la región de Puno, con un rendimiento discriminatorio excelente (AUC-ROC = 0,957 a 24 h; 0,926 a 48 h; y 0,913 a 72 h).

La meseta alta de Puno presentó una distribución heterogénea de heladas, evidenciando la escala espacial (del 7,7 % en Puno al 31,3 % en Ilave) y confirmando la influencia termorreguladora del lago Titicaca sobre esta zona, que conduce a temperaturas más templadas y a menos heladas que en las áreas continentales circundantes.

El ajuste del umbral supuso una ganancia considerable para la F1-Score regional (+0,025 a 72 h), lo que indica que los sistemas de alerta agrometeorológica no deben basarse en un umbral homogéneo ($t = 0,50$),

sino en la calibración del umbral según firmas espaciales (por ejemplo, $t = 0,18$ en Ilave frente a $t = 0,82$ en Puno).

El análisis de interpretabilidad con valores SHAP indicó que la temperatura mínima de la media móvil de 3 días fue el predictor con mayor impacto (media $|\text{SHAP}| = 1,274$), seguida por la temperatura media diaria y la temperatura mínima, lo que resalta la importancia física de la acumulación de aire frío en los días previos a los eventos de helada por irradiación.

El escalado de Platt disminuyó el *Brier Score* (0,082 a 0,063 en *CatBoost*), proporcionando estimaciones probabilísticas fiables y bien ajustadas para decisiones agrícolas preventivas en cultivos estratégicos del altiplano andino.

Conflicto de intereses

Los autores declaran no tener conflictos de intereses.

Referencias

- Arias, P. A., Garreaud, R., Poveda, G., Espinoza, J. C., Molina-Carpio, J., Masiokas, M., Viale, M., Scaff, L., & van Oevelen, P. J. (2021). Hydroclimate of the Andes Part II: Hydroclimate variability and sub-continental patterns. *Frontiers in Earth Science*, 8, 505467. <https://doi.org/10.3389/feart.2020.505467>
- Baño-Medina, J., Manzanar, R., & Gutiérrez, J. M. (2021). On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections. *Climate Dynamics*, 57(11-12), 2941-2951. <https://doi.org/10.1007/s00382-021-05847-0>
- Benique-Olivera, E., & Ojeda-Tito, A. (2024). Impacto del cambio climático en la producción y rendimiento de quinua (*Chenopodium quinoa* Willd) en la provincia de Azángaro, región del Altiplano-Puno, Perú. *Revista de Investigaciones Altoandinas*, 26(3), 154-160. <https://doi.org/10.18271/ria.2024.618>
- Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3), 1937-1967. <https://doi.org/10.1007/s10462-020-09896-5>
- Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619, 533-538. <https://doi.org/10.1038/s41586-023-06185-3>
- Chen, H., Covert, I. C., Lundberg, S. M., & Lee, S.-I. (2023). Algorithms to estimate Shapley value feature attributions. *Nature Machine Intelligence*, 5(6), 590-601. <https://doi.org/10.1038/s42256-023-00657-x>
- Dewitte, S., Cornelis, J. P., Müller, R., & Munteanu, A. (2021). Artificial intelligence revolutionises weather forecast, climate monitoring and decadal prediction. *Remote Sensing*, 13(16), 3209. <https://doi.org/10.3390/rs13163209>
- Díaz, L. B., & Vera, C. S. (2024). Extreme indices of temperature and precipitation in South America: trends and intercomparison of regional climate models. *Climate Dynamics*, 62(2), 1421-1439. <https://doi.org/10.1007/s00382-022-06598-2>
- Diedrichsen, E., Schmidhalter, U., & Trevisan, R. G. (2023). Prediction of frost events using machine learning and deep learning for smart agriculture. *Computers and Electronics in Agriculture*, 209, 107815. <https://doi.org/10.1016/j.compag.2023.107815>
- Grisel, O., Mueller, A., Lars, Gramfort, A., Louppe, G., Fan, T. J., Prettenhofer, P., Blondel, M., Niculae, V., Nothman, J., Joly, A., Lemaitre, G., Estève, L., Vanderplas, J., du Boisberranger, J., Kumar, M., Qin, H., Hug, N., Varoquaux, N., . . . Lorentzen, C. (2024). scikit-learn/scikit-learn: Scikit-learn 1.4.0 (Versión 1.4.0) [Software de computación]. Zenodo. <https://doi.org/10.5281/zenodo.10532824>
- Huerta, A., Aybar, C., & Lavado-Casimiro, W. (2023). PISCO temperature v1.2: a spatiotemporal reconstruction of air temperature for Peru. *Earth System Science Data*, 15(1), 345-365. <https://doi.org/10.5194/essd-15-345-2023>
- Junquas, C., Espinoza, J. C., Segura, H., Lucas-Picher, P., Boudin, M., Condom, T., & Lebel, T. (2022). Regional climate modeling of the diurnal cycle of precipitation and associated atmospheric circulation patterns over an Andean glacier region. *Climate Dynamics*, 58(9-10), 2873-2895. <https://doi.org/10.1007/s00382-021-06079-y>
- Kaur, H., Pannu, H. S., & Malhi, A. K. (2022). A systematic review on imbalanced data challenges in machine learning: Applications and solutions. *ACM Computing Surveys*, 52(4), 1-36. <https://doi.org/10.1145/3505244>
- Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Meroze, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., & Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382(6677), 1416-1421. <https://doi.org/10.1126/science.adf2336>
- Lira, H., Martí, L., & Sanchez-Pi, N. (2022). A graph neural network with spatio-temporal attention for multi-sources time series data: An application to

- frost forecast. *Sensors*, 22(4), 1486. <https://doi.org/10.3390/s22041486>
- MINAGRI. (2021). Plan de gestión de riesgo agrario frente a heladas y friaje 2021-2030. Ministerio de Agricultura y Riego del Perú. <https://www.midagri.gob.pe>
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>
- Nguyen, T., Shah, R., Sharma, P., Vu, T., Nguyen, H., Burke, J., Bonfils, C., Entekhabi, D., & Bhattacharya, A. (2023). ClimaX: A foundation model for weather and climate. Proceedings of the 40th International Conference on Machine Learning (ICML 2023), PMLR 202, 25904-25938. <https://doi.org/10.48550/arXiv.2301.10343>
- Quispe-Mamani, J. F., Yucra-Quispe, L. F., Loza-Murguía, M., & Chipana-Cutipa, E. (2023). Climate variability and agricultural production in the Peruvian Altiplano: impacts on food security. *Agriculture*, 13(4), 847. <https://doi.org/10.3390/agriculture13040847>
- Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. (2024). WeatherBench 2: A benchmark for the next generation of data-driven global weather models. *Journal of Advances in Modeling Earth Systems*, 16(1), e2023MS004019. <https://doi.org/10.1029/2023MS004019>
- Rozante, J. R., Silveira Coelho, C. A., Silveira Pires, L., & Pilloso, F. M. (2023). Improved frost forecast using machine learning. *Artificial Intelligence in Geosciences*, 4, 100058. <https://doi.org/10.1016/j.aiig.2023.10.001>
- Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., Mozaffari, A., & Stadler, S. (2021). Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A*, 379(2194), 20200097. <https://doi.org/10.1098/rsta.2020.0097>
- SENAMHI. (2023). Atlas de heladas del Perú. Servicio Nacional de Meteorología e Hidrología del Perú. <https://www.senamhi.gob.pe>
- Zippenfenig, P. (2023). Open-Meteo.com Weather API [software]. Zenodo. <https://doi.org/10.5281/zenodo.7970649>