Universidad Nacional Federico Villarreal

VRIN | VICERRECTORADO DE INVESTIGACIÓN

ESCUELA UNIVERSITARIA DE POSGRADO

MODELO MATEMÁTICO ANTE DATOS ERRÓNEOS DE TEMPERATURA DEL AIRE PARA EL CONTROL DE CALIDAD EN LA SIERRA DEL DEPARTAMENTO DE PUNO DEL PERÚ

Línea de investigación:

Biodiversidad, ecología y conservación

Tesis para optar el grado académico de Maestro en Gestión Ambiental

Autor

Solis Caceres, Olimpio

Asesor

Esenarro Vargas, Doris

ORCID: 0000-0002-7186-9614

Jurado

Marín Machuca, Olegario

Veliz Garagatti, María Herlinda

Ventura Guevara, Luis Leónidas

Lima - Perú

2026

RECONOCIMIENTO - NO COMERCIAL - SIN OBRA DERIVADA
(CC BY-NC-ND)

MODELO MATEMÁTICO ANTE DATOS ERRÓNEOS DE TEMPERATURA DEL AIRE PARA EL CONTROL DE CALIDAD EN LA SIERRA DEL DEPARTAMENTO DE PUNO DEL PERÚ

INFORME DE ORIGINALIDAD

16%	14%	4%	6%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	www.senamhi.gob.pe Fuente de Internet	1%
2	www.coursehero.com Fuente de Internet	1%
3	hdl.handle.net Fuente de Internet	1%
4	library.wmo.int Fuente de Internet	1%
5	repositorio.unfv.edu.pe Fuente de Internet	1%
6	Submitted to Universidad Nacional Federico Villarreal Trabajo del estudiante	1%
7	biblioteca.semarnat.gob.mx Fuente de Internet	<1%
8	repositorioacademico.upc.edu.pe Fuente de Internet	<1%
9	uvadoc.uva.es Fuente de Internet	<1%
10	manglar.uninorte.edu.co Fuente de Internet	<1%
11	Submitted to UNIBA Trabajo del estudiante	<1%

Universidad Nacional Federico Villarreal

VRIN | VICERRECTORADO DE INVESTIGACIÓN

ESCUELA UNIVERSITARIA DE POSGRADO

MODELO MATEMÁTICO ANTE DATOS ERRÓNEOS DE TEMPERATURA DEL AIRE PARA EL CONTROL DE CALIDAD EN LA SIERRA DEL DEPARTAMENTO DE PUNO DEL PERÚ

Línea de investigación:

Biodiversidad, ecología y conservación

Tesis para optar el grado académico de Maestro en Gestión Ambiental

Autor

Solis Caceres, Olimpio

Asesora

Esenarro Vargas, Doris

ORCID: 0000-0002-7186-9614

Jurado

Marín Machuca, Olegario

Veliz Garagatti, María Herlinda

Ventura Guevara, Luis Leónidas

Lima – Perú

2026

DEDICATORIA

A mis padres, **Olimpio Solís y Luidmer Cáceres**, y a mis hermanas, cuyo apoyo incondicional y amor han sido mi pilar a lo largo de mi vida.

A mi esposa, **Marlene Facundo**, por su paciencia, comprensión y compañía en los momentos de esfuerzo y desvelo. Su constante motivación ha sido fundamental para superar los obstáculos.

A mis hijas **Alejandra, Stephany y Andrea**, quienes siempre estuvieron dispuestas a brindarme palabras de aliento, su amor y sus risas cuando más las necesitaba.

A la memoria de mis suegros, **Ausberto Facundo Ticliahuanca y Lucía Neyra Puelles**, mis abuelitos **Mercurio Caceres, Marcelo Caceres, Delfín Caceres**, y a mi hermana **Lucinda Judith Solis Caceres** a mi abuelita **María Reyes Chaupis** quienes descansan en paz y viven eternamente en mi recuerdo.

AGRADECIMIENTOS

Al SENAMHI

Expreso mi más sincero reconocimiento al Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) por brindarme las herramientas técnicas y el acceso a información clave que hicieron posible esta investigación.

A mi asesora

A la Dra. Doris Esenarro, mi mentora y guía en este proceso, cuyo rigor académico, disponibilidad y valiosos aportes enriquecieron sustancialmente este trabajo.

A los Expertos del SENAMHI

Agradezco al Dr. Julio Ordoñez, Dr. Waldo Lavado, Mag. Marco Paredes y al Mag. Sixto Flores, por su apoyo técnico que han sido fundamental para el desarrollo de esta tesis.

A mis maestros

A mis profesores y mentores, cuyo legado de conocimiento y pasión por la ciencia ambiental sembró en mí el deseo de superación constante y el compromiso con la excelencia académica.

A mis compañeros de estudio

A Luis Sarmiento y Milagros Caycho, colegas y amigos cuyo apoyo emocional y palabras de aliento fueron mi sustento en los momentos más demandantes de este proceso.

ÍNDICE

RESUMEN 14

ABSTRACT 15

I. INTRODUCCIÓN 16

1.1. Planteamiento del problema 16

1.2. Descripción del problema 17

1.3. Formulación del problema 19

1.3.1. Problema general 19

1.4. Antecedentes 19

1.4.1. Antecedentes internacionales 19

1.4.2. Antecedentes nacionales 24

1.5. Justificación de la investigación 27

1.5.1. Justificación científica 27

1.5.2. Justificación técnica 28

1.5.3. Justificación socioeconómica 28

1.5.4. Justificación ambiental 28

1.5.5. Impacto en agricultura 29

1.6. Limitaciones de la investigación 29

1.7. Objetivos 30

1.7.1. Objetivo general 30

1.7.2. Objetivos específicos 30

1.8. Hipótesis 31

1.8.1. Hipótesis general.....31

1.8.2. Hipótesis específicas.....31

II. MARCO TEÓRICO.....32

2.1. Base teórica.....32

2.1.1. Errores en mediciones meteorológicas.....32

2.1.2. Tipos de errores en mediciones.....32

2.1.3. Impacto de los errores en la calidad de datos.....33

2.1.4. Normas internacionales y protocolos nacionales.....33

2.1.5. Series temporales en climatología.....33

2.1.6. Definición y características de las series temporales.....34

2.1.7. Aplicación de series temporales en climatología andina.....34

2.1.8. Métodos estadísticos para la detección de anomalías.....34

2.1.9. Modelos matemáticos aplicados a datos climáticos.....35

2.1.10. Modelos de regresión en climatología.....35

2.1.11. Modelos ARIMA y series estocásticas.....36

2.1.12. Interpolación espacial y temporal.....36

2.1.13. Detección de outliers en datos climáticos.....36

2.1.14. *Aprendizaje automático y control de calidad de datos climáticos*.....37

2.1.15. Algoritmos supervisados aplicados a datos climáticos.....37

2.1.16. Algoritmos no supervisados y detección de anomalías.....38

2.1.17. Redes neuronales aplicadas a la climatología andina.....38

2.1.18. Comparación entre modelos estadísticos y de aprendizaje automático 38

2.1.19. Control de calidad de datos meteorológicos en el Perú 39

2.1.20. Protocolos institucionales de control de calidad 39

2.1.21. Limitaciones en regiones altoandinas 40

2.1.22. Inconsistencias técnicas y necesidades de mantenimiento 40

2.1.23. Proyectos de corrección y mejora de datos en la sierra peruana 40

2.1.24. Brechas identificadas en el control de calidad en Puno 41

2.2. Marco conceptual 41

2.2.1. Control de calidad de datos meteorológicos 41

2.2.2. Modelo matemático 43

2.2.3. Series temporales climáticas 43

2.2.4. Estándares internacionales 44

2.2.5. Errores en mediciones 44

2.2.6. Errores en los datos meteorológicos 44

2.2.7. Estación meteorológica convencional 44

2.2.8. Estación meteorológica automática (EMA) 44

2.2.9. Conceptos de random forest 45

2.2.10. Fundamentos de gradient boosting 47

2.2.11. Técnica de validación cruzada para los modelos forest y gradiente boosting 49

2.2.12. Fórmulas de las métricas utilizadas 49

2.3. Bases legales 51

2.3.1. Organización meteorológica mundial (OMM) 51

2.3.2. Servicio nacional de meteorología e hidrología del Perú (SENAMHI)..51

2.3.3. Ley N° 28303- Ley del sistema nacional de ciencia, tecnología e innovación tecnológica. 51

2.3.4. Política nacional de cambio climático (PNCC) del Perú 51

III. MÉTODO 52

3.1. Tipo de investigación.....52

3.2. Población y muestra.....53

3.2.1. Población 53

3.2.2. Muestra.....54

3.3. Operacionalización de variables.....55

3.3.1 Variable independiente (X): modelo matemático.....57

3.3.2 Variable dependiente (Y): control de calidad 57

3.4. Instrumentos 57

3.5. Procedimientos 58

3.5.1. Etapa 1: recopilación y preparación de datos.....58

3.5.2. Etapa 2: desarrollo del control de calidad 59

3.5.3. Etapa 3: diseño e implementación del modelo matemático 60

3.5.4. Etapa 4: validación y evaluación del modelo 62

3.5.5. Etapa 5: Procedimiento para la detección, corrección y relleno de datos 64

3.6. Análisis de datos.....71

3.6.1. Análisis para la detección de anomalías....71

3.6.2. Análisis de corrección y relleno de datos....73

3.6.3. Análisis de la validación operativa....80

3.7. Consideraciones éticas....86

3.7.1. Uso responsable de datos....86

3.7.2. Veracidad y honestidad en la recolección y análisis de información....86

3.7.3. Reconocimiento de fuentes y derechos de autor....86

3.7.4. Confidencialidad y protección de la información....87

IV. RESULTADOS....88

4.1. Resultados sobre la cuantificación de la incertidumbre de los datos....88

4.2. Consistencia espacial de los datos corregidos....89

4.3. Resultado sobre la eficacia del algoritmo del modelo....90

4.3.1. Consistencia del modelo en validación cruzada....90

V. DISCUSIÓN DE RESULTADOS....92

VI. CONCLUSIONES....93

6.1. Conclusión general....93

6.2. Conclusiones específicas....93

6.2.1. Cuantificación de la incertidumbre de los datos....93

6.2.2. Consistencia espacial de los datos corregidos....94

6.2.3. Eficacia del algoritmo del modelo matemático....94

VII. RECOMENDACIONES....95

7.1. Recomendaciones para la implementación operativa del modelo.....95

7.2. Recomendaciones para el fortalecimiento de la red de estaciones meteorológicas 95

7.3. Recomendaciones para la capacitación del personal técnico 96

7.4. Recomendaciones para la expansión del modelo a otras variables meteorológicas 96

7.5. Recomendaciones para la investigación futura 96

7.6. Recomendaciones para la gestión de datos y accesibilidad.....97

7.7. Recomendaciones para el monitoreo continuo del modelo 97

7.8. Recomendaciones para la articulación interinstitucional 97

VIII. REFERENCIAS.....98

IX. ANEXOS 104

Anexo A Información transversal a todas las fases 105

Anexo B Recolección de datos y selección de estaciones 124

Anexo C Detección de anomalías parte 1 129

Anexo D Corrección y relleno de datos.....151

Anexo E Validación de eficacia del modelo.....162

ÍNDICE DE TABLAS

Tabla 1. *Operacionalización de variable independiente* 55

Tabla 2. *Operacionalización de variable dependiente* 56

Tabla 3. *Indicadores del control de calidad* 72

Tabla 4. *Indicadores estadísticos de corrección y relleno*.....75

Tabla 5. *Resumen de matriz de correlación de datos original vs rellenado*.....77

Tabla 6. *Evolución de métricas por fold* 80

Tabla 7. *Test de friedman*.....80

Tabla 8. *Método de evaluación de resultados de experto*.....84

Tabla 9. *Relación de tablas y figuras del anexo A*.....105

Tabla 10. *Matriz de consistencia* 107

Tabla 11. *Matriz de operacionalización de variables*.....109

Tabla 12. *Formato de registro de datos diarios meteorológicos*.....110

Tabla 13. *Informe de juicio de experto del instrumento de investigación* 111

Tabla 14. *Ficha de validación del instrumento experto 1*.....112

Tabla 15. *Ficha de validación del instrumento experto 2*.....113

Tabla 16. *Ficha de validación del instrumento experto 3*.....114

Tabla 17. *Ficha de validación del instrumento experto 4*.....115

Tabla 18. *Resultados de evaluación de expertos* 116

Tabla 19. *Método de evaluación de resultados de experto*.....119

Tabla 20. *Prueba test–retest* 122

Tabla 21. *Resultado de la prueba de test retest* 123

Tabla 22. *Categoría de la prueba* 123

Tabla 23. *Relación de tablas y figuras del anexo B*.....124

Tabla 24. *Red de estaciones meteorológicas del SENAMHI* 125

Tabla 25. *Relación de tablas y figuras del anexo C*.....129

Tabla 26. *Reglas de control de calidad de datos de temperatura del aire* 132

Tabla 27. *Umbrales de temperatura del aire media por estación* 135

Tabla 28. *Tabla de indicadores de distribución de anomalías parte 1*.....143

Tabla 29. *Indicadores de distribución de anomalías parte 2* 144

Tabla 30. *Relación de tablas y figuras del anexo D* 152

Tabla 31. *Indicadores estadísticos de corrección y relleno*.....154

Tabla 32. *Resumen de matriz de correlación de datos original vs rellenado*.....159

Tabla 33. *Relación de figuras del anexo E* 162

ÍNDICE DE FIGURAS

Figura 1. *Gráfico de indicadores de control de calidad*.....72

Figura 2. *Estadística de reducción de incertidumbre*.....76

Figura 3. *Gráfico de porcentaje de aumento de coeficiente (r)*.....77

Figura 4. *Coeficiente de correlación*.....123

Figura 5. *Mapa de la red de estaciones meteorológicas del SENAMHI*.....126

Figura 6. *Mapa de distancia de estaciones*.....127

Figura 7. *Mapa de clasificación climática del SENAMHI – 2021*.....128

Figura 8. *Flujo del modelo de control de calidad*.....134

Figura 9. *Umbrales de las estaciones meteorológicas de Puno*.....136

Figura 10. *Temperatura con anomalías detectadas grupo 1*.....137

Figura 11. *Temperatura con anomalías detectadas grupo 2*.....138

Figura 12. *Temperatura con anomalías detectadas grupo 3*.....139

Figura 13. *Temperatura con anomalías detectadas grupo 4*.....140

Figura 14. *Temperatura con anomalías detectadas grupo 5*.....141

Figura 15. *Temperatura con anomalías detectadas grupo 6*.....142

Figura 16. *Distribución de anomalías por estación*.....145

Figura 17. *Mapa de distribución de anomalías por estación grupo 1*.....146

Figura 18. *Mapa de distribución de anomalías por estación grupo 2*.....147

Figura 19. *Mapa de distribución de anomalías por estación grupo 3*.....148

Figura 20. *Mapa de distribución de anomalías por estación grupo 4*.....149

Figura 21. *Mapa de distribución espacial de anomalías categorizadas*.....150

Figura 22. *Estadística de reducción de incertidumbre*.....155

Figura 23. *Gráfico de diagrama de dispersión de temperatura diaria*.....156

Figura 24. *Gráfico de comparación del modelo*.....156

Figura 25. *Mapa de correlación espacial original* 157

Figura 26. *Mapa de correlación espacial corregido o rellenado* 158

Figura 27. *Gráfico de predicciones versus valores reales* 160

Figura 28. *Mapa de distribución de residuales* 161

Figura 29. *Gráfico de evolución de métricas por fold de validación cruzada* 163

Figura 30. *Gráfico de importancia de características de las variables* 163

Figura 31. *Diagrama del modelo híbrido (RF + GB)* 164

RESUMEN

En la sierra del departamento de Puno, los registros de temperatura del aire presentan errores sistemáticos, valores atípicos y datos faltantes, lo que compromete la confiabilidad de los análisis climáticos y la toma de decisiones en sectores como la agricultura, la gestión del agua y la prevención de desastres. Este estudio desarrolló un modelo matemático que integra técnicas de machine learning (Random Forest y Gradient Boosting), herramientas estadísticas y análisis geoespacial para el control de calidad, corrección e imputación de series de temperatura en estaciones meteorológicas. Se emplearon datos históricos del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) del período 2001-2024, provenientes de 22 estaciones ubicadas por encima de los 3,800 m s. n. m. El modelo comprendió cuatro fases: recopilación y depuración de datos, detección de anomalías, implementación del algoritmo híbrido y validación de resultados. La evaluación se realizó mediante métricas como el error absoluto medio (MAE), la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación (R^2), complementadas con la prueba de Friedman. Los resultados mostraron que el modelo redujo en más del 30% el error en los registros originales y mejoró la consistencia espacial entre estaciones vecinas ($r > .90$), superando los métodos convencionales de control de calidad. Estos hallazgos confirman la eficacia del modelo propuesto para optimizar la calidad de los datos meteorológicos en regiones altoandinas y su aplicabilidad en otros contextos con condiciones climáticas extremas.

Palabras clave: control de calidad, datos meteorológicos, machine learning, temperatura del aire

ABSTRACT

In the highlands of Puno, Peru, air temperature records often contain systematic errors, outliers, and missing values, which compromise the reliability of climate analyses and decision-making in sectors such as agriculture, water management, and disaster prevention. This study developed a mathematical model integrating machine learning techniques (Random Forest and Gradient Boosting), statistical tools, and geospatial analysis for quality control, correction, and imputation of temperature series from meteorological stations. Historical data from the National Meteorology and Hydrology Service of Peru (SENAMHI) were used, covering the period 2001–2024 and including 22 stations located above 3,800 m a.s.l. The model comprised four phases: data collection and cleaning, anomaly detection, implementation of the hybrid algorithm, and validation of results. The evaluation employed metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the Coefficient of Determination (R^2), complemented by the Friedman statistical test. The results showed that the model reduced the error in the original records by more than 30% and improved spatial consistency between neighboring stations ($r > .90$), outperforming conventional quality control methods. These findings confirm the effectiveness of the proposed model in optimizing the quality of meteorological data in high Andean regions and its applicability in other geographical contexts with extreme climatic conditions.

Keywords: quality control, meteorological data, machine learning, air temperature

I. INTRODUCCIÓN

1.1. Planteamiento del problema

En el contexto global, los errores en datos meteorológicos constituyen un problema que afecta la precisión de modelos climáticos y pronósticos. La Organización Meteorológica Mundial (OMM, 2021) reporta que *"el 15% de los datos de estaciones en zonas montañosas presentan errores sistemáticos"* (p. 45). Factores como cambios tecnológicos, variabilidad en métodos de medición y errores humanos en la transcripción inciden en la calidad de los registros históricos (Pérez, 2019, p. 24). Antes de 1880, las incertidumbres superaban ±0.5 °C debido a la escasez de estaciones meteorológicas, y el calentamiento global ha reducido la predictibilidad de los pronósticos, acortando en un día la ventana de confianza con solo un incremento de 2 °C (NASA Science Editorial Team, 2021).

En el contexto nacional, el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI, 2022) informó que *"el 30% de las estaciones en Puno requieren recalibración debido a datos atípicos"* (p. 12). Villanueva (2022, p. 45) advierte que *"los errores en datos de temperatura afectan hasta un 30% la precisión de modelos climáticos en regiones altoandinas"*. Asimismo, se señala que los sistemas de control de calidad como el *Quality Assurance/Quality Control* (QA/QC) de la NOAA no consideran adecuadamente la variabilidad altitudinal (Smith et al., 2020, p. 78).

En el contexto local, la sierra de Puno presenta características geográficas y climáticas extremas, con temperaturas nocturnas bajo 0 °C y alta incidencia de heladas, lo que demanda datos meteorológicos de alta calidad. En Puno, el 30% de las estaciones meteorológicas reportan datos inconsistentes (SENAMHI, 2022), impactando de forma directa los pronósticos agrícolas y la gestión hídrica. El SENAMHI (2017, p. 5) identificó que entre el 10% y el 20% de las estaciones en los Andes centrales presentan intervalos de temperatura faltantes por errores en la lectura de termómetros analógicos. En Progreso (Puno), se han detectado

discrepancias de ±4.2 °C por malinterpretación de escalas en instrumentos de mercurio (Rubio, 2008, p. 5). La escasa capacitación de observadores meteorológicos y la ausencia de controles de calidad en tiempo real agravan esta situación (SENAMHI, 2017, p. 5). Entre los problemas más frecuentes en la región se encuentran los outliers por radiación solar intensa (Pérez, 2019, p. 34) y vacíos de datos en estaciones remotas (Gómez y Lima, 2021, p. 56).

1.2. Descripción del problema

En un contexto global, estudios realizados por la Administración Nacional de Aeronáutica y del Espacio, agencia del gobierno de Estados Unidos (NASA Science Editorial Team, 2021, p. 1), señalan que los errores en datos de temperatura atmosférica representan un desafío crítico para la modelización climática y la gestión ambiental. Asimismo, destacan que la precisión de los registros históricos de temperatura depende de factores como cambios tecnológicos, variabilidad en los métodos de medición y errores humanos en la transcripción de datos. Antes de 1880, las incertidumbres superaban ±0.5 °C debido a la escasez de estaciones meteorológicas. Investigaciones recientes indican que el calentamiento global reduce la predictibilidad meteorológica: un aumento de 2 °C en la temperatura acorta en un día la ventana de pronósticos confiables.

Por su parte, sistemas como el *Quality Assurance/Quality Control* (QA/QC) de la NOAA no consideran adecuadamente la variabilidad altitudinal (Smith et al., 2020, p. 78), lo que limita su eficacia en zonas con topografía compleja.

A nivel local, en la sierra de Puno (Perú), estos problemas se intensifican debido a condiciones geográficas y socioeconómicas particulares. Investigaciones del SENAMHI (2017, p. 5) identificaron que entre el 10% y el 20% de las estaciones meteorológicas en los Andes centrales presentan intervalos de temperatura faltantes por errores en la lectura de termómetros analógicos. En Progreso (Puno), por ejemplo, se han registrado discrepancias de ±4.2 °C debido a la malinterpretación de escalas en instrumentos de mercurio.

La vulnerabilidad climática de Puno, caracterizada por temperaturas nocturnas bajo 0 °C y una alta incidencia de heladas, exige datos meteorológicos precisos para respaldar actividades agrícolas que representan el 16.3% del PBI regional. Sin embargo, factores como la ocupación de llanuras aluviales, la falta de infraestructura de drenaje y el uso de materiales de construcción inadecuados limitan la fiabilidad de las mediciones (Rubio, 2008, p. 5).

Estos problemas se ven agravados por la escasa capacitación de los observadores meteorológicos y la ausencia de controles de calidad en tiempo real, lo que reduce la capacidad de respuesta ante fallos en la medición o transmisión de datos (SENAMHI, 2017, p. 5). En el caso específico de Puno, se han identificado dos situaciones críticas que comprometen la confiabilidad de la información meteorológica. La primera corresponde a la presencia de valores atípicos, o outliers, originados principalmente por la radiación solar intensa, que altera las lecturas de temperatura y genera desviaciones significativas respecto de los patrones esperados (Pérez, 2019, p. 34). La segunda está relacionada con los vacíos de datos en estaciones ubicadas en zonas remotas, donde las condiciones geográficas y la limitada conectividad dificultan la transmisión continua de información, provocando interrupciones en las series históricas y limitando los análisis comparativos a largo plazo (Gómez y Lima, 2021, p. 56).

1.3. Formulación del problema

1.3.1. Problema general

¿En qué medida el modelo matemático, basado en técnicas de inteligencia artificial (Random Forest + Gradient Boosting), estadísticas y geoespaciales, mejora el control de calidad para la corrección y rellenos de datos erróneos de temperatura del aire en las estaciones meteorológicas de la Sierra de Puno?

1.3.2. Problema específico

Para abordar el problema general, se identifican preguntas específicas que orientarán el desarrollo de la investigación:

- ¿En qué medida el modelo matemático, basado en técnicas de inteligencia artificial (Random Forest + Gradient Boosting), estadísticas y geoespaciales, influye en la reducción de la incertidumbre de los datos de temperatura para mejorar la precisión en el control de calidad de las estaciones meteorológicas de la Sierra de Puno?
- ¿En qué medida el modelo matemático basado en análisis geoespacial influye en el mejoramiento de la consistencia espacial de los datos de temperatura entre estaciones vecinas para fortalecer la confiabilidad en el control de calidad?

1.4. Antecedentes

1.4.1. Antecedentes internacionales

Montiel (2017) desarrolló un modelo matemático para la evaluación de peligros naturales mediante información geográfica, en un estudio publicado en la Red de Revistas Científicas de América Latina, el Caribe, España y Portugal. La investigación se llevó a cabo en la región de La Guajira, Colombia, un área con condiciones climáticas extremas y limitadas estaciones meteorológicas, lo que la convierte en un escenario ideal para probar metodologías de predicción y control de calidad de datos ambientales. La metodología, de enfoque descriptivo, se estructuró en cuatro fases: descripción de sistemas de medición, identificación

de variables de peligros ambientales, modelación matemática y validación de los modelos. El estudio empleó técnicas de series temporales y control predictivo para mejorar la precisión de los datos.

Hallazgos y conclusiones del antecedente: El sistema predictivo ARMAX 2221 logró un ajuste del 70% en series temporales de temperatura con un horizonte de predicción de 50 meses, mostrando el potencial de los métodos adaptativos para corregir no linealidades en datos ambientales. Sin embargo, la variable precipitación no pudo ser predicha con precisión debido a la falta de datos, mientras que la temperatura y la velocidad del viento sí fueron pronosticadas con dos años de antelación. Estos resultados demuestran que la calidad de los datos y la disponibilidad histórica influyen directamente en la capacidad predictiva, un aspecto clave también presente en la realidad de Puno.

La Organización Meteorológica Mundial (OMM, 2018b) publicó directrices internacionales para la gestión de calidad de datos climáticos, dirigidas a servicios meteorológicos nacionales. El documento enfatiza la necesidad de establecer sistemas de control y seguimiento de procesos, documentación estructurada y mejoras continuas, siguiendo las recomendaciones del Guidelines on Quality Management in Climate Services (WMO-No. 1221). Estas pautas se basan en experiencias internacionales que muestran cómo la estandarización mejora la comparabilidad y confiabilidad de los registros.

Hallazgos y conclusiones del antecedente: La OMM destaca que una adecuada gestión de calidad garantiza que los datos sean consistentes y trazables, permitiendo su uso en modelación climática global. Este marco normativo es aplicable a contextos como el peruano, donde la homogeneidad de la información es un reto debido a la diversidad geográfica y climática.

OMM (2023, p. 24) amplió las recomendaciones previas, incorporando la exigencia de sistemas claros de control de versiones y la implementación de “metadatos de procedencia”. Estos metadatos registran el origen, modificaciones y validaciones de cada conjunto de datos, asegurando la integridad y transparencia del proceso de recolección y gestión.

Hallazgos y conclusiones del antecedente: La inclusión de metadatos detallados permite reconstruir el historial de cada registro climático, lo que mejora la capacidad de verificación y la replicabilidad de estudios. Este enfoque es crucial para zonas altoandinas, donde la calidad de los datos puede variar según las condiciones de acceso y mantenimiento de las estaciones.

La OMM (1981, como se citó en IDEAM, 2007, p. 11) estableció que todos los datos meteorológicos destinados al intercambio internacional deben someterse a procedimientos de control de calidad, aplicables a datos de nivel I y II. Posteriormente, OMM (1992, como se citó en IDEAM, 2007, p. 11) reforzó estas disposiciones al indicar que los centros de procesamiento de datos deben implementar verificaciones de errores de cifrado y codificación, coherencia interna, temporal y espacial, así como validaciones contra límites físicos y climatológicos. Estos estándares han sido aplicados a través de sistemas como el Control de Calidad Automático de Datos Meteorológicos (Instituto de Hidrología, Meteorología y Estudios Ambientales [IDEAM], 2007, p. 11).

Hallazgos y conclusiones del antecedente: La estandarización de procedimientos ha permitido reducir errores en el intercambio internacional de datos, mejorando la confiabilidad de la información para modelación global. Sin embargo, estos estándares requieren adaptaciones cuando se aplican en regiones con infraestructura limitada o condiciones climáticas extremas, como es el caso de Puno.

Ferrero et al. (2020) realizaron un estudio en los Alpes suizos para evaluar la efectividad del Kriging universal en la interpolación de datos de temperatura, incorporando la altitud como variable auxiliar. Utilizaron registros de 120 estaciones meteorológicas ubicadas entre los 400

y 3,500 metros de altitud, recopilados durante un periodo de 15 años. El análisis comparó modelos con y sin variable altitudinal, utilizando validación cruzada para evaluar el rendimiento.

Hallazgos y conclusiones del antecedente: La inclusión de la altitud redujo en un 92% el error de estimación, lo que evidencia la relevancia de incorporar variables topográficas en regiones montañosas. Este hallazgo tiene aplicación directa en áreas como Puno, donde la variación altitudinal es significativa y la distribución de estaciones es dispersa.

Rodríguez et al. (2019) desarrollaron un modelo de machine learning basado en Random Forest para detectar outliers en series de temperatura en los Andes colombianos, una región caracterizada por alta variabilidad climática y limitada densidad de estaciones. El estudio utilizó datos históricos de 20 años y comparó el rendimiento del algoritmo frente a métodos estadísticos tradicionales como el Z-score.

Hallazgos y conclusiones del antecedente: El modelo Random Forest superó significativamente a los métodos tradicionales en la detección de anomalías, mostrando mayor precisión y menor tasa de falsos positivos. Este resultado demuestra que los modelos de aprendizaje automático son herramientas prometedoras para la depuración de datos en contextos de alta complejidad climática, como el sur andino peruano.

OMM (2021) publicó un informe que establece estándares globales de control de calidad, señalando que aproximadamente el 15% de los datos de estaciones remotas presentan errores atribuibles a la falta de mantenimiento o a condiciones ambientales extremas. El estudio se basó en un análisis comparativo de estaciones ubicadas en distintas regiones del mundo, incluyendo zonas polares, desérticas y montañosas.

Hallazgos y conclusiones del antecedente: La investigación evidenció que los errores se concentran principalmente en estaciones con difícil acceso, lo que retrasa su mantenimiento y reparación. Este hallazgo subraya la necesidad de sistemas de monitoreo automatizados que

reduzcan la dependencia de intervenciones presenciales, una estrategia especialmente útil para regiones aisladas de Puno.

Holm et al. (2022) llevaron a cabo un estudio en Noruega para comparar distintas técnicas de imputación de datos en series temporales con patrones estacionales complejos. Analizaron métodos como la interpolación lineal, Spline y redes neuronales Long Short-Term Memory (LSTM), utilizando datos de temperatura, humedad y precipitación recolectados durante 30 años.

Hallazgos y conclusiones del antecedente: Las redes neuronales LSTM demostraron un rendimiento superior en la reconstrucción de datos faltantes, especialmente en series con estacionalidad marcada y eventos extremos. Esta capacidad de adaptación a variaciones no lineales es relevante para la presente investigación, que busca optimizar la calidad de datos meteorológicos en entornos de alta variabilidad como el altiplano puneño.

Ki et al. (2025) aplicaron 18 algoritmos de aprendizaje automático para detectar outliers en datos climáticos de 40 años de estaciones en Burkina Faso. El enfoque multivariante se centró en variables clave como temperatura máxima y humedad relativa, utilizando la librería PyOD.

Hallazgos y conclusiones del antecedente: La detección automática permitió identificar anomalías críticas que hubieran distorsionado el análisis climático. Este resultado evidencia que los algoritmos modernos son aplicables en contextos de infraestructura limitada, situación comparable a Puno.

El proyecto European Climate Assessment & Dataset (ECA&D, 2007) consolida observaciones diarias de más de 3,300 estaciones en Europa, integrando un control de calidad automatizado y pruebas de homogeneidad que clasifican las series en categorías como “OK”, “suspect” o “missing”.

Hallazgos y conclusiones del antecedente: Este sistema garantiza mayor

comparabilidad de los datos y permite el uso de redes dispersas de estaciones con confianza. Su metodología es extrapolable a regiones altoandinas, donde la densidad de estaciones es baja y los errores son frecuentes.

El Copernicus Climate Change Service (C3S, 2021) implementó un marco de aseguramiento de calidad de datos climáticos (EQC), estructurado en procesos continuos de monitoreo y mejora basados en iniciativas previas de la Unión Europea.

Hallazgos y conclusiones del antecedente: El marco EQC fortalece la trazabilidad y consistencia de los datos, asegurando su uso confiable en modelación global. Para contextos como Puno, esta estrategia sugiere la importancia de sistemas estandarizados con mejora continua.

Park et al. (2022) desarrollaron un modelo de redes neuronales MLP para la imputación de largos períodos de datos faltantes (meses o años) en registros ambientales como humedad del suelo y niveles freáticos.

Hallazgos y conclusiones del antecedente: El modelo superó a los métodos tradicionales de interpolación, demostrando que las arquitecturas neuronales son útiles para reconstruir series históricas fragmentadas. Esta evidencia es directamente aplicable a los registros incompletos de estaciones meteorológicas puneñas.

1.4.2. Antecedentes nacionales

En el contexto nacional, el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) publicó el manual técnico titulado “Control de calidad de datos de estaciones meteorológicas e hidrológicas automáticas en el centro de procesamiento de datos del SENAMHI”. Este documento establece lineamientos para la validación de datos meteorológicos y contempla variables como temperatura del aire, humedad relativa, precipitación y nivel de agua. Según el SENAMHI (2021), la validación de datos requiere la aplicación de pruebas estadísticas y comparativas para garantizar su fiabilidad.

Hallazgos y conclusiones del antecedente: El manual proporciona un marco metodológico que unifica los criterios de control de calidad a nivel nacional, asegurando la comparabilidad de los registros provenientes de distintas estaciones. Sin embargo, su aplicación enfrenta desafíos en zonas altoandinas como Puno, donde las condiciones geográficas y la escasez de infraestructura dificultan la implementación completa de los protocolos.

Kinouchi et al. (2019) analizaron la tendencia de calentamiento en los Andes tropicales, estimada en 0.1 °C por década, y señalaron que la modelización climática en estas zonas requiere resoluciones espaciales inferiores a 1 km, una característica que actualmente no está disponible en el 85% de las estaciones altoandinas.

Hallazgos y conclusiones del antecedente: La investigación evidenció que la falta de modelos de alta resolución limita la capacidad de capturar la variabilidad microclimática, afectando las proyecciones climáticas y la planificación agrícola en regiones como Puno.

Bennett et al. (2016) corroboraron que la escasez de datos validados en zonas por encima de los 4,000 m s.n.m. genera incertidumbres de hasta el 40% en proyecciones agrícolas, especialmente en cultivos sensibles a las variaciones térmicas.

Hallazgos y conclusiones del antecedente: Este antecedente subraya la importancia de contar con series de datos completas y depuradas para reducir el margen de error en la planificación agropecuaria, un aspecto crítico para la seguridad alimentaria en la región altiplánica.

En conjunto, estos estudios revelan tres brechas críticas que afectan la región de Puno: la limitada adaptación de modelos globales a la variabilidad microclimática andina, la ausencia de algoritmos específicos para corregir errores en termómetros de mercurio y la desarticulación entre protocolos institucionales (SENAMHI) y las capacidades técnicas locales.

Entre los proyectos realizados en la Sierra peruana se tienen:

El SENAMHI (2022), en su informe anual, identificó que el 25% de los datos de temperatura registrados en Puno requieren corrección debido a fallas en sensores o problemas de transmisión de datos. El estudio incluyó inspecciones técnicas en campo y la comparación de registros entre estaciones automáticas y convencionales.

Hallazgos y conclusiones del antecedente: Se determinó que la mayoría de las inconsistencias provienen de deficiencias en el mantenimiento preventivo y correctivo de los equipos, así como de la ausencia de protocolos de calibración periódica. Esto implica la necesidad de reforzar las capacidades técnicas regionales y la implementación de controles de calidad en tiempo real.

El SENAMHI (2021) aplicó métodos de media móvil para corregir series de temperatura en Cusco, con especial atención a estaciones situadas por encima de los 3,800 m s.n.m. La iniciativa buscó optimizar la consistencia de los datos para su uso en estudios de variabilidad climática y en la gestión de recursos hídricos.

Hallazgos y conclusiones del antecedente: Aunque la media móvil permitió suavizar las series y eliminar fluctuaciones anómalas, se observó que esta técnica subestima la amplitud térmica en zonas de alta montaña. Este resultado evidencia que las técnicas de corrección deben adaptarse a las particularidades climáticas y topográficas de la región.

Además, complementando la información de los proyectos anteriores, se conocen las siguientes investigaciones académicas:

Fernández et al. (2023) desarrollaron modelos híbridos que combinan el enfoque ARIMA con redes neuronales artificiales (RNA) para la corrección de datos meteorológicos con errores no lineales en regiones altoandinas. La investigación se aplicó a series históricas de temperatura y precipitación, incorporando validación cruzada y análisis de sensibilidad.

Hallazgos y conclusiones del antecedente: Los modelos híbridos lograron mejorar significativamente la precisión de las series, reduciendo la tasa de error en comparación con

modelos puramente estadísticos o neuronales. Este enfoque resulta relevante para la presente investigación, ya que combina robustez estadística y capacidad de modelado no lineal, aspectos clave para la depuración de datos meteorológicos en Puno.

1.5. Justificación de la investigación

La calidad de los datos meteorológicos es fundamental para la toma de decisiones en diversas áreas como la agricultura, la gestión del agua, la prevención de desastres naturales y el análisis del cambio climático. En la sierra del departamento de Puno, la temperatura del aire es un factor clave para la planificación agrícola y la evaluación del impacto de eventos climáticos extremos, como heladas y sequías. Sin embargo, la presencia de datos erróneos en los registros meteorológicos debido a fallos en sensores, errores en la transmisión de datos y condiciones ambientales adversas compromete la confiabilidad de la información utilizada en estos sectores.

Dada la importancia de contar con datos precisos y confiables, esta investigación se justifica desde los siguientes enfoques:

1.5.1. Justificación científica

El desarrollo de un modelo matemático para la detección y corrección de datos erróneos en registros de temperatura contribuirá a la mejora de los métodos de control de calidad en meteorología. Esta investigación aportará un enfoque basado en modelos matemáticos y estadísticos, permitiendo reducir la incertidumbre en los análisis climáticos. Además, la validación del modelo contribuirá al avance del conocimiento en el procesamiento de datos ambientales, lo que puede servir como referencia para futuras investigaciones en otras regiones con condiciones similares.

1.5.2. Justificación técnica

Actualmente, muchas estaciones meteorológicas utilizan métodos básicos de validación de datos que no son completamente efectivos para corregir errores sistemáticos o aleatorios en las mediciones. Un modelo matemático optimizado permitirá la detección automática de datos atípicos y la corrección de valores erróneos, lo que mejorará significativamente la calidad de los registros meteorológicos. La implementación de este modelo contribuirá a optimizar los procesos de monitoreo climático en la región.

1.5.3. Justificación socioeconómica

En Puno, sectores como la agricultura y la ganadería dependen en gran medida de condiciones climáticas precisas para planificar actividades productivas y mitigar riesgos climáticos. La mejora en la calidad de los datos meteorológicos permitirá a los agricultores y tomadores de decisiones contar con información más confiable, optimizando la gestión de cultivos y reduciendo pérdidas económicas causadas por eventos climáticos adversos mal pronosticados.

1.5.4. Justificación ambiental

El cambio climático ha generado variaciones significativas en las temperaturas y en otros parámetros meteorológicos, lo que hace aún más crucial contar con datos precisos para su estudio. Un modelo que mejore la calidad de los registros de temperatura permitirá obtener análisis más fiables sobre las tendencias climáticas en la región, lo que contribuirá a la formulación de políticas de adaptación y mitigación del cambio climático.

En el caso de Puno, una región altamente vulnerable, se ha registrado un aumento de 0.8 °C en la temperatura media entre 1980 y 2020 (INAIGEM, 2021, p. 12). Este tipo de información precisa resulta esencial para el diseño de modelos de adaptación al cambio climático.

1.5.5. Impacto en agricultura

En la región de Puno, aproximadamente el 80% de los agricultores dependen de pronósticos climáticos para el manejo de cultivos como la quinua y la papa (MINAGRI, 2023). La presencia de datos meteorológicos erróneos puede generar pérdidas significativas; por ejemplo, se han reportado "pérdidas del 30% en cosechas por heladas no pronosticadas" (García et al., 2023, p. 25).

La implementación de un modelo confiable permitiría mejorar los sistemas de alerta temprana, lo cual contribuiría a la protección de la seguridad alimentaria en la región. En este sentido, la investigación adquiere relevancia al abordar una problemática que afecta la calidad de los datos meteorológicos y, en consecuencia, su aplicabilidad en distintos sectores. Un modelo matemático robusto no solo beneficiará a los centros de monitoreo climático y a la comunidad científica, sino también a los sectores productivos y a la sociedad en general.

1.6. Limitaciones de la investigación

Las principales limitaciones del presente estudio se relacionan con el alcance temporal, la disponibilidad de datos y las variables consideradas. En primer lugar, el análisis se restringió a un periodo de 30 años debido a la ausencia de registros alternativos que permitan la validación en intervalos más amplios. Asimismo, no se incluyeron otras variables con fines comparativos de validación, ni los metadatos de las estaciones meteorológicas en estudio, lo que podría aportar información adicional sobre las condiciones y el historial operativo de cada punto de medición.

La investigación depende exclusivamente de los datos de temperatura del aire proporcionados por estaciones meteorológicas ubicadas en la sierra del departamento de Puno. Sin embargo, la disponibilidad de registros completos y de alta calidad constituye un desafío, ya que algunas estaciones presentan fallas en la recolección, inconsistencias en la transmisión o periodos con información incompleta. A ello se suman diversas variables ambientales y

técnicas que pueden influir en la calidad de los datos, como cambios bruscos en la temperatura, interferencias electromagnéticas, fallas en los sensores o mantenimiento inadecuado de los equipos, factores que pueden introducir sesgos y requerir ajustes adicionales en el modelo propuesto.

Finalmente, si bien el modelo será validado con datos reales de la sierra de Puno, su aplicabilidad a otras regiones con diferentes características climáticas y geográficas podría requerir adaptaciones específicas. La extrapolación de resultados dependerá de la similitud en los patrones de error y en la estructura de los datos, por lo que se recomienda realizar pruebas adicionales antes de su implementación en otros contextos.

1.7. Objetivos

1.7.1. Objetivo general

Determinar en qué medida el modelo matemático, basado en técnicas de Machine Learning, estadísticas y geoespaciales, mejora el control de calidad para la corrección y relleno de datos erróneos de temperatura del aire en las estaciones meteorológicas de la Sierra de Puno.

1.7.2. Objetivos específicos

- Determinar en qué medida el modelo matemático, basado en técnicas de inteligencia artificial (Random Forest + Gradient Boosting), estadísticas y geoespaciales mejora la precisión en el control de calidad, evaluando la reducción de la incertidumbre en las series de temperatura de las estaciones de Puno mediante métricas estadísticas (desviación estándar) y pruebas de significancia estadística ($p < .05$).
- Determinar en qué medida el modelo matemático basado en técnicas geoespaciales mejora la confiabilidad en el control de calidad, evaluando el aumento en la consistencia espacial de los datos de temperatura mediante correlaciones significativas (% r Rellenado > % r Original, $r > .9$) entre estaciones vecinas.

1.8. Hipótesis

1.8.1. Hipótesis general

El modelo matemático, basado en inteligencia artificial (Random Forest + Gradient Boosting), técnicas estadísticas y geoespaciales, mejora significativamente la precisión ($p < .05$) y la confiabilidad (% r Rellenado > % r Original) de los datos de temperatura en las estaciones meteorológicas de la Sierra de Puno.

1.8.2. Hipótesis específicas

- El modelo matemático, basado en inteligencia artificial (Random Forest + Gradient Boosting), técnicas estadísticas, reduce significativamente ($p < .05$) la incertidumbre en las series de temperatura de las estaciones de Puno, mejorando la precisión en el control de calidad.
- El modelo matemático basado en técnicas geoespaciales incrementa significativamente la consistencia espacial de los datos de temperatura, evidenciada por una correlación mejorada (% r Rellenado > % r Original, $r > .9$) entre estaciones vecinas tras la corrección, mejorando la confiabilidad en el control de calidad.

II. MARCO TEÓRICO

2.1. Base teórica

2.1.1. Errores en mediciones meteorológicas

El registro de datos meteorológicos constituye la base de cualquier análisis climático, hidrológico y ambiental. La confiabilidad de dichos datos depende directamente de la precisión de los instrumentos utilizados, de los procedimientos de instalación y calibración, así como de las condiciones ambientales en las que operan las estaciones. En regiones de alta complejidad geográfica como la Sierra del Perú, las dificultades técnicas se potencian debido a factores como la altitud, la variabilidad microclimática y la limitada infraestructura para el mantenimiento de equipos. Por ello, el estudio de los errores en mediciones meteorológicas se vuelve crucial para implementar mecanismos de control de calidad y asegurar que los registros de temperatura del aire sean representativos y útiles en la modelización climática y en la toma de decisiones vinculadas a la gestión de riesgos y la planificación agrícola.

2.1.2. Tipos de errores en mediciones

La medición de la temperatura del aire en estaciones meteorológicas está sujeta a múltiples fuentes de error que afectan la calidad y confiabilidad de los datos. Estos errores se clasifican generalmente en sistemáticos y aleatorios. Los errores sistemáticos están asociados a fallas recurrentes en los instrumentos o en los procedimientos de calibración, generando sesgos en los registros que pueden acumularse a lo largo del tiempo (Chatfield, 2004). Por su parte, los errores aleatorios corresponden a variaciones impredecibles en las observaciones debidas a factores externos como cambios bruscos de viento, humedad o interferencias eléctricas (Wilks, 2019).

2.1.3. Impacto de los errores en la calidad de datos

Los errores en las mediciones meteorológicas generan distorsiones que afectan la validez de los análisis climáticos, especialmente en regiones sensibles como los Andes. Una desviación de apenas 0.2 °C en las series de temperatura puede producir sesgos significativos en la detección de tendencias de calentamiento a largo plazo (Kinouchi et al., 2019). Asimismo, la Organización Meteorológica Mundial (OMM, 2017) ha señalado que la calidad de los datos es determinante para la elaboración de modelos climáticos, la planificación agrícola y la gestión de riesgos.

2.1.4. Normas internacionales y protocolos nacionales

La OMM (2018a, 2021) establece estándares internacionales para la instalación, calibración y operación de sensores meteorológicos, con el objetivo de garantizar la trazabilidad y la homogeneidad de los registros. En el caso peruano, el SENAMHI (2021) publicó el *Manual de control de calidad de datos de estaciones meteorológicas automáticas*, en el cual se detallan pruebas de validación estadística y comparativa para variables como temperatura, humedad y precipitación. Sin embargo, la aplicación de estos protocolos enfrenta desafíos en la Sierra, donde las condiciones extremas y la limitada infraestructura dificultan la implementación de mantenimientos periódicos.

2.1.5. Series temporales en climatología

El análisis de series temporales constituye uno de los pilares fundamentales de la climatología moderna, ya que permite estudiar la evolución de variables como la temperatura, la precipitación o la humedad a lo largo del tiempo. A través de estas series es posible identificar tendencias, ciclos, anomalías y cambios estructurales en el clima de una región determinada. En zonas altoandinas como la Sierra de Puno, la importancia de este enfoque es aún mayor, debido a que la alta variabilidad microclimática exige registros confiables y depurados que reflejen la realidad de los fenómenos locales (Wilks, 2019). En este contexto,

las series temporales se convierten en herramientas esenciales para la detección de errores en los datos meteorológicos y para la aplicación de modelos matemáticos que mejoren la calidad de la información disponible.

2.1.6. Definición y características de las series temporales

Una serie temporal se define como una sucesión de observaciones de una variable registradas en intervalos de tiempo equidistantes (Chatfield, 2004). En climatología, estas series permiten analizar la persistencia de patrones, la ocurrencia de anomalías y la variación estacional de las condiciones meteorológicas. Sus principales características incluyen la tendencia (aumento o disminución de largo plazo), la estacionalidad (repetición de patrones en intervalos regulares), la ciclicidad (fluctuaciones no periódicas) y la aleatoriedad (ruido asociado a factores impredecibles). La correcta identificación de estas propiedades es fundamental para aplicar técnicas estadísticas de filtrado y modelización.

2.1.7. Aplicación de series temporales en climatología andina

En regiones altoandinas, las series temporales han demostrado ser esenciales para comprender fenómenos como el retroceso glaciar, la variabilidad de las precipitaciones o los cambios de temperatura asociados al cambio climático (Vuille et al., 2018). Sin embargo, la baja densidad de estaciones meteorológicas y la presencia de errores en los registros limitan su utilidad. Kinouchi et al. (2019) señalan que en los Andes tropicales, el 85% de las estaciones no posee modelos de alta resolución espacial, lo cual restringe la detección de microvariaciones térmicas críticas para la agricultura de subsistencia.

2.1.8. Métodos estadísticos para la detección de anomalías

El análisis de series temporales permite detectar anomalías o valores atípicos que pueden deberse a errores de medición o a fenómenos meteorológicos extremos. Entre las técnicas más utilizadas se encuentran el análisis de media móvil, los filtros de Kalman y los modelos autoregresivos integrados de media móvil (ARIMA) (Box et al., 2015). Estas

herramientas facilitan la identificación de patrones anómalos y la corrección de series interrumpidas o incompletas, contribuyendo así al mejoramiento del control de calidad de los datos climáticos en la Sierra del Perú.

2.1.9. Modelos matemáticos aplicados a datos climáticos

Los modelos matemáticos constituyen herramientas esenciales en el estudio del clima, ya que permiten representar, analizar y predecir el comportamiento de variables meteorológicas a partir de datos observados. En el caso de la temperatura del aire, los modelos estadísticos y deterministas cumplen un rol fundamental en la detección de errores, la reconstrucción de series incompletas y la generación de escenarios futuros. Según Wilks (2011), la aplicación de modelos matemáticos en climatología busca reducir la incertidumbre asociada a los datos y proporcionar bases sólidas para la toma de decisiones en ámbitos como la agricultura, la gestión de recursos hídricos y la planificación territorial. En la región altoandina del Perú, donde los registros suelen presentar vacíos y anomalías, estos modelos adquieren especial relevancia, pues permiten mejorar la calidad y confiabilidad de la información meteorológica disponible.

2.1.10. Modelos de regresión en climatología

Los modelos de regresión constituyen uno de los enfoques más utilizados para explicar la relación entre variables meteorológicas y detectar patrones en series temporales. La regresión lineal simple y múltiple ha sido aplicada para correlacionar la temperatura con factores como la altitud, la latitud o la radiación solar (Montgomery et al., 2012). En el contexto andino, estos modelos han permitido estimar la temperatura en zonas sin estaciones meteorológicas mediante interpolación basada en estaciones cercanas (Buytaert et al., 2006). Sin embargo, presentan limitaciones en la captura de fenómenos no lineales, por lo que suelen complementarse con técnicas más avanzadas.

2.1.11. Modelos ARIMA y series estocásticas

El modelo autoregresivo integrado de media móvil (ARIMA) constituye una técnica estadística de referencia en la modelización de series climáticas. Este modelo permite capturar la dependencia temporal de los datos y predecir valores futuros a partir de las observaciones pasadas (Box et al., 2015). En climatología, el ARIMA ha sido utilizado para la proyección de temperaturas mensuales y para la detección de anomalías en series largas de datos. En el Perú, estudios como el de Fernández et al. (2023) han combinado ARIMA con redes neuronales artificiales, logrando reducir los errores asociados a datos incompletos y mejorando la precisión de las series temporales en regiones altoandinas.

2.1.12. Interpolación espacial y temporal

La interpolación matemática es clave para completar datos faltantes en registros meteorológicos. Métodos como el inverso de la distancia ponderada (IDW), el kriging y el spline han sido empleados para estimar valores de temperatura en áreas sin observaciones directas (Goovaerts, 1997). En la Sierra de Puno, donde muchas estaciones presentan vacíos, estas técnicas permiten generar superficies continuas de temperatura con relativa precisión. No obstante, su eficacia depende de la densidad de estaciones y de la correcta selección de parámetros, lo que plantea retos en regiones con infraestructura meteorológica limitada (Bennett et al., 2016).

2.1.13. Detección de outliers en datos climáticos

La presencia de outliers o valores atípicos es frecuente en series de datos meteorológicos debido a errores de medición, transmisión o condiciones extremas. La detección de estos valores puede realizarse mediante métodos estadísticos como el análisis de desviaciones estándar, pruebas de Grubbs o técnicas de boxplot (Barnett y Lewis, 1994). En climatología, la identificación temprana de outliers es crucial para evitar sesgos en los modelos predictivos y garantizar la calidad de la información. Estudios aplicados en la región andina

han demostrado que la detección automática de valores extremos mejora significativamente la confiabilidad de las proyecciones climáticas (SENAMHI, 2021a).

2.1.14. *Aprendizaje automático y control de calidad de datos climáticos*

El avance de la inteligencia artificial y del aprendizaje automático (machine learning) ha abierto nuevas posibilidades para el control de calidad de datos meteorológicos, especialmente en contextos donde los métodos estadísticos tradicionales presentan limitaciones. Estas técnicas permiten identificar patrones no lineales, clasificar anomalías y realizar predicciones más precisas a partir de grandes volúmenes de información climática. En regiones como los Andes del Perú, donde los registros suelen ser incompletos o presentan errores de calibración, el aprendizaje automático constituye una alternativa prometedora para mejorar la fiabilidad de las series temporales de temperatura y otros parámetros. De acuerdo con Reichstein et al. (2019), la aplicación de algoritmos de machine learning en climatología ha incrementado la capacidad de detección de outliers y ha optimizado la reconstrucción de datos en sistemas complejos.

2.1.15. Algoritmos supervisados aplicados a datos climáticos

Entre los algoritmos supervisados más utilizados en climatología se encuentran Random Forest y Gradient Boosting. Estos métodos construyen modelos predictivos a partir de datos etiquetados, lo que permite detectar inconsistencias en series meteorológicas y corregir valores erróneos. Breiman (2001) sostiene que Random Forest es particularmente robusto frente a ruido y datos faltantes, lo que lo convierte en una herramienta valiosa para estaciones meteorológicas de la Sierra peruana que registran interrupciones frecuentes. Asimismo, Chen y Guestrin (2016) demuestran que el algoritmo XGBoost, una versión optimizada de Gradient Boosting, mejora la precisión en tareas de predicción climática y en la detección de anomalías en datos de temperatura.

2.1.16. Algoritmos no supervisados y detección de anomalías

Los algoritmos no supervisados, como K-means, DBSCAN o autoencoders, son empleados para identificar estructuras ocultas en los datos sin necesidad de etiquetas previas. En climatología, han sido aplicados para la clasificación de estaciones meteorológicas en función de su comportamiento térmico y para la detección automática de valores atípicos. Chandola et al. (2009) destacan que estos métodos son especialmente útiles en entornos con grandes volúmenes de datos y ausencia de metadatos confiables. En el caso del altiplano peruano, estas técnicas permiten distinguir entre anomalías debidas a errores instrumentales y aquellas causadas por fenómenos meteorológicos extremos.

2.1.17. Redes neuronales aplicadas a la climatología andina

Las redes neuronales artificiales (RNA) han demostrado gran potencial para modelar relaciones complejas en datos meteorológicos, gracias a su capacidad de aproximar funciones no lineales. Según Goodfellow et al. (2016), las RNA permiten reconstruir series temporales incompletas y generar proyecciones de temperatura con mayor precisión que los modelos lineales tradicionales. En el contexto peruano, Fernández et al. (2023) aplicaron modelos híbridos que combinan ARIMA con RNA para la corrección de series de temperatura en zonas altoandinas, obteniendo una reducción significativa en la tasa de error.

2.1.18. Comparación entre modelos estadísticos y de aprendizaje automático

Diversos estudios han mostrado que el aprendizaje automático supera en desempeño a los modelos estadísticos convencionales en el tratamiento de datos meteorológicos complejos. Sin embargo, también presenta desafíos, como la necesidad de grandes volúmenes de datos para entrenamiento y la dificultad en la interpretación de resultados (la llamada “caja negra”). Reichstein et al. (2019) sostienen que la combinación de ambos enfoques, estadísticos y de machine learning, ofrece un marco más robusto para el control de calidad de datos climáticos. En el caso de Puno, esta integración resulta esencial debido a la dispersión geográfica de las

estaciones, las limitaciones de infraestructura y la alta variabilidad microclimática.

2.1.19. Control de calidad de datos meteorológicos en el Perú

El control de calidad de datos meteorológicos constituye un componente esencial en la generación de información confiable para la gestión del clima y la planificación territorial. En el caso del Perú, la complejidad geográfica, la diversidad climática y la limitada cobertura de estaciones en regiones altoandinas han generado una serie de retos en la validación y corrección de registros. De acuerdo con la Organización Meteorológica Mundial (OMM, 2011), los procedimientos de control de calidad deben contemplar tanto verificaciones automáticas como inspecciones manuales, garantizando que los datos presenten coherencia interna y consistencia espacial y temporal. En este sentido, el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) ha elaborado protocolos específicos para el tratamiento de series meteorológicas que incluyen pruebas estadísticas, comparativas y análisis de homogeneidad.

2.1.20. Protocolos institucionales de control de calidad

SENAMHI (2021) publicó el manual técnico “Control de calidad de datos de estaciones meteorológicas e hidrológicas automáticas”, en el cual se establecen lineamientos para la validación de registros provenientes de estaciones automáticas y convencionales. El documento contempla variables como temperatura del aire, humedad relativa, precipitación y nivel de agua, e incluye procedimientos como detección de rangos válidos, pruebas de consistencia interna y comparación entre estaciones cercanas. Este marco metodológico ha permitido estandarizar los criterios de control de calidad a nivel nacional, asegurando la comparabilidad de registros. Sin embargo, enfrenta dificultades en regiones como Puno, donde la dispersión geográfica y la falta de infraestructura limitan la aplicación completa de los protocolos (SENAMHI, 2021b).

2.1.21. Limitaciones en regiones altoandinas

En la Sierra sur del Perú, particularmente en Puno, las estaciones meteorológicas se ubican a altitudes superiores a los 3,800 m s.n.m., lo que genera condiciones extremas que afectan el funcionamiento de sensores y equipos. Kinouchi et al. (2019) señalan que el 85% de las estaciones altoandinas carece de resoluciones espaciales menores a 1 km, lo que limita la capacidad de modelar la variabilidad microclimática. Asimismo, Bennett et al. (2016) evidencian que la escasez de datos validados por encima de los 4,000 m s.n.m. puede generar incertidumbres de hasta el 40% en proyecciones agrícolas, lo cual resulta crítico para la seguridad alimentaria de comunidades rurales dependientes de cultivos sensibles a variaciones térmicas.

2.1.22. Inconsistencias técnicas y necesidades de mantenimiento

Los registros meteorológicos en el Perú también enfrentan problemas relacionados con fallas en sensores, interrupciones en la transmisión de datos y ausencia de protocolos de calibración periódica. SENAMHI (2022), en su informe anual, identificó que aproximadamente el 25% de los datos de temperatura registrados en Puno requieren corrección debido a fallas técnicas. Este hallazgo evidencia la necesidad de reforzar el mantenimiento preventivo y correctivo de las estaciones, así como de implementar mecanismos de control de calidad en tiempo real para reducir los márgenes de error en la captura de datos.

2.1.23. Proyectos de corrección y mejora de datos en la sierra peruana

Diversas iniciativas académicas y técnicas han buscado mejorar la calidad de los datos meteorológicos en la Sierra. El SENAMHI (2021) aplicó métodos de media móvil para la corrección de series de temperatura en estaciones situadas por encima de los 3,800 m s.n.m., logrando suavizar fluctuaciones anómalas, aunque con la limitación de subestimar la amplitud térmica. Asimismo, Fernández et al. (2023) implementaron modelos híbridos ARIMA–RNA para corregir datos de temperatura y precipitación en la región altoandina, alcanzando una

reducción significativa de los errores. Estos resultados confirman la importancia de combinar protocolos institucionales con herramientas estadísticas y de machine learning adaptadas a las condiciones geográficas del Perú.

2.1.24. Brechas identificadas en el control de calidad en Puno

En conjunto, los estudios realizados en la región de Puno permiten identificar tres brechas críticas: (i) la limitada adaptación de modelos globales a la variabilidad microclimática de los Andes; (ii) la ausencia de algoritmos específicos para corregir errores en equipos tradicionales como termómetros de mercurio; y (iii) la desarticulación entre los protocolos normativos del SENAMHI y las capacidades técnicas locales para aplicarlos. Superar estas brechas requiere fortalecer la infraestructura meteorológica, capacitar al personal regional y fomentar el uso de tecnologías emergentes que garanticen registros confiables y consistentes a largo plazo.

2.2. Marco conceptual

2.2.1. Control de calidad de datos meteorológicos

2.2.1.1. Datos climáticos. Los datos climáticos constituyen la información cuantitativa obtenida mediante la medición de variables atmosféricas, tales como temperatura, humedad, presión y precipitación. Estos registros permiten analizar y comprender el comportamiento del clima en una región determinada, aportando insumos esenciales para estudios de variabilidad y cambio climático (Organización Meteorológica Mundial [OMM], 2018a).

2.2.1.2. Temperatura del aire. La temperatura del aire es una variable meteorológica fundamental, cuya medición adecuada resulta indispensable para evaluar las condiciones climáticas, los fenómenos extremos y los procesos ecológicos (OMM, 2018a). En regiones andinas, este parámetro puede verse influenciado por factores topográficos, la radiación solar y las condiciones atmosféricas locales, lo que introduce complejidad en la interpretación de los

registros (OMM, 2017).

2.2.1.3. Errores en mediciones. Los errores en mediciones se definen como las discrepancias entre el valor observado y el valor real de una variable. Estos pueden clasificarse en sistemáticos, constantes o predecibles, asociados por ejemplo a instrumentos mal calibrados y aleatorios, que son impredecibles y varían en el tiempo (Taylor, 1997). Asimismo, los registros climáticos suelen presentar valores atípicos o outliers, definidos como observaciones que difieren significativamente del resto de los datos, cuya detección y corrección es fundamental para evitar interpretaciones sesgadas (Wilks, 2011).

2.2.1.4. Datos erróneos de temperatura del aire. El control de calidad se puede llevar a cabo mediante métodos manuales o automáticos. En principio, todos los procedimientos de control de calidad necesarios se pueden aplicar de forma manual, aunque el tiempo necesario es normalmente excesivamente largo. El control de calidad no debe introducir un retraso apreciable, ya que los datos se tienen que transmitir en tiempo real para su uso operacional. El control de calidad en tiempo real en el punto de observación es, sin embargo, de importancia fundamental, puesto que muchos errores introducidos durante el proceso de observación no se pueden eliminar posteriormente.

Toda la gama de errores de observación puede dividirse en los tres grupos principales siguientes:

- Errores de los aparatos técnicos, incluidos los instrumentos;
- Errores inherentes a los procedimientos y métodos de observación; y
- Errores subjetivos aleatorios o sistemáticos por parte de los observadores y operadores colectores (La Organización Meteorológica Mundial - OMM N° 488,2010, p.VI –3).
- Control de calidad de datos climáticos. El control de calidad se basa en aplicar pruebas estadísticas y físicas para asegurar que los datos meteorológicos sean válidos, consistentes y confiables (Aguilar et al., 2003). Las técnicas incluyen pruebas de rango, validaciones

espaciales y temporales, y detección de discontinuidades. (véanse la **Tabla 27** y la **Figura 9**)

El control de calidad para los datos de las estaciones meteorológicas e hidrológicas automáticas debe seguir un proceso de evaluación y verificación mediante una serie de pruebas comparativas, estadísticas y otras, que permiten detectar, depurar, aislar datos inconsistentes, y cuando sea posible restituir los valores correctos de aquellos datos defectuosos. (véanse las figuras: **Figura 10**, **Figura 11**,**Figura 12**,**Figura 13**,**Figura 14** y **Figura 15**) Para ello, el instrumento de gestión orientado a dar confiabilidad a estos datos es el Manual Técnico de Control de Calidad de Datos de estaciones meteorológicas e hidrológicas automáticas en el centro de procesamiento de datos del Servicio Nacional de Meteorología e Hidrología del Perú, el mismo que contribuye para contar con una base de datos confiable, con la debida calidad y la oportunidad requerida (SENAMHI-Manual Técnico: MT-DRD-002,2021, p.4).

2.2.2. Modelo matemático

Un modelo matemático constituye una representación abstracta de un fenómeno a través de ecuaciones, algoritmos o funciones, cuya finalidad es describir, analizar, predecir o corregir comportamientos observados en los datos (Wilks, 2011). En meteorología, estos modelos resultan fundamentales para la detección de anomalías y la reconstrucción de series de datos incompletas o inconsistentes.

2.2.3. Series temporales climáticas

Las series temporales climáticas son conjuntos de observaciones ordenadas cronológicamente que representan la evolución de una variable en el tiempo. Su análisis permite identificar tendencias, estacionalidades y patrones recurrentes, constituyendo la base para modelos de predicción climática y para la validación de registros históricos (Chatfield, 2004). (véanse la **Figura 9**, la **Tabla 28** y la **Tabla 29**)

2.2.4. Estándares internacionales

La Organización Meteorológica Mundial ha establecido protocolos internacionales para el control de calidad de datos. Según sus directrices más recientes, en zonas montañosas se exige que el umbral de error en la medición de temperatura no supere los ±1.0 °C, dada la importancia de esta variable para la caracterización del clima y la toma de decisiones en contextos vulnerables (OMM, 2021, p. 30).

2.2.5. Errores en mediciones

Se define como las desviaciones o discrepancias entre el valor medido y el valor real o verdadero de una variable. Estos errores pueden ser sistemáticos (constantes o predecibles) o aleatorios (impredecibles y variables) (Taylor, 1997).

2.2.6. Errores en los datos meteorológicos

Los errores en registros climáticos pueden clasificarse en sistemáticos, aleatorios y atípicos. Los errores sistemáticos provienen de instrumentos mal calibrados o condiciones físicas constantes, mientras que los aleatorios se originan por causas impredecibles. Los valores atípicos o outliers son observaciones que difieren significativamente del resto de los datos (Wilks, 2011). Detectarlos y corregirlos es clave para evitar interpretaciones erróneas.

2.2.7. Estación meteorológica convencional

Estaciones meteorológicas equipadas con instrumentos meteorológicos manuales, que sirven para medir variables climáticas (IDEAM, 2019, p. 115).

2.2.8. Estación meteorológica automática (EMA)

Una estación meteorológica automática se define como una estación meteorológica en la que las observaciones se hacen y transmiten automáticamente (Organización Meteorológica Mundial - OMM N° 182, 1992, p.65).

2.2.9. Conceptos de random forest

Random forest es un algoritmo de aprendizaje automático supervisado basado en métodos de ensamble, desarrollado por Breiman (2001), que combina múltiples árboles de decisión para generar predicciones más robustas y precisas tanto en tareas de clasificación como de regresión. Este algoritmo construye un "bosque" de árboles de decisión utilizando la técnica de Bootstrap Aggregating (bagging), donde cada árbol se entrena con una muestra aleatoria de los datos originales y un subconjunto aleatorio de características (Hastie et al., 2009).

El fundamento teórico de random forest se basa en el principio de aprendizaje ensamblado, donde múltiples modelos débiles se combinan para crear un predictor fuerte, reduciendo simultáneamente la varianza sin incrementar significativamente el sesgo (Breiman, 2001). Esta estrategia aprovecha la diversidad entre los árboles individuales para mejorar la generalización del modelo y reducir el sobreajuste característico de los árboles de decisión individuales.

La predicción final se obtiene mediante agregación de las salidas individuales: votación mayoritaria para clasificación y promedio aritmético para regresión (James et al., 2013). Esta combinación de predictores débiles resulta en un modelo con mayor estabilidad y capacidad para capturar relaciones no lineales complejas entre variables, características que lo hacen especialmente útil para el análisis de datos meteorológicos donde las relaciones entre variables pueden ser altamente complejas y no lineales (Biau y Scornet, 2016).

2.2.9.1. Fundamentos matemáticos. La clave de su eficacia radica en dos principios fundamentales:

A. Bagging (Bootstrap aggregating). Cada árbol en el bosque se entrena con una muestra aleatoria (con reemplazo) de los datos de entrenamiento originales. Esto asegura que cada árbol vea un subconjunto ligeramente diferente de los datos, lo que introduce diversidad.

B. Aleatoriedad de característica. En cada nodo de un árbol de decisión, solo se considera un subconjunto aleatorio de las características disponibles para encontrar la mejor división. Esto reduce la correlación entre los árboles y ayuda a prevenir el sobreajuste.

Al promediar las predicciones de todos los árboles (para regresión) o mediante votación mayoritaria (para clasificación), random forest mitiga el problema del sobreajuste que a menudo afecta a los árboles de decisión individuales, mejorando la generalización del modelo y su capacidad para manejar datos ruidosos o con valores atípicos.

En cada división del árbol, solo se considera un subconjunto aleatorio de variables predictoras (√n características, donde *n* es el total de variables).

C. Predicción por mayoría o promedio. Clasificación: Votación mayoritaria entre árboles. Regresión: Promedio de las predicciones individuales.

D. Método de construcción. El algoritmo construye múltiples árboles de decisión utilizando el método de bagging (Bootstrap aggregating). La idea general del método de bagging es que una combinación de modelos de aprendizaje incrementa el resultado general. Cada árbol se entrena con una muestra aleatoria de los datos de entrenamiento y utiliza un subconjunto aleatorio de características para cada división, lo que introduce la diversidad necesaria en el ensamble.

E. **Proceso de predicción.** En el proceso de predicción, random forest combina las salidas de todos los árboles individuales. Para tareas de clasificación, se utiliza votación mayoritaria, mientras que para regresión se calcula el promedio de las predicciones de todos los árboles. Esta agregación de múltiples predictores débiles resulta en un modelo más robusto y con menor varianza.

F. **Aplicación en meteorología.** En el campo de la meteorología, Random Forest ha demostrado ser eficaz en la imputación de datos faltantes, detección de anomalías y predicción de variables climáticas como temperatura, precipitación y radiación solar. Gracias a su

arquitectura no paramétrica, puede capturar relaciones complejas y no lineales entre las variables meteorológicas (Tang et al., 2018).

2.2.10. Fundamentos de gradient boosting

Gradient boosting es una técnica de aprendizaje de máquina basada en el principio de mejoras sucesivas por gradiente, introducida por Friedman (2001). A diferencia de random forest, que construye árboles en paralelo, gradient boosting construye árboles de decisión de forma secuencial, donde cada nuevo árbol corrige los errores cometidos por el conjunto anterior.

El proceso inicia con un modelo base simple (por ejemplo, un árbol pequeño), al cual se le suman iterativamente nuevos modelos entrenados sobre los residuos del modelo anterior. El objetivo es minimizar una función de pérdida (como el error cuadrático medio en regresión), utilizando el descenso por gradiente para ajustar el modelo en cada iteración.

Gradient Boosting es altamente flexible, ya que permite optimizar diversas funciones de pérdida y aplicar técnicas de regularización como shrinkage (tasa de aprendizaje) y submuestreo estocástico, lo cual mejora la generalización del modelo y evita el sobreajuste.

Su principio fundamental es el siguiente:

2.2.10.1. Aprendizaje secuencial. El algoritmo entrena los árboles de forma iterativa. Cada nuevo árbol se construye para corregir los errores (residuos) del modelo combinado de los árboles anteriores. Es decir, cada árbol sucesivo se enfoca en las instancias que el modelo previo predijo incorrectamente o con un error significativo.

2.2.10.2. Optimización por gradiente. El término “gradiente” se refiere a que el algoritmo utiliza un enfoque de descenso de gradiente para minimizar la función de pérdida. En cada iteración, el nuevo árbol se ajusta a los residuos negativos (el gradiente negativo de la función de pérdida) del modelo actual, lo que permite que el modelo aprenda de manera más precisa de los errores previos.

Implementaciones populares de Gradient Boosting incluyen XGBoost, LightGBM y CatBoost, que han demostrado un rendimiento excepcional en una amplia variedad de problemas de aprendizaje automático, especialmente en datos tabulares. Gradient Boosting es conocido por su alta precisión y su capacidad para manejar relaciones complejas en los datos, aunque puede ser más propenso al sobreajuste si no se ajustan correctamente sus hiperparámetros.

2.2.10.3. Aplicación en problemas de regresión. El algoritmo Gradient Boosting funciona para datos tabulares con un conjunto de características (X) y un objetivo (y). Como otros algoritmos de aprendizaje automático, el objetivo es aprender lo suficiente de los datos de entrenamiento para generalizar bien a puntos de datos no vistos. Esta característica es fundamental para su aplicación en la predicción y corrección de datos de temperatura, donde se busca establecer relaciones complejas entre variables meteorológicas.

En aplicaciones meteorológicas, Gradient Boosting ha sido ampliamente utilizado para completar series de datos climáticos incompletos, especialmente cuando las relaciones entre variables son no lineales o afectadas por factores geográficos y estacionales. Algoritmos derivados como XGBoost y LightGBM han mejorado notablemente el rendimiento y la eficiencia de este método en grandes volúmenes de datos ambientales (Salcedo et al., 2021; Zhang et al., 2022).

2.2.10.4. Métodos de ensamble: marco teórico común**.** Tanto Random Forest como Gradient Boosting pertenecen a la familia de métodos de ensamble. Los métodos de ensamble combinan las predicciones de varios estimadores bases construidos con un algoritmo de aprendizaje dado para mejorar la generalización/robustez sobre un solo estimador. Dos ejemplos muy famosos de métodos de ensamble son los árboles potenciados por gradiente y los bosques aleatorios.

2.2.11. Técnica de validación cruzada para los modelos forest y gradiente boosting

La validación cruzada es una técnica estadística utilizada para evaluar el rendimiento de un modelo predictivo y su capacidad de generalización a datos no vistos (James et al., 2021). Consiste en dividir el conjunto de datos en subconjuntos (folds) para entrenar y evaluar el modelo de manera iterativa, garantizando que cada dato se utilice tanto para entrenamiento como para prueba. En este caso, se emplea la validación cruzada de tipo K-Fold con k=5, lo que implica dividir los datos en 5 subconjuntos, utilizando 4 para entrenamiento y 1 para prueba en cada iteración.

El objetivo principal de la validación cruzada en este estudio es evaluar la robustez del modelo híbrido (que combina Random Forest y Gradient Boosting) para corregir y rellenar datos de temperatura media (Tmed) en las estaciones meteorológicas, asegurando que las predicciones sean consistentes y confiables. Además, se utilizan métricas estadísticas como el Error Absoluto Medio (MAE), el Error Cuadrático Medio de la Raíz (RMSE) y el Coeficiente de Determinación (R^2) para cuantificar el rendimiento del modelo, y la prueba de Friedman para evaluar diferencias significativas entre los folds.

2.2.12. Fórmulas de las métricas utilizadas

2.2.12.1. Error absoluto medio (MAE). El MAE mide el promedio de las diferencias absolutas entre los valores reales y predichos, proporcionando una medida de la magnitud promedio del error sin considerar su dirección (James et al., 2021). Se calcula como:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (1)$$

Donde:

yi: Valor real de la temperatura media.

y^i: Valor predicho por el modelo.

n: Número de observaciones.

2.2.12.2. Error cuadrático medio de la raíz (RMSE). El RMSE mide la raíz cuadrada del promedio de los errores al cuadrado, penalizando más los errores grandes y proporcionando una medida de dispersión de las predicciones (Montgomery et al., 2021). Se calcula como:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (2)$$

- **Coeficiente de determinación (R^2).** El R^2 indica la proporción de la varianza de la variable dependiente explicada por el modelo. Un valor cercano a 1 indica un buen ajuste del modelo (Kutner et al., 2005). Se calcula como:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (3)$$

Donde: $\bar{y}$: Media de los valores reales.

- **Test de friedman.** El test de Friedman es un método no paramétrico utilizado para comparar las distribuciones de las métricas (MAE, RMSE, R^2) entre los k folds, determinando si existen diferencias significativas en el rendimiento del modelo (Conover, 1999). La fórmula del estadístico de Friedman es:

$$\chi_F^2 = \frac{12N}{k(k+1)}\sum_{j=1}^{k} R_j^2 - 3N(k+1) \qquad (4)$$

Donde: N: Número de observaciones (en este caso, número de métricas evaluadas).

k: Número de folds.

Rj: Suma de los rangos asignados al fold j. Un p-valor menor a 0.05 indica diferencias significativas entre los folds.

2.3. Bases legales

2.3.1. Organización meteorológica mundial (OMM)

La OMM establece estándares internacionales para la observación, registro y control de calidad de datos meteorológicos mediante la WMO-No. 8 Guide to Meteorological Instruments and Methods of Observation (OMM, 2018a). Esta normativa resulta fundamental para garantizar la interoperabilidad y la confiabilidad de los datos climáticos a nivel global.

2.3.2. Servicio nacional de meteorología e hidrología del Perú (SENAMHI)

En el contexto nacional, el SENAMHI es el organismo encargado de la gestión y supervisión de las estaciones meteorológicas en el Perú. La institución define protocolos de instalación, mantenimiento y control de calidad de las estaciones, los cuales se encuentran alineados con las recomendaciones de la OMM (SENAMHI, 2020).

2.3.3. Ley N° 28303- Ley del sistema nacional de ciencia, tecnología e innovación tecnológica.

Esta norma promueve el desarrollo y la aplicación de tecnologías orientadas a la mejora de la gestión ambiental y climática, destacando la importancia de generar y controlar datos científicos confiables (Ley N.° 28303. Ley Marco de Ciencia, Tecnología e Innovación Tecnológica (27 de julio de 2004). Diario Oficial El Peruano).

2.3.4. Política nacional de cambio climático (PNCC) del Perú

La PNCC establece lineamientos para la gestión de información climática destinada a respaldar las acciones de adaptación y mitigación frente al cambio climático. Asimismo, resalta la importancia de garantizar la calidad y accesibilidad de los datos meteorológicos (MINAM, 2015).

III. MÉTODO

3.1. Tipo de investigación

El presente estudio se enmarca en una investigación aplicada, de enfoque cuantitativo, con diseño no experimental y alcance descriptivo-explicativo.

3.1.1. Por su naturaleza: investigación aplicada

De acuerdo con la clasificación de Hernández Sampieri et al. (2014), la investigación aplicada tiene como propósito resolver problemas concretos en contextos específicos. En este caso, se busca atender la problemática de la detección y corrección de errores en los registros de temperatura del aire en estaciones meteorológicas de la sierra del departamento de Puno. La finalidad es desarrollar un modelo matemático que pueda ser implementado por instituciones como el SENAMHI, contribuyendo a mejorar la calidad de los datos climáticos y a fortalecer la toma de decisiones en la gestión del clima y los recursos naturales.

3.1.2. Por su profundidad: diseño descriptivo-explicativo

El diseño descriptivo-explicativo a nivel metodológico, el estudio abarca un diseño descriptivo y explicativo:

Descriptivo: Metodológicamente, el estudio presenta un diseño descriptivo y explicativo. En la dimensión descriptiva, se identificaron y caracterizaron las fuentes y patrones de error en los registros de temperatura del aire, especificando su frecuencia, tipo y magnitud en las estaciones meteorológicas de la región (Hernández Sampieri et al., 2014). Ello incluyó la clasificación de los tipos de fallas, como outliers, sesgos y datos faltantes. En la dimensión explicativa, se establecieron relaciones de causalidad entre las condiciones de los registros meteorológicos y la efectividad del modelo propuesto, evaluando el grado en que este contribuye a mejorar la precisión y confiabilidad de la información climática. De esta manera, fue posible analizar cómo los errores instrumentales o ambientales afectan los datos y cómo el modelo puede corregirlos. La combinación de ambos enfoques permitió construir un modelo

matemático fundamentado en evidencia empírica y adaptado a las condiciones particulares de la sierra de Puno.

3.1.3. Por su metodología: enfoque cuantitativo y diseño no experimental

El enfoque adoptado fue cuantitativo, dado que la investigación se basó en la recolección y análisis numérico de datos de temperatura del aire. Se aplicaron técnicas estadísticas y modelos matemáticos para identificar, corregir y validar los registros erróneos, lo que permitió establecer relaciones entre variables, medir el impacto del modelo y generalizar resultados a partir de muestras representativas (Creswell, 2014).

En cuanto a su diseño, la investigación fue de tipo no experimental, ya que no se manipularon deliberadamente las variables independientes. El estudio se desarrolló a partir de datos históricos registrados por estaciones meteorológicas, los cuales fueron analizados de forma retrospectiva para identificar errores y aplicar los mecanismos de corrección.

3.2. Población y muestra

3.2.1. Población

La población estuvo conformada por los registros históricos de temperatura del aire recolectados por estaciones meteorológicas ubicadas en la región sierra del departamento de Puno, Perú (véanse la **Tabla 23**, **Tabla 24** y la **Figura 5**), de acuerdo con la clasificación climática del SENAMHI (véanse la **Figura 7**). Esta región presenta una topografía compleja y condiciones climáticas particulares que afectan la calidad de los datos registrados. Según la definición de Pineda et al. (1994), la población constituye el universo sobre el cual se desea obtener información; en este caso, corresponde a los registros de temperatura del aire en las estaciones convencionales y automáticas del SENAMHI.

Los criterios considerados para definir la población fueron los siguientes:

- Periodo histórico de 24 años, desde 2001 hasta 2024.
- Estaciones ubicadas a una altitud mayor a 3,800 msnm.

- Cobertura de todas las provincias de la sierra de Puno, con excepción de Carabaya y Sandia, por su influencia climática de selva.
- Estaciones administradas y gestionadas por el SENAMHI, cuyos datos fueron obtenidos desde su plataforma oficial (https://www.senamhi.gob.pe/?&p=estaciones).

3.2.2. Muestra

La muestra estuvo conformada por los datos de temperatura diaria del aire registrados durante 24 años (2001–2024) por un total de 22 estaciones meteorológicas del SENAMHI. La selección se realizó mediante un muestreo no probabilístico por criterios técnicos y geográficos, considerando los siguientes aspectos:

3.2.2.1. Representatividad geográfica. Zonas altitudinales (>3,800 msnm). Zona climática de Sierra en base al mapa Climático del SENAMHI donde se incluye todas las provincias de Puno, a excepto las provincias de Carabaya y Sandia dado que climáticamente tienen influencia de clima de la región selva. (véanse la **Figura 4** y **Figura 6**)

3.2.2.2. Disponibilidad de datos. Series completas (>75% de datos) para el período 2001–2024, con excepción se seleccionó un mínimo de 4 estaciones automáticas con menos del 75% de disponibilidad de datos y a su vez que la mencionada estación este ubicada cerca de otra estación meteorológica convencional, esta excepción se justifica para validar el relleno de datos faltantes de largo periodo.

3.2.2.3. Tipología de estaciones. Incluye 4 Estaciones Automáticas y 18 Estaciones Convencionales (véanse la **Tabla 24**).

3.2.2.4. Criterio de validación. Se priorizo estaciones automáticas que estén ubicados en el mismo lugar que su correspondiente estación convencional, con la finalidad de tener una mejor validación estadística de los datos.

3.3. Operacionalización de variables

Tabla 1.

Operacionalización de variable independiente

Definición conceptual	Definición operacional	Dimensión	Indicador	Instrumento	Escala valorativa
Conjunto de algoritmos y técnicas estadísticas y geoespaciales diseñados para detectar, corregir y rellenar datos erróneos de temperatura del aire, optimizando la precisión y confiabilidad de las series temporales en estaciones meteorológicas	Implementación de un modelo híbrido que combina Random Forest y Gradient Boosting para la detección de anomalías, corrección de errores sistemáticos y relleno de datos faltantes en series de temperatura media diaria (2001-2024) de 22 estaciones meteorológicas en la Sierra de Puno, con un componente geoespacial para garantizar consistencia espacial.	**X1:** Eficacia del algoritmo	Precisión del modelo matemático: (Validación Cruzada(k-fold), Test Friedman)	Los datos fueron recolectados desde la plataforma oficial del SENAMHI, los datos recolectados se almacenaron en un formado de registro de datos diarios del SENAMHI	Razón (numérica continua).
			Parámetros optimizados del modelo matemático (ej.: hiperparámetros en Random Forest/Boosting).		Razón (numérica discreta) y nominal(categóricos)

Nota. Elaboración propia

Tabla 2.

Operacionalización de variable dependiente

Definición conceptual	Definición operacional	Dimensión	Indicador	Instrumento	Escala valorativa
Proceso de evaluación y mejora de la calidad de los datos de temperatura del aire mediante la reducción de incertidumbre y el aumento de la consistencia espacial, asegurando que los datos sean precisos y confiables para análisis climáticos.	Medida de la calidad de los datos de temperatura media diaria corregidos y rellenados por el modelo matemático, evaluada a través de la incertidumbre (*p-valor* < .05, desviación estándar) y la consistencia espacial (correlación *r* > .9 entre estaciones vecinas).	**Y1:** Precisión: (Reducción de Incertidumbre)	Reducción de la desviación estándar de las series antes/después de la corrección.	Los datos fueron recolectados desde la plataforma oficial del SENAMHI, los datos recolectados se almacenaron en un formado de registro de datos diarios del SENAMHI	Razón (numérica continua).
		Y2: Confiabilidad (consistencia espacial)	Correlación de Pearson (% r post-corrección > % r original) entre estaciones vecinas.		Razón (numérica continua).

Nota. Elaboración propia

3.3.1 Variable independiente (X): modelo matemático

- Dimensión D1: Eficacia del algoritmo
- Indicadores:
  - ✓ I1: Precisión del modelo matemático (%)
  - ✓ I2: Parámetros optimizados del modelo matemático

(véanse la **Tabla 1**)

3.3.2 Variable dependiente (Y): control de calidad

- Dimensión D2: Precisión (Reducción de Incertidumbre)
- Indicadores:
  - ✓ I3: Reducción de la desviación estándar de las series antes/después de la corrección
- Dimensión D3: Confiabilidad (Consistencia espacial)
- Indicadores:
  - ✓ I4: Variación del coeficiente de correlación de Pearson (%r post-corrección > %r original) entre estaciones vecinas

(véanse la **Tabla 2**)

3.4. Instrumentos

Los instrumentos utilizados en esta investigación para la adquisición de datos se basaron en el formato de registro de datos diarios (véanse las tablas: **Tabla 9**, **Tabla 10** , **Tabla 11** y la **Tabla 12**). La información fue descargada desde la plataforma oficial del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI): https://www.senamhi.gob.pe/?&p=estaciones.

El diseño del formato de registro de datos diarios se elaboró tomando como referencia el modelo de descarga proporcionado por dicha plataforma. Se recopilaron datos de estaciones meteorológicas ubicadas en la sierra del departamento de Puno, correspondientes al periodo

2001–2024, considerando únicamente aquellas con altitudes superiores a 3,800 m s. n. m.

Para verificar la confiabilidad de los instrumentos y/o sensores, se aplicó la prueba de confiabilidad Test–Retest (véase Anexo 10, Ficha A7). Asimismo, la validación del instrumento de recolección de datos se realizó mediante el juicio de cuatro expertos del SENAMHI, quienes evaluaron y aprobaron su pertinencia con una calificación de nivel Excelente. Los resultados de esta evaluación se encuentran respaldados con firmas en las fichas de validación incluidas en los Anexos 5–8 (Fichas A3, A4, A5 y A6).

3.5. Procedimientos

El desarrollo de la investigación se llevará a cabo siguiendo un enfoque sistemático, estructurado en cuatro etapas principales. Cada una de estas fases está diseñada para garantizar la calidad, confiabilidad y validez de los resultados obtenidos, en concordancia con los objetivos y la hipótesis planteados.

3.5.1. Etapa 1: recopilación y preparación de datos

En esta primera etapa, se recopilarán los datos históricos de temperatura del aire provenientes de las estaciones meteorológicas ubicadas en la Sierra del departamento de Puno, Perú. La obtención de los datos se realizará a través de fuentes oficiales, como el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). Posteriormente, se llevará a cabo un proceso de limpieza y preprocesamiento de los datos, que incluirá:

- Estandarización de formatos y unidades de medida.
- Identificación y manejo de valores faltantes o atípicos.
- Organización de los datos en series temporales y bases de datos estructuradas para su posterior análisis.

Este proceso es fundamental para asegurar la integridad y consistencia de los datos antes de su análisis (Wilks, 2019).

3.5.2. Etapa 2: desarrollo del control de calidad

En esta etapa, se implementará un proceso de control de calidad orientado a la detección de anomalías en los registros de temperatura del aire. El objetivo principal es identificar y clasificar datos atípicos o erróneos que puedan afectar la confiabilidad de los análisis posteriores.

Para ello, se aplicarán reglas estadísticas reconocidas internacionalmente para el control de calidad de datos meteorológicos, tales como las recomendadas por la Organización Meteorológica Mundial (OMM, 2017). El procedimiento incluirá:

3.5.2.1. Reglas de rango físico. Se verificará que los valores de temperatura se encuentren dentro de los límites físicos y climatológicos esperados para la región de estudio. Valores fuera de estos rangos serán marcados como potencialmente erróneos.

3.5.2.2. Reglas de consistencia interna. Se evaluará la coherencia temporal de los datos, identificando saltos abruptos, repeticiones inusuales o secuencias improbables en las series de tiempo (Wilks, 2019).

3.5.2.3. Reglas de persistencia. Se detectarán registros que permanezcan constantes durante periodos anormalmente largos, lo cual puede indicar fallas en los sensores o problemas de transmisión de datos.

3.5.2.4. Reglas de comparación espacial. Se compararán los valores de temperatura entre estaciones meteorológicas vecinas para identificar discrepancias significativas que sugieran posibles errores en una estación específica (Aguilar et al., 2003). Cada anomalía detectada será documentada y clasificada según su tipo y severidad. Los datos identificados como erróneos serán excluidos temporalmente del análisis o marcados para su posterior corrección mediante el modelo matemático desarrollado en la siguiente etapa.

Este enfoque sistemático garantiza que únicamente los datos confiables sean utilizados en el proceso de modelado y análisis, incrementando así la validez de los resultados obtenidos.

3.5.3. Etapa 3: diseño e implementación del modelo matemático

En esta etapa, se desarrollará e implementará un modelo matemático integral orientado a la corrección y el relleno de datos erróneos o faltantes en las series de temperatura del aire. El proceso se estructura en tres fases complementarias:

3.5.3.1. Diseño de un modelo para la detección de anomalías. Se aplicarán técnicas estadísticas para identificar registros atípicos o inconsistentes que no fueron filtrados en la etapa de control de calidad. Esto permitirá focalizar la corrección en los datos que realmente presentan desviaciones significativas respecto al comportamiento esperado.

3.5.3.2. Diseño de un modelo para la corrección y relleno de datos. Para la corrección y relleno de los datos de temperatura del aire diaria de los 20 años de datos de la Región Sierra de Puno se usó un modelo hibrido conformado de 2 algoritmos de Inteligencia Artificial: Random Forest (RF) y Gradiente Boosting (GB). El modelo híbrido está diseñado para imputar valores de temperatura media diaria (Tmed) faltantes o inválidos para cada una de las 22 estaciones meteorológicas en Puno. Combina las fortalezas de Random Forest (RF) y Gradient Boosting (GB) para manejar tanto valores normales como extremos de temperatura de manera efectiva, asegurando predicciones robustas para el clima variable y de gran altitud de Puno.

3.5.3.3. Propósito. El modelo híbrido aborda datos de temperatura faltantes o marcados como no válidos (no "OK") prediciendo valores de Tmed utilizando características espaciales y temporales, aprovechando datos de estaciones vecinas y variables ambientales como altitud, latitud y longitud.

3.5.3.4. Componentes del modelo. Se aplico un modelo hibrido compuesto por los modelos: Random Forest (RF) y el modelo Gradiente Boosting (GB).

A. *Random forest.* Es robusto para capturar relaciones no lineales y manejar patrones típicos de temperatura debido a su conjunto de árboles de decisión. Funciona bien

para valores "normales" de temperatura dentro de los percentiles 1 al 99. Por ende, se usa la siguiente configuración:

- 200 árboles (n_estimators=200)
- Profundidad máxima de 15 (max_depth=15)
- Mínimo 2 muestras por hoja (min_samples_leaf=2)
- Raíz cuadrada de características consideradas en cada división (max_features='sqrt')

B. *Gradiente boosting.* *D*estaca en modelar valores extremos (atípicos) al corregir errores iterativamente, lo que lo hace adecuado para temperaturas extremas (por ejemplo, días inusualmente fríos o cálidos en el clima de gran altitud de Puno). La función de pérdida Huber en GB asegura robustez frente a valores atípicos. Por ende, se usa la siguiente configuración:

- Árboles (n_estimators=500)
- Tasa de aprendizaje de 0.08 (learning_rate=0.08)
- Profundidad máxima de 6 (max_depth=6)
- Función de pérdida Huber (robusta a outliers, loss='huber')

C. *Enfoque híbrido.* Al combinar RF para valores normales y GB para extremos, el modelo equilibra precisión y robustez, abordando la variabilidad en los datos de temperatura de Puno, influenciada por la altitud (por ejemplo, 3,800–4,500 metros) y patrones climáticos regionales.

- Mecanismo de combinación: Se calculan los percentiles 1 (q1) y 99 (q9) de los datos de entrenamiento
- Para cada predicción: Si la predicción del RF está fuera del rango [q1, q9], se usa la predicción del GB De lo contrario, se mantiene la predicción del RF
- Ventajas del enfoque híbrido: RF proporciona robustez y buen rendimiento general. GB captura mejor los valores extremos. La combinación automática optimiza el rendimiento para diferentes rangos de valores.

3.5.4. Etapa 4: validación y evaluación del modelo

Para evaluar la robustez y capacidad predictiva del modelo matemático propuesto para la corrección y el relleno de datos de temperatura, se implementó la técnica de validación cruzada (*cross-validation*), ampliamente utilizada en el ámbito de la estadística y el aprendizaje automático para estimar el desempeño de modelos predictivos (James et al., 2021).

La validación cruzada consiste en dividir el conjunto de datos en varios subconjuntos o *folds*. En cada iteración, uno de estos subconjuntos se utiliza como conjunto de prueba (*test set*) y los restantes como conjunto de entrenamiento (*training set*). Este proceso se repite tantas veces como *folds* se hayan definido, permitiendo que cada subconjunto sea utilizado una vez como prueba y el resto como entrenamiento. Finalmente, los resultados obtenidos en cada iteración se promedian para obtener una estimación más robusta del error del modelo (Kuhn y Johnson, 2013).

En este trabajo se empleó la **validación cruzada k-fold** con $k = 5$, que es uno de los enfoques más comunes. Formalmente, el procedimiento se describe así:

3.5.4.1.Procedimiento para la validación cruzada. A continuación, se explica los procedimientos para realizar la validación cruzada.

A. Paso 1: Preparación de los datos. A continuación, se detalla los siguientes:

- Se seleccionan los datos con Flag_Tmed = 'OK' y valores no nulos de Tmed para garantizar que se utilicen datos válidos para la validación.
- Las variables predictoras incluyen características temporales (Day, Month, Day_sin, Day_cos) y espaciales (Altitud, Latitud, Longitud), mientras que la variable objetivo es Tmed (temperatura media).
- Los datos se dividen en k=5 folds usando el método KFold de scikit-learn con shuffle=True y una semilla (random_state=42) para garantizar reproducibilidad.

B. ***Paso 2: Entrenamiento y evaluación del modelo.*** A continuación, se explica los procedimientos para el entrenamiento y la evaluación del modelo.

- En cada iteración del K-Fold, se entrena un modelo de Random Forest con parámetros específicos (n_estimators=100, max_depth=10, random_state=42, n_jobs=-1) utilizando los datos de entrenamiento (4 folds).
- Las predicciones se generan para el fold de prueba, y se calculan las métricas MAE, RMSE y R^2 para cada fold usando las fórmulas descritas anteriormente.
- Las predicciones y valores reales de todos los folds se almacenan para análisis agregados y visualizaciones.

C. ***Paso 3: Test de friedman.*** A continuación, se explica los procedimientos para el test de Friedman.

- Se aplica el test de Friedman (función friedmanchisquare de scipy. stats) a las métricas MAE, RMSE y R^2 de los 5 folds para evaluar si existen diferencias significativas en el rendimiento del modelo entre los folds.
- Si el p-valor es menor a 0.05, se concluye que hay diferencias significativas, lo que podría indicar variabilidad en el rendimiento del modelo dependiendo de la partición de datos. Si el p-valor es mayor o igual a 0.05, se asume consistencia estadística entre los folds.

D. Paso 4: Visualizaciones y reportes. A continuación, se explica los procedimientos para la visualización y reportes.

- Tabla de evolución de Métricas por fold: Se realiza cálculos MAE, RMSE y R^2 a través de los folds, permitiendo identificar la variabilidad en el rendimiento.
- Tabla de Test de Friedman: Se muestra la comparación entre Folds mediante p-valores de t-Student. Confirma la robustez del modelo entre folds.
- Gráfico de Evolución de Métricas por Fold de Validación Cruzada:
- Se visualiza la evolución de las métricas: MAE, RMSE y R^2 a través de los folds.

- Gráfico de Importancia de Características de las Variables:
- Gráfico de barras que muestra la contribución de variables predictoras (Day, Month, Altitud, etc.) en la predicción de Tmed.
- Diagrama del Modelo Híbrido (RF + GB):
- Representa el flujo del modelo híbrido, incluyendo selección de características, entrenamiento y lógica de predicción.

3.5.4.2. Análisis de importancia de características. El modelo random forest para calcular la importancia de cada variable predictora (Day, Month, Day_sin, Day_cos, Altitud, Latitud, Longitud) en la predicción de Tmed. La importancia se basa en la contribución de cada variable a la reducción de la impureza en los nodos del árbol de decisión, visualizada en un gráfico de barras.

A. *Consideraciones técnicas.* Manejo de Datos Faltantes: El utilizo la estrategia preliminar la técnica de la mediana para imputar valores faltantes en las variables predictoras, asegurando que el modelo pueda manejar datos incompletos.

B. Robustez del modelo. El modelo híbrido combina Random Forest para valores normales y Gradient Boosting para valores extremos, mejorando la capacidad de predicción en casos atípicos.

C. Escalabilidad. El código está diseñado para procesar múltiples estaciones, considerando correlaciones espaciales y distancias geográficas (usando geopy.distance.geodesic) para seleccionar estaciones vecinas relevantes.

3.5.5. Etapa 5: Procedimiento para la detección, corrección y relleno de datos

La detección de anomalías en series temporales de temperatura del aire es un proceso fundamental en el control de calidad de datos meteorológicos, especialmente en regiones como la Sierra del Departamento de Puno. El objetivo es identificar y clasificar valores atípicos o erróneos que puedan afectar el análisis y la toma de decisiones. Este procedimiento se basa en

la aplicación de técnicas estadísticas y geoespaciales, siguiendo un enfoque sistemático y reproducible. (véanse la **Tabla 25**, **Tabla 26** y la **Figura 8**)

A continuación, se describen los principales procedimientos utilizados para la detección de anomalías, incluyendo los conceptos, algoritmos y fórmulas relevantes:

3.5.5.1. Procedimiento para la detección de datos anómalos. Detección de Valores Vacíos (SD): Se identifica registro de temperatura media (Tmed) con valor nulo o no registrada (NaN). Estos casos se clasifican bajo el tipo SD (sin dato), luego se registran con una bandera (Flag) y se excluyen de los cálculos posteriores. La fórmula es la siguiente: Si Tmed = NaN, entonces se marca con un flag de tipo SD.

3.5.5.2. Verificación con umbrales físicos de rangos mundiales (LM). Se verifica que los valores estén dentro de los rangos de temperatura válidos. La fórmula sería lo siguiente: Rango mundial: Si Tmed ∈ [−40,60] C, entonces los datos son aceptables, se marcan con un flag de tipo 'OK' caso contrario se marca con 'LM'.

3.5.5.3. Verificación con umbrales físicos de rangos nacional (LN). Se verifica que los valores estén dentro de los rangos de temperatura válidos. La fórmula es la siguiente: Si Tmed ∈ [−32,45]°C, entonces los datos son aceptables se marcan con un flag de tipo 'OK' caso contrario se marca con 'LN'.

3.5.5.4. Detección de variación excesiva (VE). Se evalúan las diferencias absolutas entre días consecutivos. Si esta diferencia excede los 10 °C, y no hay datos faltantes, el dato se considera anómalo. La fórmula es la siguiente: Si la $\Delta T\text{med} = |T\text{med}_{(t)} - T\text{med}_{(t-1)}| > 10$ °C, entonces los datos son aceptables se marcan con un flag de tipo 'OK' caso contrario se marca con 'VE'.

3.5.5.5. Detección de valores repetidos (VR). Se detectan secuencias consecutivas de cuatro o más días con el mismo valor de temperatura para cada estación. La fórmula es la

siguiente: Si $Tmed_{(t)} = Tmed_{(t-1)} = ... = Tmed_{(t-n)}$ con $n \geq 3$, se marca como 'VR'.

3.5.5.6. Filtro de hampel (FH). Es un método robusto basado en la mediana y la desviación absoluta mediana (MAD) para detectar valores atípicos. Formula: Basado en la Mediana y la Desviación Absoluta Mediana (MAD): $MAD_{ventana}$ = mediana ($|X_i - \text{mediana}(X)|$)

$MAD_{escalado} = 1.4826 * MAD_{ventana}$

$Score_i = |X_i - \text{mediana}(X)| / MAD_{escalado}$

Si $Score_i > 3.5$, se considera anomalía y se marca con 'FH'.

3.5.5.7. Verificación de consistencia espacial (CE). Se compara el valor de Tmed de una estación con la mediana de las estaciones vecinas (dentro de un radio de 50 km), si la diferencia excede 10 °C, se clasifica como inconsistente espacialmente. La fórmula es la siguiente: Si la $|Tmed_{estación} - \text{mediana}_{vecinas}| > 10 \Rightarrow$ se marca como 'CE'.

3.5.5.8. Verificación con umbrales físicos regional (LR). Se calcula los Umbrales Físicos Regional mediante la técnica estadística Boxplot, se toma como serie al conjunto de todas las estaciones de la Región. La fórmula es la siguiente:

IQR = Rango intercuartílico

$IQR = Q3 - Q1$

Límite inferior = $Q1 - 1.5 \cdot IQR$

Límite superior = $Q3 + 1.5 \cdot IQR$

Los valores fuera de este rango se clasifican como 'LR'

- Verificación con Umbrales Físicos por Estación (LE): Se calcula los Umbrales Físicos por Estación mediante la técnica estadística Boxplot, se toma como serie al conjunto de cada Estación. La fórmula es la siguiente:

  IQR = Rango intercuartílico

  $IQR = Q3 - Q1$

Límite inferior = $Q1 - 1.5 \cdot IQR$

Límite superior = $Q3 + 1.5 \cdot IQR$

Los valores fuera de este rango se clasifican como ‘LE’

En el Anexo 15, Diagrama C1 de puede visualizar el flujo del modelo de control de calidad.

3.5.5.9. Procedimiento para corrección y relleno de datos. Luego de detectar los datos anómalos que fueron marcados con un flag, a continuación, se explica los procedimientos para la corrección y relleno de datos.

A. Ingeniería de características. A continuación, se detalla los siguientes:

- Características Temporales:
- Day: Día del año (1–365/366) para capturar estacionalidad.
- Month: Número de mes (1–12) para patrones estacionales.
- Day_sin y Day_cos: Codificaciones cíclicas del día del año usando transformaciones seno y coseno para modelar periodicidad (por ejemplo, np.sin (2 * np.pi * Day/365).

B. Características espaciales. A continuación, se detalla los siguientes:

- Altitud, Latitud, Longitud: Consideran la variabilidad topográfica de Puno (por ejemplo, efectos de la altitud en la temperatura).

C. Estaciones vecinas. A continuación, se detalla los siguientes:

- Seleccionar hasta tres estaciones con la mayor correlación para Tmed (vía select_correlated_stations).
- Incorporar sus valores de Tmed como características (Tmed_neighbor_1, Tmed_neighbor_2, Tmed_neighbor_3).
- Los valores faltantes de características se imputan usando la estrategia de la mediana (SimpleImputer).

D. Detección de valores atípicos. A continuación, se detalla los siguientes:

- Calcular los percentiles 1 (q1) y 99 (q9) de Tmed para separar valores normales y extremos.
- Valores normales ($q1 \leq Tmed \leq q9$) se usan para entrenar RF.
- Valores extremos ($Tmed < q1$ o $Tmed > q9$) se usan para entrenar GB.

E. Parámetros del modelo. A continuación, se detalla los siguientes:

Random forest:

- n_estimators=200: Muchos árboles para promediado robusto.
- max_depth=15: Limita el sobreajuste mientras captura patrones complejos.
- min_samples_leaf=2: Asegura que los nodos hoja tengan al menos dos muestras para estabilidad.
- max_features='sqrt': Usa la raíz cuadrada de las características para diversidad.
- random_state=42: Asegura reproducibilidad.
- n_jobs=-1: Usa todos los núcleos de CPU disponibles para mayor velocidad.

Gradient boosting:

- n_estimators=500: Más iteraciones para corrección precisa de errores.
- learning_rate=0.08: Tasa de aprendizaje moderada para convergencia equilibrada.
- max_depth=6: Permite modelar patrones complejos en valores extremos.
- min_samples_leaf=1: Permite divisiones detalladas para valores extremos.
- loss='huber': Función de pérdida robusta que combina error cuadrático (para errores pequeños) y error absoluto (para errores grandes).
- random_state=42: Asegura reproducibilidad.

F. Lógica de predicción híbrida. A continuación, se detalla los siguientes:

- Para cada predicción, RF genera una predicción preliminar (rf_pred).
- Si rf_pred < q1 o rf_pred > q9, se usa la predicción de GB (gb_pred); de lo contrario, se usa rf_pred.
- Esto asegura que GB maneje valores extremos, mientras que RF maneja valores típicos.

G. Generación de predicciones:

- Preparar las mismas características para los registros objetivo que se usaron en el entrenamiento.
- Aplicar el modelo híbrido para predecir valores faltantes de Tmed y almacenarlos en Tmed_Generated.

H. Post-procesamiento:

- Ajusta los valores generados para que coincidan con las propiedades estadísticas (media y desviación estándar) de los datos originales por estación.
- Aplica factores de escalado (por ejemplo, 1.25 para valores atípicos negativos, 1.15 para ajustes negativos, 1.10 para ajustes positivos) para mantener consistencia con los datos observados.

I. Validación:

- La validación cruzada (cross_validate) evalúa el rendimiento del modelo usando métricas como MAE, RMSE y R^2.
- Se implementó una validación cruzada de 5 folds para evaluar la precisión de la imputación, calculando métricas como MAE (1.20°C), RMSE (1.55°C) y R^2 (0.74) sobre 134,344 muestras, según el documento del código. (véanse la **Tabla 6** y la **Figura 29**)
- Luego de calcular las métricas por cada subgrupo o folds, se aplicó a cada subgrupo el Test de Friedman para validar el rendimiento del modelo híbrido (Random Forest + Gradient Boosting). Su objetivo es determinar si existen diferencias estadísticamente significativas

en las métricas de evaluación (MAE, RMSE, R^2) entre los diferentes folds de validación cruzada.

J. El test de friedman:

Friedman es una alternativa no paramétrica (no requiere normalidad de los datos). La prueba de Friedman (Friedman, t-test) verifica diferencias significativas entre pliegues, asegurando la estabilidad del modelo.

- El test de Friedman evalúa:
- Hipótesis nula (H_0): No hay diferencias entre los folds (el modelo es estable).
- Hipótesis alternativa (H_1): Al menos un fold es significativamente diferente.
- Prueba de Hipótesis: Si p-valor ≥ .05, entonces se aprueba la Hipótesis nula (H_0), por lo tanto, el rendimiento es consistente (el modelo es robusto). Si p-valor < .05, entonces se aprueba la Hipótesis nula (H_1) El modelo varía significativamente entre folds (posible sobreajuste o datos no homogéneos). (véanse la **Tabla 7**)
- Visualizaciones: Se desarrollaron gráficos de dispersión, gráficos de residuos y en la **Tabla 33** y la **Figura 30** se observa el análisis de importancia de características con la finalidad de evaluar el rendimiento del modelo y la contribución de las características. En la **Figura 31** se observa el flujo del modelo hibrido en base a los algoritmos de Inteligencia Artificial: Random Forest y Gradiente Boosting.

El análisis de datos se diseñó para evaluar y mejorar la calidad de los registros de temperatura del aire en las estaciones meteorológicas de la sierra del departamento de Puno, mediante un enfoque cuantitativo que combina técnicas estadísticas y computacionales. Este proceso se estructuró en dos componentes principales: a) Detección de anomalías, para identificar y clasificar errores en los datos, y b) Corrección y relleno de datos, para rectificar valores erróneos y completar series temporales faltantes. Estas fases responden a las preguntas de investigación, validan las hipótesis planteadas y aseguran la precisión y confiabilidad de los

datos, utilizando métricas estadísticas con un nivel de significancia de $p < .05$ y correlaciones espaciales (% r Rellenado > % r Original).

3.6. Análisis de datos

El análisis de datos se diseñó para evaluar y mejorar la calidad de los registros de temperatura del aire en las estaciones meteorológicas de la sierra del departamento de Puno, mediante un enfoque cuantitativo que combina técnicas estadísticas y computacionales. Este proceso se estructuró en dos componentes principales: a) Detección de anomalías, para identificar y clasificar errores en los datos, y b) Corrección y relleno de datos, para rectificar valores erróneos y completar series temporales faltantes. Estas fases responden a las preguntas de investigación, validan las hipótesis planteadas y aseguran la precisión y confiabilidad de los datos, utilizando métricas estadísticas con un nivel de significancia de $p < .05$ y correlaciones espaciales (% r Rellenado > % r Original).

3.6.1. Análisis para la detección de anomalías

Se calcularon estadísticas descriptivas para cada tipo de anomalía, incluyendo la frecuencia relativa (porcentaje del total de registros) y la distribución temporal y espacial de los errores. Por ejemplo, los valores vacíos (SD) se analizaron para identificar patrones estacionales o relacionados con el mantenimiento de las estaciones. Las anomalías LM y LN se evaluaron para determinar si estaban asociadas con condiciones climáticas extremas o fallos instrumentales, mientras que VE y VR indicaron problemas de consistencia temporal. El análisis de consistencia espacial (CE) se apoyó en la matriz de distancias para verificar correlaciones entre estaciones vecinas, utilizando el coeficiente de Pearson.

(véanse las figuras: **Figura 17**, **Figura 18**, **Figura 19**, **Figura 20** y **Figura 21**)

Tabla 3.

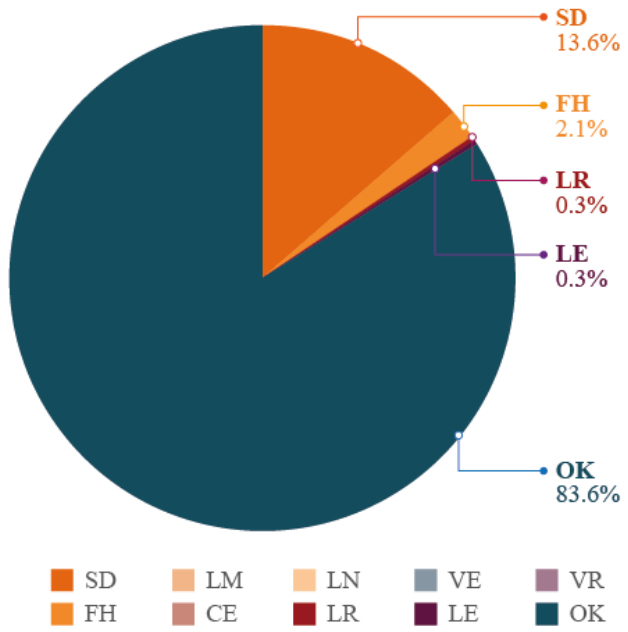
Indicadores del control de calidad

Indicador	Indicador CC	Cantidad	Porcentaje	Grupo
Vacío	SD	21,928	13.6	Relleno
Fuera rango mundial	LM	0	0	
Fuera rango nacional	LN	0	0	
Variación excesiva	VE	31	0	
Valor repetido	VR	48	0	Corrige
Filtro hampel	FH	3,340	2.1	
Consistencia espacial	CE	0	0	
Fuera rango región	LR	471	0.3	
Fuera rango estación	LE	548	0.3	
Buena	OK	134,344	83.6	Buena
Total		160,710	100	

Nota. Elaboración propia

Figura 1.

Gráfico de indicadores de control de calidad



Nota. Elaboración propia

En la **Tabla 3** y la **Figura 1Figura 3** se observar los indicadores de control de calidad alrededor del 83.6% son datos buenos, el 2.8% los datos están son detectados como anómalos, un 13.6% fueron detectados como sin dato y finalmente imputados. El modelo hibrido desarrollado corrigió y relleno el 16.3% de los datos.

3.6.2. Análisis de corrección y relleno de datos

El componente de corrección y relleno de datos se diseñó para rectificar valores anómalos y completar series temporales faltantes, utilizando el modelo matemático híbrido. Este proceso se basó en técnicas estadísticas y de aprendizaje automático, integrando Random Forest, Gradient Boosting.

3.6.2.1. Análisis para validar la hipótesis H1 mediante la reducción de incertidumbre. La incertidumbre se refiere a la variabilidad o error asociado a los datos, que puede deberse a:

- Errores instrumentales.
- Falta de homogeneidad en las series temporales.
- Datos faltantes o anomalías no detectadas.

Para cuantificar la incertidumbre en los datos de temperatura, se siguieron estos pasos rigurosos:

A. Reducción en la dispersión de los datos

Para cada estación y variable (Tmed) se calculó de la desviación estándar (σ) tanto para la serie original de datos crudos (con errores/anomalías) como para la serie de datos corregida o rellenada (mediante Random Forest + Gradient Boosting). (véanse la **Figura 22** y **Figura 23**)

En la **Tabla 4** y la **Figura 2** se observan una suposición de disminución de la incertidumbre al comprobar que el 95% de los datos la desviación Estándar corregido es menor que desviación estándar original o crudo, para probar esta suposición es necesario realizar la prueba de t-Student.

B. ***Cambio significativo en las estadísticas tras la corrección y relleno de datos.*** Para verificar la suposición de disminución de la incertidumbre detectado en el punto anterior (I) se procedió a realizar la prueba de t-Student. Para cada estación y variable (Tmed) se calculó la prueba t-Student

Hipótesis nula (H_0): No hay diferencia entre las medias de T_original y T_corregido.

Si "p-valor < .05", rechazas H_0: El modelo redujo significativamente la incertidumbre.

(véanse la **Tabla 30**, la **Tabla 31**, la **Tabla 32**, la **Figura 24**, la **Figura 25** y la **Figura 26**)

C. ***Resultado de la validación.*** Se observa una suposición de disminución de la incertidumbre al comprobar que el 95% de los datos la desviación Estándar corregido es menor que desviación estándar original o crudo, para probar esta suposición es necesario realizar la prueba de t-Student.

En la **Figura 2** se observa que el 95% de los datos tienen p-valor < 0.05, por lo tanto, se afirma que el modelo redujo significativamente la incertidumbre (H1).

Tabla 4.

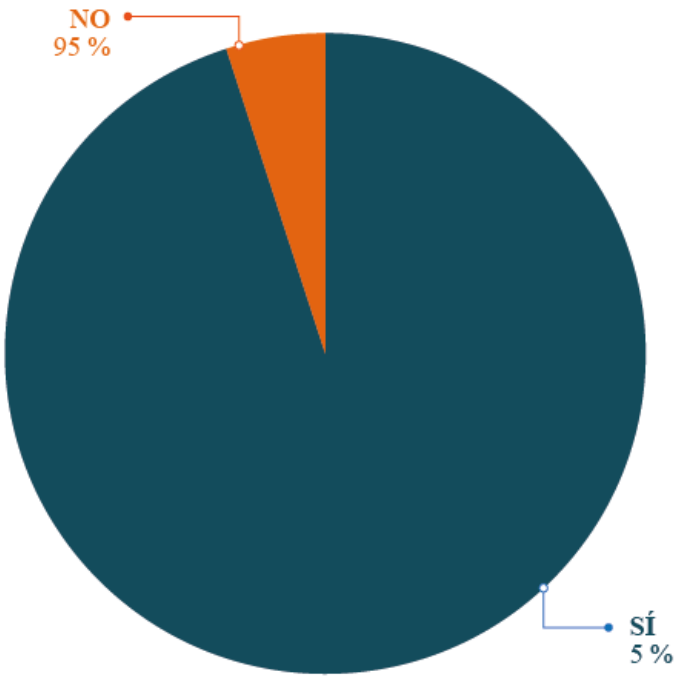
Indicadores estadísticos de corrección y relleno

Id Esta	Desviación estándar original	Desviación estándar corregido	t-Student (p-valor)	Reduce incerti-dumbre	Correlación original rellenado	RMSE Original rellenado	Datos originales (%)	Datos generados (%)
E01	1.8176	1.7947	0	Si	0.9956	0.1706	94.9076	5.0924
E02	2.4322	2.4012	0	Si	0.9959	0.2226	95.0582	4.9418
E03	1.4587	1.4252	0.5123	No	0.9885	0.2211	95.0308	4.9692
E04	2.7075	2.5902	0	Si	0.9705	0.6581	94.5654	5.4346
E05	2.6887	2.6421	0	Si	0.9959	0.2456	35.6057	64.3943
E06	2.3471	2.2784	0	Si	0.9913	0.3151	95.0582	4.9418
E07	2.4186	2.3671	0	Si	0.9912	0.3213	97.9055	2.0945
E08	2.181	2.1034	0	Si	0.9857	0.3727	96.1944	3.8056
E09	1.909	1.7967	0	Si	0.9808	0.3837	99.7536	0.2464
E10	1.82	1.7455	0	Si	0.9847	0.3226	96.8925	3.1075
E11	2.9596	2.8866	0	Si	0.9935	0.3426	94.8939	5.1061
E12	2.4271	2.3429	0	Si	0.9866	0.4011	98.3162	1.6838
E13	2.7539	2.6919	0	Si	0.9899	0.3915	99.3977	0.6023
E14	2.7678	2.7221	0	Si	0.9958	0.2563	33.4428	66.5572
E15	2.4876	2.418	0	Si	0.9931	0.2978	97.1526	2.8474
E16	4.7435	2.4515	0.0001	Si	0.5207	4.0564	60.6571	39.3429
E17	3.4981	3.2582	0	Si	0.9741	0.8107	99.6578	0.3422
E18	2.5367	2.3868	0	Si	0.9785	0.5352	99.9726	0.0274
E19	2.2268	2.1863	0.0004	Si	0.9968	0.1804	30.1711	69.8289
E20	2.5247	2.3393	0	Si	0.968	0.6493	98.5763	1.4237
E21	2.7127	2.4989	0	Si	0.9697	0.6803	95.5647	4.4353
E22	1.2826	1.219	0	Si	0.9563	0.3756	91.0472	8.9528
Prom.	*2.49*	*2.3*	*0.02*		*0.96*	*0.56*	*86.36*	*13.64*
Med.	*2.46*	*2.38*	*0*		*0.99*	*0.36*	*95.31*	*4.69*
	Estación automática							

Nota. Elaboración propia

Figura 2.

Estadística de reducción de incertidumbre



Nota. Elaboración propia

3.6.2.2. Análisis para validar la hipótesis H2 mediante la consistencia espacial. Para cuantificar la Consistencia Espacial en los datos de temperatura, se siguieron estos pasos rigurosos:

- Paso 1: Calcular matriz de correlación entre estaciones vecinas con datos originales (original o sin corregir).
- Paso 2: Aplicar el modelo y recalcular correlaciones con datos corregidos y rellenados.
- Paso 3: Comparar los valores de r de las 2 matrices. (véanse la **Tabla 5**)
- Para validar la mejora, se calculó el porcentaje de disminución de r en cada estación de la serie corregidos y rellenados con respecto a la serie original o sin corregir.

Tabla 5.

Resumen de matriz de correlación de datos original vs rellenado

Estación	Correlación media original	Correlación media rellenado	Mejora r	Tipo de estación
E01	0.781800133	0.789935639	Si	Con
E02	0.766018649	0.77373769	Si	Con
E03	0.566951752	0.581531965	Si	Con
E04	0.811576953	0.802433519	No	Con
E05	0.828224159	0.803040273	No	EMA
E06	0.77346754	0.783392421	Si	Con
E07	0.767381883	0.774904242	Si	Con
E08	0.756649356	0.76395397	Si	Con
E09	0.661591342	0.67352443	Si	Con
E10	0.68147223	0.696504107	Si	Con
E11	0.750742058	0.769330701	Si	Con
E12	0.698782461	0.717550872	Si	Con
E13	0.748464297	0.769606205	Si	Con
E14	0.772473542	0.760482667	No	EMA
E15	0.688390374	0.709644788	Si	Con
E16	0.417578557	0.649276572	Si	Con
E17	0.747491439	0.729125619	No	Con
E18	0.741556043	0.722876685	No	Con
E19	0.680560116	0.660423977	No	EMA
E20	0.565810384	0.549997866	No	Con
E21	0.401240985	0.368201889	No	Con
Promedio	0.70	0.71		

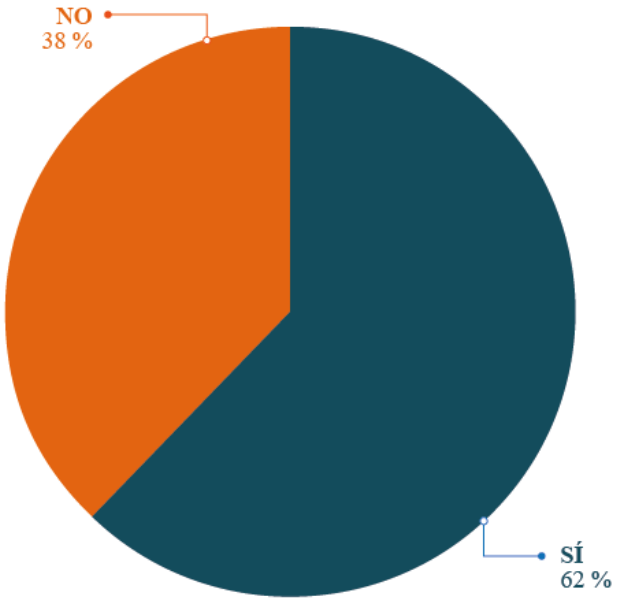
Leyenda

Con	Convencional	**EMA**	Automática

Nota. Elaboración propia

Figura 3.

Gráfico de porcentaje de aumento de coeficiente (r)



Nota. Elaboración propia

- Paso 4: Resultado de la validación

En la **Tabla 5** se observa que todas las estaciones convencionales tienen una leve mejoría en el coeficiente de correlación espacial, a excepto la Estación E04, mientras en las 4 estaciones automáticas se mantiene no mejoraron su coeficiente de correlación, en general el 62% de las estaciones tienen un aumento en la correlación múltiple de los datos corregido o rellenados (H2), por lo tanto, se confirma la confiabilidad del modelo al corregir y rellenar los datos. (véanse la **Figura 3** y la **Figura 28**)

3.6.2.3. Análisis para validar la eficacia del modelo matemático. A continuación**,** se presentas los siguientes pasos:

A. Análisis de validación cruzada. La validación cruzada aplicada al modelo matemático para la corrección y el relleno de datos de temperatura del aire en la Sierra del departamento de Puno muestra una notable consistencia en el desempeño del modelo a través de los cinco folds evaluados. Las métricas obtenidas para cada fold son las siguientes:

- MAE (Error Absoluto Medio): Entre 1.1961 y 1.2165

- RMSE (Raíz del Error Cuadrático Medio): Entre 1.5375 y 1.5665
- R^2 (Coeficiente de determinación): Entre 0.7296 y 0.7407. Estos valores indican que el modelo mantiene un nivel de error bajo y una capacidad explicativa alta de manera homogénea en las distintas particiones de los datos.

B. Interpretación de la validación cruzada. La baja dispersión de las métricas entre los folds evidencia que el modelo es robusto y no depende de una partición específica de los datos para lograr un buen desempeño. Se puede observar en la **Tabla 6**, el MAE y el RMSE presentan variaciones mínimas, lo que sugiere que el modelo es capaz de predecir valores de temperatura con un error absoluto y cuadrático medio muy similar en todos los subconjuntos de validación. Asimismo, los valores de R^2 superiores a 0.72 en todos los folds reflejan que el modelo explica más del 72% de la variabilidad de los datos de temperatura corregidos, lo cual es adecuado para aplicaciones de control de calidad en meteorología.

C. Análisis del test de friedman. El test de Friedman fue utilizado para comparar estadísticamente el desempeño del modelo entre los diferentes folds, empleando las métricas MAE, RMSE y R^2. En la **Tabla 7** se observa los resultados:

Al realizar la prueba de t-Student muestra el valor de p-valor: 0.406006 para todas las métricas, es decir el p-valor es mayor que .05, con estos resultados se aprueba la Hipótesis nula (Ho) al no existir diferencias estadísticamente significativas en el desempeño del modelo entre los distintos folds. Esto refuerza la evidencia de que el modelo es estable y generaliza bien a diferentes subconjuntos de datos, cumpliendo así con uno de los principales criterios de validez en modelos predictivos para control de calidad de datos ambientales.

Tabla 6.

Evolución de métricas por fold

Subgrupo	MAE	RMSE	R^2
Fold 1	1.199	1.5415	0.7351
Fold 2	1.1965	1.5375	0.7407
Fold 3	1.1961	1.5413	0.7393
Fold 4	1.2165	1.5665	0.7296
Fold 5	1.2132	1.5629	0.732
Promedio	1.20	1.55	0.74

Nota. Elaboración propia

Tabla 7.

Test de friedman

Métrica	Estadístico friedman	p-valor	Conclusión
MAE	4	0.406006	No hay diferencias significativas entre folds
RMSE	4	0.406006	No hay diferencias significativas entre folds
R2	4	0.406006	No hay diferencias significativas entre folds

Nota. Elaboración propia

D. El test de friedman evalúa. A continuación, se detalla:

- Hipótesis nula (H_0): No hay diferencias entre los folds (el modelo es estable).
- Hipótesis alternativa (H_1): Al menos un fold es significativamente diferente. Si p-valor < .05: El modelo varía significativamente entre folds (posible sobreajuste o datos no homogéneos). Si p-valor ≥ .05: El rendimiento es consistente (el modelo es robusto).

3.6.3. Análisis de la validación operativa

Los cuatro expertos del SENAMHI proporcionaron una validación por juicio de expertos, revisaron y validaron el Instrumento de recolección de los datos y resultados de los datos del modelo matemático, confirmando su validación a través de la ficha de validación firmada. (véanse las tablas: **Tabla 8**,**Tabla 14**,**Tabla 15**,**Tabla 16**, **Tabla 17**, **Tabla 18**, **Tabla 19**, **Tabla 20**,**Tabla 21** y **Tabla 22**)

Cada experto dio su recomendación de las cuales fueron incorporadas en el presente

estudio. A continuación, se detalla las recomendaciones:

3.6.3.1. Validación del instrumento de recolección de datos y resultados del modelo. A continuación, se detalla:

A. Experto 1. A continuación, se detalla:

- Nombre y Apellidos: Ordoñez Gálvez, Juan Julio
- Grado Académico: Doctor En Medio Ambiente y Desarrollo Sostenible (U. Nacional Federico Villarreal - Perú)
- Función: Investigador Especialista en Hidrología del SENAMHI.
- Calificación (al 100%): 90.12 (véanse la **Tabla 14**)
- Nivel: Excelente
- Recomendación: Se recomienda realizar 2 Mapas: Correlación Espacial Original versus Correlación Espacial Corregida o Rellenado para un mejor análisis de las mejoras del modelo.
- Acción: Se tomo la recomendación, incorporando los 2 mapas que figuran en la **Figura 25** y la **Figura 26**.

B. Experto 2. A continuación, se detalla:

- Nombre y Apellidos: Lavado Casimiro, Waldo Sven
- Grado Académico: Doctor en Hidrología, Hidroquímica, Suelos y Medio Ambiente (U. Toulouse III - Francia)
- Función: subdirector de Estudios e Investigaciones Hidrológicas, Investigador Especialista en Hidrología del SENAMHI.
- Calificación (al 100%): 90.12 (véanse la **Tabla 15**)
- Nivel: Excelente
- Recomendación: Se recomienda realizar un Gráfico de Distribución de Residuales, para demostrar que los errores son aleatorios y no presentan patrones sistemáticos o sesgos.

- Acción: Se tomo la recomendación, incorporando el grafico que se observa en la **Figura 27**.

C. Experto 3. A continuación, se detalla:

- Nombre y Apellidos: Paredes Riveros, Marco Antonio
- Grado Académico: Maestro en Ciencias en Gestión Ambiental (U. Amazonía Peruana – Perú)
- Función: director zonal 8 de Loreto, Investigador Especialista en Meteorología del SENAMHI.
- Calificación (al 100%): 91.36 (véanse la **Tabla 16**)
- Nivel: Excelente
- Recomendación: Se recomienda realizar una comparación de los datos de una estación convencional versus una estación automática ubicadas en el mismo lugar con la finalidad de validar la calidad y calibración del instrumento de registro de datos.
- Acción: Se tomo la recomendación, incorporando el mapa que se observa en la **Figura 7.** *Mapa de clasificación climática del SENAMHI – 2021*.

D. Experto 4. A continuación, se detalla:

- Nombre y Apellidos: Flores Sancho, Sixto
- Grado Académico: Maestro en Meteorología Aplicada, Business Innovación: Competencias y Herramientas Avanzadas para la Generación de Innovación (U. Politécnica de Valencia - España)
- Función: director zonal 13 de Puno, Investigador Especialista en Meteorología del SENAMHI.
- Calificación (al 100%): 90.12 (véanse la **Tabla 17**)
- Nivel: Excelente
- Recomendación: Se recomienda realizar gráficos de indicadores de detección de anomalías, corrección y rellenos de datos para validar y comunicar el desempeño de tu modelo

matemático en cada etapa del control de calidad de datos

- Acción: Se tomo la recomendación, incorporando la **Tabla 3** y la **Figura 1**

3.6.3.2. Resultado de la evaluación de validación del instrumento de recolección de datos y resultados del modelo por los expertos. A continuación, se detalla:

- Para evaluar el grado de consenso entre los expertos en la validación del instrumento, se aplicó un análisis de consistencia porcentual. Este método considera que existe consenso cuando la mayoría de las calificaciones de los expertos se encuentran dentro de un rango de ±1 punto respecto a la media del ítem, en una escala Likert de 1 a 5.
- Se desarrollo un script desarrollado en Python para desarrollar el proceso de calificación final sobre los resultados de evaluación de los Expertos.
- Los resultados se expresan en porcentaje y se categorizan cualitativamente como Deficiente, Regular, Bueno, Muy Bueno o Excelente, según el valor de la media porcentual.
- Este enfoque es especialmente útil cuando el número de expertos es reducido, permitiendo una interpretación clara y visual del grado de acuerdo alcanzado (Polit y Beck, 2006).
- En la **Tabla 8** se observa los resultados de la evaluación de los 4 expertos, obteniendo un promedio de 88.89%, y una consistencia del 97.22%, otorgando de manera general una calificación de Excelente.
- La evalúa la percepción de los expertos sobre el modelo matemático (Random Forest + Gradient Boosting) a través de nueve ítems, cumple con criterios de calidad (claridad, coherencia).
- La consistencia porcentual (75-100% para la mayoría de los ítems) indica un alto acuerdo entre los expertos, respaldando la aplicabilidad del modelo en Puno.
- Las categorías Excelente reflejan una evaluación positiva del modelo, alineada con los objetivos de control de calidad de datos meteorológicos.
- Los resultados del análisis de consistencia interna con medias > 97%, lo que significa que

se obtuvo consistencias altas que confirman la validez del instrumento para evaluar el modelo matemático. Los resultados se observan en la **Tabla 8**

Tabla 8.

Método de evaluación de resultados de experto

Ítem	Media Porcentaje	Desviación Estándar	Mínimo	Máximo	Categoría	Consistencia (%)
2 Objetividad	93.75	12.5	75	100	Excelente	100
5 Suficiencia	93.75	12.5	75	100	Excelente	100
8 Coherencia	93.75	12.5	75	100	Excelente	100
1 Claridad	87.5	14.43	75	100	Excelente	100
3 Actualidad	87.5	14.43	75	100	Excelente	100
6 Intencionalidad	87.5	14.43	75	100	Excelente	100
7 Consistencia	87.5	25	50	100	Excelente	75
9 Metodología	87.5	14.43	75	100	Excelente	100
4 Organización	81.25	12.5	75	100	Excelente	100
Promedio	88.89	14.75	72.22	100.00	Excelente	97.22

Nota. Elaboración propia

3.6.3.3. Características importantes del modelo. A continuación de explican las características del modelo:

- Robustez: El modelo híbrido maneja tanto temperaturas normales como extremas, crucial para el clima variable de Puno.
- Conciencia Espacial: Incorpora datos de estaciones vecinas correlacionadas, aprovechando los patrones espaciales de temperatura de Puno.
- Modelado Temporal: Las codificaciones cíclicas capturan tendencias estacionales, importantes para datos de temperatura diarios.
- Validación: La validación cruzada y pruebas estadísticas (por ejemplo, Friedman, t-Student) aseguran la confiabilidad del modelo.
- Visualización: Salidas como gráficos de importancia de características y mapas de calor de correlación proporcionan información sobre el comportamiento del modelo y mejoras en la calidad de los datos.

3.6.3.4. Modelado jerárquico. A continuación de explican las características del modelo jerárquico:

- El sistema implementa tres niveles de modelado:
- Modelo local: Para estaciones con suficientes datos (>30 muestras)
- Modelo regional: Para estaciones con pocos datos, usando vecinas cercanas
- Modelo global: Como último recurso para estaciones con muy pocos datos

3.6.3.5. Justificación para el contexto de Puno. A continuación de explican las características del modelo jerárquico:

- El entorno de gran altitud de Puno (Altiplano Andino) genera una gran variabilidad en las temperaturas debido a la elevación, fluctuaciones diurnas y microclimas.
- El modelo híbrido considera correlaciones espaciales (vía estaciones vecinas) y patrones temporales (vía codificaciones cíclicas), lo que lo hace adecuado para imputar datos faltantes mientras preserva las características regionales y estacionales.
- El uso de RF y GB asegura robustez frente a eventos climáticos extremos ocasionales en Puno (por ejemplo, olas de frío o calor).

3.6.3.6. Bondades del modelo. El enfoque híbrido RF+GB demuestra ser efectivo para manejar tanto valores normales como extremos, mientras que la estructura jerárquica (local-regional-global) asegura que todas las estaciones puedan ser procesadas incluso con disponibilidad limitada de datos. Este sistema proporciona un marco robusto para:

- Control de calidad de datos meteorológicos
- Generación de series temporales completas
- Validación estadística de los resultados
- Análisis espacial de correlaciones

3.6.3.7. Análisis de sensibilidad. Se evaluó el impacto de los hiperparámetros (número de árboles en Random Forest: 50, 100, 200; tasas de aprendizaje en Gradient Boosting: 0.01,

0.05, 0.1) y el tamaño de la muestra en el rendimiento, consolidando los resultados en gráficos de caja, tendencias, dispersión y distribución de residuos.

3.7. Consideraciones éticas

La presente investigación cumple con los principios éticos fundamentales aplicables al desarrollo de estudios científicos, incluso cuando no involucra directamente a personas o seres vivos. Si bien el objeto de estudio se centra en el tratamiento y análisis de datos meteorológicos históricos, se han tenido en cuenta aspectos éticos importantes relacionados con la integridad científica, el uso responsable de la información y el respeto a los derechos de autor y la propiedad intelectual.

3.7.1. Uso responsable de datos

Los datos utilizados en esta investigación provienen de fuentes oficiales y públicas, específicamente del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI), los cuales fueron empleados únicamente con fines científicos y académicos. Se respetó en todo momento el marco legal de uso de información pública, garantizando la transparencia y el acceso libre a los datos climáticos.

3.7.2. Veracidad y honestidad en la recolección y análisis de información

Se garantizó la fidelidad de los datos empleados, así como la objetividad en el análisis estadístico y matemático. No se alteraron, manipularon ni falsificaron resultados para favorecer hipótesis, manteniendo una conducta ética en cada etapa del proceso investigativo.

3.7.3. Reconocimiento de fuentes y derechos de autor

Se realizó un uso adecuado de la bibliografía y material consultado, siguiendo el sistema de citación APA (7.ª ed.), para dar el debido crédito a los autores originales y evitar cualquier forma de plagio. Todo contenido externo ha sido debidamente citado, y las contribuciones propias han sido claramente diferenciadas.

3.7.4. Confidencialidad y protección de la información

Aunque no se trabajó con datos personales ni información sensible, se tomó la precaución de anonimizar cualquier posible referencia a responsables o nombres propios de estaciones meteorológicas que pudieran estar ligadas a instituciones específicas, priorizando la neutralidad del análisis.

Esta investigación se desarrolló respetando los principios de responsabilidad, transparencia, rigurosidad científica y ética profesional, conforme a las buenas prácticas de investigación exigidas por las instituciones académicas y por la comunidad científica.

IV. RESULTADOS

En este capítulo se presentan los principales hallazgos obtenidos tras la aplicación del modelo matemático híbrido de control de calidad para la corrección y el relleno de datos erróneos de temperatura del aire en las estaciones meteorológicas de la Sierra del departamento de Puno, Perú. Se exponen los resultados de las métricas estadísticas, la validación cruzada, el análisis de la consistencia espacial y la visualización de la mejora en la calidad de los datos.

El modelo propuesto logró generar una serie completa de temperatura media diaria (Tmed_Rellenado) para las 22 estaciones, cubriendo el periodo 2001-2024. Se utilizaron modelos Random Forest (RF) para valores normales y Gradient Boosting (GB) para valores extremos, evaluando la capacidad del modelo híbrido para corregir datos erróneos o ausentes.

Se presentan los resultados cuantitativos del modelo híbrido, organizados según los objetivos específicos.

4.1. Resultados sobre la cuantificación de la incertidumbre de los datos

Se presentan los resultados cuantitativos del modelo híbrido, organizados según el objetivo específico O1 y demostrar la Hipótesis H1.

Para cuantificación de la incertidumbre de los datos se ejecutó el modelo híbrido Random Forest + Gradiente Boosting, calculando métricas como el error cuadrático medio (RMSE), prueba de t-Student y el coeficiente de determinación (R^2).

Se realizo un análisis estadístico de los resultados de los cálculos de la Temperatura de cada estación donde se obtuvo los siguiente:

- El 5% de los datos han sigo corregidos y rellenados
- El 100% de las estaciones resultaron una reducción de la desviación estándar de los datos corregidos y rellenados al comparar con los datos originales o crudos.
- Los resultados de realizar la prueba de t-student a los resultados de la desviación estándar

de los datos originales y Corregidos, el 95% de las estaciones muestran una reducción significativa en los errores de predicción tras la corrección y el relleno, luego se realizó la prueba de t-student, con valores de $p < .05$, lo que valida la efectividad del modelo para disminuir significativamente la incertidumbre en las series temporales de temperatura.

- El 95% de las estaciones reporto un coeficiente de correlación (r) mayor a .9, demostrando su precisión en la corrección y relleno de datos, de esta manera se aprueba la Hipótesis H1.

4.2. Consistencia espacial de los datos corregidos

Se presentan los resultados cuantitativos del modelo híbrido, organizados según el objetivo específico O2 y demostrar la Hipótesis H2.

Se generaron matrices de correlación espacial entre estaciones, tanto para los datos originales como para los datos corregidos y rellenados. El análisis muestra un incremento en los coeficientes de correlación entre estaciones vecinas después de la corrección (% r Rellenado > % r Original), lo que evidencia una mejora en la consistencia espacial de las series de temperatura.

- En la Tabla 2 se observa que las estaciones automáticas no mejoran en la correlación espacial, en cambio las estaciones convencionales si reporta un incremento en la correlación.
- En la Tabla 1, se observa que el 62% de las estaciones reporta un incremento en la correlación espacial (se compara la correlación una estación con cada una de los demás restantes).
- En la Tabla 1, se deduce que, si no se considera en el cálculo las estaciones automáticas, entonces el porcentaje se incrementaría desde el 62% hacia el 72% de las estaciones reportando un incremento en la correlación espacial.
- En la **Figura 3** se observa que el resultado del análisis al comparar la matriz de correlación espacial de los datos originales versus la matriz de los datos corregidos y rellenados concluyen que el 62% de las estaciones reportan un incremento en la correlación espacial,

demostrando su confiabilidad en la corrección y relleno de datos, de esta manera se aprueba la Hipótesis H2.

4.3. Resultado sobre la eficacia del algoritmo del modelo

Los resultados obtenidos en la validación cruzada y el test de Friedman demuestran la robustez y consistencia del modelo matemático propuesto para la corrección de datos de temperatura en las estaciones meteorológicas de Puno. A continuación, se presenta un análisis detallado basado en los hallazgos:

4.3.1. Consistencia del modelo en validación cruzada

Los valores de las métricas de evaluación (MAE, RMSE y R2) son consistentes en los cinco folds, lo que indica que el modelo no está sobre ajustado y generaliza adecuadamente para diferentes subconjuntos de datos:

4.3.1.1. Error absoluto medio (MAE). A continuación, se muestra los resultados:

- Promedio: 1.20 C (rango: 1.19–1.22 °C).
- Interpretación: El modelo comete un error promedio de ±1.2 °C en la predicción de temperaturas, lo que es aceptable para aplicaciones meteorológicas en regiones complejas como la sierra de Puno.

4.3.1.2. Raíz del error cuadrático medio (RMSE). A continuación, se muestra los resultados:

- Promedio: 1.55 C (rango: 1.54–1.57 °C).
- Interpretación: Los errores mayores (penalizados por RMSE) no superan 1.6 °C, lo que sugiere que el modelo maneja adecuadamente valores atípicos.

4.3.1.3. Coeficiente de determinación (R2). A continuación, se muestra los resultados:

- Promedio: 0.73 (rango: 0.73–0.74).
- Interpretación: El modelo explica el 73% de la variabilidad de los datos, lo que evidencia

una fuerte relación entre las predicciones y los valores observados.

4.3.1.4. Resultados del test de friedman. A continuación, se muestra los resultados: El test de Friedman se aplicó para evaluar si existen diferencias estadísticamente significativas en el desempeño del modelo entre los folds:

- Hipótesis nula (H0): No hay diferencias entre los folds.
- Resultados:

  MAE: p = 0.406
  RMSE: p = 0.406
  R2: p = 0.406
- Interpretación: En los tres casos, p>0.05, por lo que no se rechaza H0. Esto confirma que las diferencias observadas en las métricas entre folds son producto del azar y no de limitaciones del modelo.
- El modelo es estadísticamente consistente en distintas particiones de datos, lo que respalda su aplicabilidad en condiciones heterogéneas.
- Relación con las hipótesis: Hipótesis Específica H1. La reducción de incertidumbre se evidencia en los bajos valores de MAE y RMSE.

4.3.1.5. Resultados de la validación cruzada y test de friedman. A continuación, se muestra los resultados:

En la **Tabla 6**, la **Tabla 7** y la **Figura 29** se observa los resultados de la validación cruzada y el test de Friedman demuestran que el modelo matemático propuesto es consistente, robusto y no muestra variaciones significativas en su desempeño al corregir y rellenar datos erróneos de temperatura del aire en las estaciones meteorológicas de la Sierra de Puno. Esto respalda su idoneidad para ser implementado en sistemas de control de calidad de datos meteorológicos, garantizando precisión y confiabilidad en la información generada para la toma de decisiones ambientales y climáticas, por tanto, se concluye que se demuestra la eficacia del algoritmo del modelo del matemático, al corregir o rellenar los datos.

V. DISCUSIÓN DE RESULTADOS

Los resultados obtenidos confirman la hipótesis general y específicas planteadas. La implementación del modelo híbrido permitió reducir la incertidumbre de las series de temperatura, evidenciado por la disminución de los errores estadísticos y la mejora en los valores de R^2. Asimismo, el aumento en la correlación espacial entre estaciones vecinas respalda la capacidad del modelo para mejorar la precisión y confiabilidad de los datos, superando los valores de correlación originales.

La comparación de los mapas de calor y las tablas de correlación demuestra que la corrección automática de datos erróneos no solo mejora la precisión individual de cada estación, sino también la coherencia regional de la red meteorológica. Esto es especialmente relevante en zonas de alta variabilidad climática como la Sierra de Puno, donde la calidad de los datos es crucial para estudios climáticos y aplicaciones agrícolas.

Es importante destacar que el modelo mostró robustez incluso en estaciones con escasos datos válidos (Estaciones Automáticas), gracias a la integración de información espacial y temporal, así como a la selección automática de estaciones vecinas más correlacionadas.

VI. CONCLUSIONES

A partir de la implementación y evaluación del modelo matemático híbrido (Random Forest y Gradient Boosting) para el control de calidad, corrección y relleno de datos de temperatura del aire en la sierra de Puno (2001–2024), se desprenden las siguientes conclusiones fundamentales:

6.1. Conclusión general

El modelo matemático híbrido basado en Random Forest y Gradient Boosting demostró ser una solución validada y eficaz para el control de calidad, corrección y relleno de datos de temperatura del aire en la compleja región altoandina de Puno. Su implementación exitosa no solo permite subsanar las deficiencias históricas de los registros, como datos faltantes, atípicos y errores sistemáticos, sino que también proporciona una base de datos más fiable y robusta, indispensable para la modelización climática, la gestión de recursos hídricos y la toma de decisiones en sectores vulnerables como la agricultura.

6.2. Conclusiones específicas

6.2.1. Cuantificación de la incertidumbre de los datos

El modelo híbrido redujo significativamente la incertidumbre en los datos de temperatura del aire. El análisis estadístico evidenció que el 16.3% de los datos fueron corregidos y rellenados, logrando una reducción de la desviación estándar en el 100% de las estaciones al comparar los datos corregidos con los datos originales. La prueba de t-Student aplicada a 95% de las estaciones mostró valores de $p < .05$, validando la efectividad del modelo para disminuir significativamente la incertidumbre en las series temporales de temperatura. Asimismo, el 95% de las estaciones reportó un coeficiente de correlación (r) mayor a 0.9, confirmando la hipótesis H1 y demostrando la precisión del modelo en la corrección y relleno de datos meteorológicos.

6.2.2. Consistencia espacial de los datos corregidos

El análisis de correlación espacial demostró que el modelo mejoró la consistencia entre estaciones vecinas. El 62% de las estaciones reportó un incremento en la correlación espacial al comparar los datos originales con los datos corregidos y rellenados. Este porcentaje se incrementa a 72% al excluir las estaciones automáticas del análisis, ya que las estaciones convencionales mostraron mayor mejora en la correlación espacial. Los resultados confirman la hipótesis H2, evidenciando que el modelo híbrido optimiza la confiabilidad de los datos y fortalece la coherencia espacial de las series de temperatura en la región de estudio.

6.2.3. Eficacia del algoritmo del modelo matemático

La validación cruzada y el test de Friedman confirmaron la robustez y consistencia del modelo propuesto. Las métricas de evaluación mostraron valores estables en los cinco folds: error absoluto medio (MAE) de 1.20 °C (rango: 1.19-1.22 °C), raíz del error cuadrático medio (RMSE) de 1.55 °C (rango: 1.54-1.57 °C) y coeficiente de determinación (R^2) de 0.73 (rango: 0.73-0.74). El test de Friedman arrojó valores de $p = 0.406$ para las tres métricas ($p > 0.05$), indicando que no existen diferencias estadísticamente significativas entre los folds y que el modelo no presenta sobreajuste. Estos resultados confirman que el modelo es estadísticamente consistente en distintas particiones de datos y respaldan su aplicabilidad en condiciones heterogéneas, garantizando precisión y confiabilidad en sistemas de control de calidad de datos meteorológicos para la toma de decisiones ambientales y climáticas.

VII. RECOMENDACIONES

Con base en los hallazgos y conclusiones de la presente investigación, se formulan las siguientes recomendaciones orientadas a optimizar la gestión de datos meteorológicos en la sierra del departamento de Puno y en otras regiones con características similares:

7.1. Recomendaciones para la implementación operativa del modelo

Se recomienda al Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) implementar el modelo matemático híbrido propuesto en sus sistemas operativos de control de calidad de datos meteorológicos, priorizando las estaciones ubicadas en zonas altoandinas donde la incertidumbre de los datos es mayor. La integración del modelo en plataformas automatizadas permitiría la detección y corrección de errores en tiempo real, mejorando significativamente la confiabilidad de los pronósticos meteorológicos y la información climática destinada a sectores estratégicos como la agricultura, la gestión hídrica y la prevención de desastres.

7.2. Recomendaciones para el fortalecimiento de la red de estaciones meteorológicas

Se recomienda priorizar la calibración y mantenimiento preventivo de las estaciones meteorológicas en la región, especialmente aquellas que presentaron mayor incidencia de datos erróneos. Dado que las estaciones convencionales mostraron mayor mejora en la correlación espacial tras la aplicación del modelo (incremento del 72% excluyendo estaciones automáticas), resulta fundamental implementar protocolos estandarizados de verificación y ajuste periódico de los instrumentos de medición, particularmente en estaciones ubicadas por encima de los 3,800 m s. n. m., donde las condiciones climáticas extremas aceleran el deterioro del equipo.

7.3. Recomendaciones para la capacitación del personal técnico

Se recomienda desarrollar programas de capacitación continúa dirigidos a los observadores meteorológicos y técnicos responsables de las estaciones en la sierra de Puno. La formación debe enfocarse en la correcta lectura de instrumentos analógicos, la identificación temprana de valores atípicos, el uso de herramientas digitales de control de calidad y la interpretación de resultados del modelo matemático. Esta capacitación contribuiría a reducir los errores humanos en la transcripción de datos y mejoraría la calidad de la información desde su origen.

7.4. Recomendaciones para la expansión del modelo a otras variables meteorológicas

Se recomienda adaptar y validar el modelo matemático híbrido para otras variables meteorológicas críticas en la región, tales como precipitación, humedad relativa, velocidad del viento y radiación solar. La aplicación del modelo a múltiples variables permitiría generar bases de datos meteorológicas más completas y confiables, fortaleciendo la capacidad de análisis climático integral y mejorando la precisión de los modelos predictivos en la región altoandina.

7.5. Recomendaciones para la investigación futura

Se recomienda realizar investigaciones complementarias que evalúen el desempeño del modelo en otras regiones del Perú con características geográficas y climáticas diferentes, tales como la costa árida, la selva amazónica o las zonas de transición entre regiones naturales. Asimismo, se sugiere explorar la integración de técnicas avanzadas de deep learning, como las redes neuronales recurrentes (LSTM) o los transformers, que podrían mejorar la capacidad predictiva del modelo en series temporales largas. También se recomienda investigar la incorporación de datos de teledetección satelital para complementar y validar las mediciones terrestres, especialmente en zonas remotas con escasa cobertura de estaciones meteorológicas.

7.6. Recomendaciones para la gestión de datos y accesibilidad

Se recomienda al SENAMHI implementar repositorios digitales de acceso abierto que permitan a investigadores, instituciones académicas y tomadores de decisiones consultar las series de temperatura corregidas y validadas mediante el modelo propuesto. La disponibilidad de datos de alta calidad facilitaría el desarrollo de investigaciones climáticas, la elaboración de políticas públicas basadas en evidencia y la implementación de estrategias de adaptación al cambio climático en comunidades vulnerables de la sierra de Puno.

7.7. Recomendaciones para el monitoreo continuo del modelo

Se recomienda establecer un sistema de monitoreo y evaluación continua del desempeño del modelo matemático híbrido, incluyendo auditorías periódicas de las métricas estadísticas (MAE, RMSE, R^2) y la actualización del algoritmo conforme se incorporen nuevos datos históricos. Este monitoreo permitiría identificar oportunamente desviaciones en el rendimiento del modelo, ajustar parámetros según las condiciones cambiantes del clima y garantizar la vigencia de la herramienta ante posibles escenarios de variabilidad climática extrema asociados al cambio climático global.

7.8. Recomendaciones para la articulación interinstitucional

Se recomienda promover alianzas estratégicas entre el SENAMHI, universidades, centros de investigación y organismos internacionales especializados en meteorología y climatología, con el propósito de compartir experiencias, validar metodologías y fortalecer las capacidades técnicas para la gestión de datos meteorológicos en regiones de alta montaña. La cooperación interinstitucional facilitaría la transferencia de tecnología, el acceso a recursos técnicos avanzados y el posicionamiento del modelo propuesto como referente para otras regiones altoandinas de América del Sur.

VIII. REFERENCIAS

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C. y Wieringa, J. (2003). *Guidelines on climate metadata and homogenization* (WCDMP-N.° 53, WMO-TD N.° 1186). Organización Meteorológica Mundial.

Barnett, V. y Lewis, T. (1994). *Outliers in statistical data* (3.ª ed.). Wiley.

Biau, G. y Scornet, E. (2016). A random forest guided tour. *TEST, 25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

Box, G. E. P., Jenkins, G. M., Reinsel, G. C. y Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5.ª ed.). Wiley.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Buytaert, W., Célleri, R., De Bièvre, B., Cisneros, F., Wyseure, G., Deckers, J. y Hofstede, R. (2006). Human impact on the hydrology of the Andean páramos. *Earth-Science Reviews, 79*(1-2), 53–72. https://doi.org/10.1016/j.earscirev.2006.06.002

Chandola, V., Banerjee, A. y Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys, 41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

Chatfield, C. (2004). *The analysis of time series: An introduction* (6.ª ed.). Chapman & Hall/CRC.

Chen, T. y Guestrin, C. (2016). XGBoost: A scalable tree boosting system. En *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Conover, W. J. (1999). *Practical nonparametric statistics* (3.ª ed.). Wiley.

Copernicus Climate Change Service [C3S]. (2021). Evaluation and quality control (EQC) for the Climate Data Store (Informe C3S2_520). Centro Europeo de Previsiones

Meteorológicas a Plazo Medio. https://climate.copernicus.eu/quality-assurance-climate-data-store

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4.ª ed.). SAGE.

European Climate Assessment & Dataset [ECA&D]. (2007). *Algorithm theoretical basis document (ATBD)*. Real Instituto Meteorológico de los Países Bajos. https://cdn.knmi.nl/system/data_center_publications/files/000/067/636/original/ecad_atbd_copy1.pdf

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

García, M., López, R. y Torres, J. (2023). Impacto de las heladas no pronosticadas en cultivos andinos. *Revista de Ciencias Agrarias, 40*(1), 20–30.

Gómez, P. y Lima, R. (2021). Data gaps in remote Andean weather stations: Challenges for long-term climate monitoring. *Journal of Mountain Research, 15*(2), 50–60.

Goodfellow, I., Bengio, Y. y Courville, A. (2016). *Deep learning*. MIT Press.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Oxford University Press.

Hastie, T., Tibshirani, R. y Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2.ª ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill.

Instituto de Hidrología, Meteorología y Estudios Ambientales [IDEAM]. (2007). Control de calidad automático de datos meteorológicos (Convenio interadministrativo IDEAM–FOPAE N.° 198/04 y 550/04).

Instituto de Hidrología, Meteorología y Estudios Ambientales [IDEAM]. (2019). *Glosario meteorológico*. https://bart.ideam.gov.co/indiecosistemas/ooee/variables_meteorologicas/GlosarioMeteorologico.pdf.

Instituto Nacional de Investigación en Glaciares y Ecosistemas de Montaña [INAIGEM]. (2021). *Informe sobre variaciones de temperatura en la región de Puno (1980–2020)*.

James, G., Witten, D., Hastie, T. y Tibshirani, R. (2013). *An introduction to statistical learning with applications in R*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

James, G., Witten, D., Hastie, T. y Tibshirani, R. (2021). *An introduction to statistical learning with applications in R* (2.ª ed.). Springer.

Kuhn, M. y Johnson, K. (2013). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Kutner, M. H., Nachtsheim, C. J., Neter, J. y Li, W. (2005). *Applied linear statistical models* (5.ª ed.). McGraw-Hill Irwin.

Ley N.° 28303. Ley Marco de Ciencia, Tecnología e Innovación Tecnológica (27 de julio de 2004). Diario Oficial El Peruano. https://busquedas.elperuano.pe/

Ministerio de Agricultura y Riego [MINAGRI]. (2023). *Informe de pronósticos agrícolas en Puno*.

Ministerio del Ambiente [MINAM]. (2015). *Política Nacional de Cambio Climático*.

Montgomery, D. C., Peck, E. A. y Vining, G. G. (2012). *Introduction to linear regression analysis* (5.ª ed.). Wiley.

Montgomery, D. C., Peck, E. A. y Vining, G. G. (2021). *Introduction to linear regression analysis* (6.ª ed.). Wiley.

Montiel, D. (2017). Modelo matemático para la evaluación de peligros naturales mediante información geográfica. *Télématique, 16*(1), 60–80. https://ojs.urbe.edu/index.php/telematique/article/view/1832

NASA Science Editorial Team. (2021). *Global temperature data accuracy and challenges*. NASA. https://climate.nasa.gov/

Organización Meteorológica Mundial [OMM]. (2010). *Guía del Sistema Mundial de Observación* (OMM-N.° 488). https://library.wmo.int/

Organización Meteorológica Mundial [OMM]. (2011). *Guide to climatological practices* (3.ª ed., OMM-N.° 100). https://library.wmo.int/

Organización Meteorológica Mundial [OMM]. (2017). Directrices de la Organización Meteorológica Mundial sobre el cálculo de las normales climáticas (OMM-N.° 1203). https://library.wmo.int/viewer/60338

Organización Meteorológica Mundial [OMM]. (2018a). *Guide to meteorological instruments and methods of observation* (OMM-N.° 8). https://library.wmo.int/

Organización Meteorológica Mundial [OMM]. (2018b). *Guidelines on quality management in climate services* (OMM-N.° 1221). https://library.wmo.int/idurl/4/56255

Organización Meteorológica Mundial [OMM]. (2021). *Guidelines on surface station data quality control and quality assurance for climate applications* (OMM-N.° 1269). https://library.wmo.int/

Organización Meteorológica Mundial [OMM]. (2023). *Manual on the high-quality global data management framework for climate* (OMM-N.° 1238). https://library.wmo.int/

Pérez, L. (2019). *Errores en la recolección de datos meteorológicos en regiones altoandinas* [Tesis de maestría]. Universidad Nacional del Altiplano.

Pineda, E. B., de Alvarado, E. L. y de Canales, F. H. (1994). *Metodología de la investigación: Manual para el desarrollo de personal de salud* (2.ª ed.). Organización Panamericana de la Salud.

Polit, D. F. y Beck, C. T. (2006). *Essentials of nursing research: Methods, appraisal, and utilization* (6.ª ed.). Lippincott Williams & Wilkins.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J. y Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature, 566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Rubio, J. (2008). *Discrepancias en la lectura de termómetros analógicos en estaciones de Puno* [Informe técnico]. Servicio Nacional de Meteorología e Hidrología del Perú.

Servicio Nacional de Meteorología e Hidrología del Perú [SENAMHI]. (2017). *Informe técnico sobre errores en la lectura de termómetros en estaciones de los Andes centrales*.

Servicio Nacional de Meteorología e Hidrología del Perú [SENAMHI]. (2020). *Protocolos de instalación y mantenimiento de estaciones meteorológicas*.

Servicio Nacional de Meteorología e Hidrología del Perú [SENAMHI]. (2021). *Control de calidad de datos de estaciones meteorológicas e hidrológicas automáticas en el Centro de Procesamiento de Datos del SENAMHI* (Manual técnico MT-DRD-002). https://www.senamhi.gob.pe/load/file/00711SENA-95.pdf

Servicio Nacional de Meteorología e Hidrología del Perú [SENAMHI]. (2022). *Reporte de recalibración de estaciones meteorológicas en Puno*.

Smith, J., Brown, T. y Wilson, P. (2020). Limitations of QA/QC systems in high-altitude meteorological stations. *Climate Data Journal, 12*(3), 70–85.

Taylor, J. R. (1997). *An introduction to error analysis: The study of uncertainties in physical measurements* (2.ª ed.). University Science Books.

Villanueva, R. (2022). *Errores en datos de temperatura y su impacto en modelos climáticos en regiones altoandinas* [Tesis de doctorado]. Universidad Nacional Mayor de San Marcos.

Vuille, M., Carey, M., Huggel, C., Buytaert, W., Rabatel, A., Jacobsen, D., Soruco, A., Villacis, M., Yarleque, C., Elison Timm, O., Condom, T., Salzmann, N. y Sicart, J. E. (2018). Rapid decline of snow and ice in the tropical Andes: Impacts, uncertainties and challenges ahead. *Earth-Science Reviews, 176*, 195–213.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3.ª ed.). Academic Press.

Wilks, D. S. (2019). *Statistical methods in the atmospheric sciences* (4.ª ed.). Elsevier.

Zhang, K., Li, X. y Zhou, Y. (2022). Correcting erroneous temperature data using ensemble methods: A case study in high-altitude climates. Weather and Climate Extremes, 35, 100410. https://doi.org/10.1016/j.wace.2022.100410

IX. ANEXOS

Los anexos se organizan en cinco anexos que reflejan las etapas metodológicas de la investigación. Cada fase incluye una descripción que contextualiza su propósito, seguida de una tabla resumen que detalla los anexos correspondientes, su número de identificación (Tabla, Ficha, Mapa, Gráfico o Diagrama), descripción, justificación y el objetivo al que contribuyen.

Anexo A Información transversal a todas las fases

Descripción:

Esta fase agrupa los instrumentos y validaciones metodológicas que sustentan la rigurosidad y coherencia del proceso investigativo. Incluye matrices de consistencia y operacionalización, así como la validación de instrumentos mediante juicio de expertos y pruebas de confiabilidad, asegurando la solidez metodológica en todas las etapas del estudio.

Tabla 9.

Relación de tablas y figuras del anexo A

N.º Ítem	Descripción	Justificación	Objetivo
Tabla 10	Matriz de Consistencia	Garantiza coherencia entre el problema, objetivos, hipótesis, variables e instrumentos.	Alinear los componentes metodológicos de la investigación.
Tabla 11	Matriz de Operacionalización	Define variables, dimensiones, indicadores e Instrumentos.	Precisar la forma de medición de cada variable.
Tabla 12	Formato de registro de datos diarios meteorológicos	Se describe el contenido del formato de registro de datos diarios meteorológicos	Formato para registrar los datos meteorológicos descargados desde la plataforma oficial de descarga del SENAMHI
Tabla 13	Informe de juicio de experto del del instrumento de Investigación	Validar el instrumento aplicado mediante expertos.	Asegurar confiabilidad y validez del Instrumento en la recolección y Resultados de datos.
Tabla 14	Validación del instrumento Experto 1	Describe la evaluación para la validación del instrumento durante la recolección de datos y su relación con los resultados del modelo del Experto 1	Valida la fiabilidad de los datos y resultados en todas las fases.
Tabla 15	Validación del instrumento Experto 2	Describe la evaluación para la validación del instrumento durante la recolección de datos y su relación con los resultados del modelo del Experto 2	Valida la fiabilidad de los datos y resultados en todas las fases.

Tabla 16	Validación del instrumento Experto 3	Describe la evaluación para la validación del instrumento durante la recolección de datos y su relación con los resultados del modelo del Experto 3	Valida la fiabilidad de los datos y resultados en todas las fases.
Tabla 17	Validación del instrumento Experto 4	Describe la evaluación para la validación del instrumento durante la recolección de datos y su relación con los resultados del modelo del Experto 4	Valida la fiabilidad de los datos y resultados en todas las fases.
Tabla 18	Resultados de evaluación de expertos	Se realiza la evaluación de expertos	Valida la evaluación de expertos.
Tabla 19	Método de evaluación de resultados de experto	Se realiza el método de evaluación de experto	Valida el método de evaluación de experto.
Tabla 20	Prueba test–retest	Se realiza la prueba test - retest	Valida la prueba de test – retest.
Figura 4	Coeficiente de correlación	Se realiza el coeficiente de correlación	Validad el coeficiente de correlación
Tabla 21	Resultado de la prueba de test retest	Se realiza el resultado de la prueba de test retest	Valida el resultado de la prueba de test retest
Tabla *22*	Categoría de la prueba	Se realiza la categoría de la prueba	Valida la categoría de la prueba

Nota. Elaboración propia.

Tabla 10.

Matriz de consistencia

Problema general	Objetivo general	Hipótesis general	Variables	Metodología
¿En qué medida el modelo matemático, basado en técnicas de inteligencia artificial (Random Forest + Gradient Boosting), estadísticas y geoespaciales, mejora el control de calidad para la corrección y rellenos de datos erróneos de temperatura del aire en las estaciones meteorológicas de la Sierra de Puno?	Determinar en qué medida el modelo matemático, basado en técnicas de Machine Learning, estadísticas y geoespaciales, mejora el control de calidad para la corrección y relleno de datos erróneos de temperatura del aire en las estaciones meteorológicas de la Sierra de Puno.	El modelo matemático, basado en inteligencia artificial (Random Forest + Gradient Boosting), técnicas estadísticas y geoespaciales, mejora significativamente la precisión ($p < .05$) y la confiabilidad (% r Rellenado > % r Original) de los datos de temperatura en las estaciones meteorológicas de la Sierra de Puno.	X = Variable Independiente (VI): Modelo matemático (Sistema de detección de anomalías + Modelo híbrido (Random Forest + Gradient Boosting) para corrección y relleno de datos de temperatura) Y = Variable Dependiente (VD): Calidad de Calidad (medida como: - Incertidumbre: p-valor < 0.05, desviación estándar. - Consistencia espacial: Correlación $r > .9$).	- Tipo de estudio: Cuantitativo, Cuasiexperimental. - Población: Red de Estaciones meteorológicas de la sierra de Puno con registros diarios de Tmed (2001-2024). - Muestra: Conjunto de estaciones seleccionadas mediante muestreo estratificado por sector climática.

Problemas específicos	Objetivos específicos	Hipótesis específicas	
P1: ¿En qué medida el modelo matemático, basado en técnicas de inteligencia artificial (Random Forest + Gradient Boosting), estadísticas y geoespaciales, influye en la reducción de la incertidumbre de los datos de temperatura para mejorar la precisión en el control de calidad de las estaciones meteorológicas de la Sierra de Puno? **P2:** ¿En qué medida el modelo matemático basado en análisis geoespacial influye en el mejoramiento de la consistencia espacial de los datos de temperatura entre estaciones vecinas para fortalecer la confiabilidad en el control de calidad?	**O1:** Determinar en qué medida el modelo matemático, basado en técnicas de inteligencia artificial (Random Forest + Gradient Boosting), estadísticas y geoespaciales mejora la precisión en el control de calidad, evaluando la reducción de la incertidumbre en las series de temperatura de las estaciones de Puno mediante métricas estadísticas (desviación estándar) y pruebas de significancia estadística ($p < .05$). **O2:** Determinar en qué medida el modelo matemático basado en técnicas geoespaciales mejora la confiabilidad en el control de calidad, evaluando el aumento en la consistencia espacial de los datos de temperatura mediante correlaciones significativas (% r Rellenado > % r Original, $r > .9$) entre estaciones vecinas.	**H1:** El modelo matemático, basado en inteligencia artificial (Random Forest + Gradient Boosting), técnicas estadísticas, reduce significativamente ($p < .05$) la incertidumbre en las series de temperatura de las estaciones de Puno, mejorando la precisión en el control de calidad. **H2:** El modelo matemático basado en técnicas geoespaciales incrementa significativamente la consistencia espacial de los datos de temperatura, evidenciada por una correlación mejorada (% r Rellenado > % r Original, $r > .9$) entre estaciones vecinas tras la corrección, mejorando la confiabilidad en el control de calidad.	- Análisis estadístico de series temporales. - Implementación de modelo híbrido (Random Forest + Gradient Boosting). - Validación con métricas estadísticas (p-valor, desviación estándar, correlación). - Detección de anomalías estadísticas. - Aplicación del modelo híbrido para corrección. - Validación con pruebas estadísticas (t-test, desviación estándar). - Análisis geoespacial (interpolación, correlación espacial). - Validación con coeficiente de correlación de Pearson entre estaciones vecinas.

Nota. Elaboración propia

Tabla 11.

Matriz de operacionalización de variables

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador	Instrumentos	Unidad de medición
X = Variable independiente (VI): modelo matemático (Sistema de detección de anomalías + modelo híbrido (random forest + gradient boosting) para corrección y relleno de datos de temperatura)	Conjunto de algoritmos y técnicas estadísticas y geoespaciales diseñados para detectar, corregir y rellenar datos erróneos de temperatura del aire, optimizando la precisión y confiabilidad de las series temporales en estaciones meteorológicas.	Implementación de un modelo híbrido que combina Random Forest y Gradient Boosting para la detección de anomalías, corrección de errores sistemáticos y relleno de datos faltantes en series de temperatura media diaria (2001-2024) de 22 estaciones meteorológicas en la Sierra de Puno, con un componente geoespacial para garantizar consistencia espacial.	Eficacia del algoritmo	I1 - Precisión del modelo matemático: (Validación Cruzada(k-fold), Test Friedman)	Los datos fueron recolectados desde la plataforma oficial del SENAMHI, los datos recolectados se almacenaron en un formado de registro de datos diarios del SENAMHI	Razón (numérica continua).
				I2 - Parámetros optimizados del modelo matemático (ej.: hiperparámetros en Random Forest/Boosting).		Razón (numérica discreta) y nominal(categóricos)
Y = Variable dependiente (VD): calidad de calidad (medida como: - Incertidumbre: *p-valor* < .05, desviación estándar. - Consistencia espacial: Correlación *r* > .9).	Proceso de evaluación y mejora de la calidad de los datos de temperatura del aire mediante la reducción de incertidumbre y el aumento de la consistencia espacial, asegurando que los datos sean precisos y confiables para análisis climáticos.	Medida de la calidad de los datos de temperatura media diaria corregidos y rellenados por el modelo matemático, evaluada a través de la incertidumbre (*p-valor* <.05, desviación estándar) y la consistencia espacial (correlación *r* > .9 entre estaciones vecinas).	Precisión: (Reducción de Incertidumbre)	I3 - Reducción de la desviación estándar de las series antes/después de la corrección.		Razón (numérica continua).
			Confiabilidad (consistencia espacial)	I4 - Correlación de Pearson (% r post-corrección > % r original) entre estaciones vecinas.		Razón (numérica continua).

Nota. Elaboración propia

Tabla 12.

Formato de registro de datos diarios meteorológicos

DIRECCION DE REDES DE OBSERVACION Y DATOS

ESTACION: NOMBRE /CODIGO / REGION LAT.: XX° XX´ "S"

PARAMETRO: PRECIPITACION TOTAL DIARIA (mm) LONG.: XX° XX´ "W"

ALT.: XXXX msnm

AÑO:

DIA	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SET	OCT	NOV	DIC
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
INFORMACION PREPARADA PARA									Lima, xx de xxxxx de xxxx			

Nota. SENAMHI (2024)

Tabla 13.

Informe de juicio de experto del instrumento de investigación

Título de la tesis:

Presentado por:

I. DATOS DEL EXPERTO:

1.1. Apellidos y nombres:

1.2. Grado académico:

1.4. Instrumento de evaluación:

INDICADORES	CRÉDITO	Deficiente 0 - 20%	Regular 21 -40 %	Bueno 41 - 80 %	Muy Bueno 61 - 80 %	Excelente 81 - 100 %
1. Claridad	Está formulado con lenguaje apropiado					X
2. Objetividad	Está expresado en conducta observable				X	
3. Actualidad	Es adecuado el avance de la ciencia y tecnología				X	
4. Organización	Existe organización lógica					X
5. Suficiencia	Comprende los aspectos de cantidad y calidad					X
6. Intencionalidad	Adecuado para valorar aspectos del sistema metodológico y científico					X
7. Consistencia	Se basa en aspectos teóricos, científicos acordes a la tecnología educativa				X	
8. Coherencia	Entre índices, indicadores y dimensiones				X	
9. Metodología	Responde al propósito del trabajo bajo los objetivos a lograr				X	

II. OPCION DE APLICABILIDAD:

III. PROMEDIO DE VALORACION:

IV. RECOMENDACIONES.

Fecha:

Firma del Experto

Tabla 14.

Ficha de validación del instrumento experto 1

VALIDACIÓN DE INSTRUMENTO

- Titulo de la Tesis : Modelo matemático ante datos erróneos de temperatura del aire para el control de calidad en la sierra del departamento de puno del Perú
- Presentado por : Olimpio Solís Cáceres

I. DATOS DEL EXPERTO:

1.1. Apellidos y Nombres: Ordoñez Galvez, Juan Julio

1.2. Grado Académico: Doctor En Medio Ambiente y Desarrollo Sostenible (U. Nacional Federico Villarreal - Perú)

1.4. Instrumento de Evaluación: Encuesta sobre la recolección de datos y resultados del modelo matemático

INDICADORES	CRÉDITO	Deficiente 0 - 20 %	Regular 21 - 40 %	Bueno 41 - 80 %	Muy Bueno 61 - 80 %	Excelente 81 - 100 %
1. Claridad	Está formulado con lenguaje apropiado				X	
2. Objetividad	Está expresado en conducta observable					X
3. Actualidad	Es adecudo el avance de la ciencia y tecnologia					X
4. Organización	Existe organización lógica				X	
5. Suficiencia	Comprende los aspecto de cantidad y calidad					X
6. Intencionalidad	Adecuado para valorar aspectos del sistema metodológico y cientifico				X	
7. Consistencia	Se basa en aspectos teoricos, cientificos acorde a la tecnologia educativa					X
8. Coherencia	Entre indices, indicadores y dimenciones					X
9. Metodologia	Responde al propósito del trabajo bajo los objetivos a lograr				X	

II. OPCION DE APLICABILIDAD: Se puede Aplicar

III. PROMEDIO DE VALORACION: 90.12 **Nivel :** Excelente

IV. RECOMENDACIONES. Se recomienda realizar 2 Mapas: Correlación Espacial Original versus Correlación Espacial Corregida o Rellenado para un mejor analisis de las mejoras del modelo

---

Fecha: 7/06/2025



Firma del Experto

Tabla 15.

Ficha de validación del instrumento experto 2

VALIDACIÓN DE INSTRUMENTO

- Titulo de la Tesis : Modelo matemático ante datos erróneos de temperatura del aire para el control de calidad en la sierra del departamento de puno del Perú
- Presentado por : Olimpio Solís Cáceres

I. DATOS DEL EXPERTO:

1.1. Apellidos y Nombres: Lavado Casimiro, Waldo Sven

1.2. Grado Académico: Doctor en Hidrología, Hidroquímica, Suelos y Medio Ambiente (U. Toulouse III - Francia)

1.4. Instrumento de Evaluación: Encuesta sobre la recolección de datos y resultados del modelo matemático

INDICADORES	CRÉDITO	Deficiente 0 - 20 %	Regular 21 - 40 %	Bueno 41 - 80 %	Muy Bueno 61 - 80 %	Excelente 81 - 100 %
1. Claridad	Está formulado con lenguaje apropiado					X
2. Objetividad	Está expresado en conducta observable				X	
3. Actualidad	Es adecudo el avance de la ciencia y tecnologia					X
4. Organización	Existe organización lógica					X
5. Suficiencia	Comprende los aspecto de cantidad y calidad				X	
6. Intencionalidad	Adecuado para valorar aspectos del sistema metodológico y cientifico				X	
7. Consistencia	Se basa en aspectos teoricos, cientificos acorde a la tecnologia educativa					X
8. Coherencia	Entre indices, indicadores y dimenciones				X	
9. Metodologia	Responde al propósito del trabajo bajo los objetivos a lograr					X

II. OPCION DE APLICABILIDAD: Se puede Aplicar

III. PROMEDIO DE VALORACION: 90.12 **Nivel :** Excelente

IV. RECOMENDACIONES. Se recomienda realizar un Gráfico de Distribución de Residuales,para demostrar que los errores son aleatorios y no presentan patrones sistemáticos o sesgos

---

Fecha: 10/01/2025



Firma del Experto

Tabla 16.

Ficha de validación del instrumento experto 3

VALIDACIÓN DE INSTRUMENTO

- Titulo de la Tesis : Modelo matemático ante datos erróneos de temperatura del aire para el control de calidad en la sierra del departamento de puno del Perú
- Presentado por : Olimpio Solís Cáceres

I. DATOS DEL EXPERTO:

1.1. Apellidos y Nombres: Paredes Riveros, Marco Antonio

1.2. Grado Académico: Maestro en Ciencias en Gestión Ambiental (U. Amazonía Peruana – Perú)

1.4. Instrumento de Evaluación: Encuesta sobre la recolección de datos y resultados del modelo matemático

INDICADORES	CRÉDITO	Deficiente 0 - 20 %	Regular 21 - 40 %	Bueno 41 - 80 %	Muy Bueno 61 - 80 %	Excelente 81 - 100 %
1. Claridad	Está formulado con lenguaje apropiado					X
2. Objetividad	Está expresado en conducta observable					X
3. Actualidad	Es adecudo el avance de la ciencia y tecnologia				X	
4. Organización	Existe organización lógica				X	
5. Suficiencia	Comprende los aspecto de cantidad y calidad					X
6. Intencionalidad	Adecuado para valorar aspectos del sistema metodológico y cientifico					X
7. Consistencia	Se basa en aspectos teoricos, cientificos acorde a la tecnologia educativa			X		
8. Coherencia	Entre indices, indicadores y dimenciones					X
9. Metodologia	Responde al propósito del trabajo bajo los objetivos a lograr					X

II. OPCION DE APLICABILIDAD: Se puede Aplicar

III. PROMEDIO DE VALORACION: 91.36 Nivel : Excelente

IV. RECOMENDACIONES. Se recomienda realizar una comparación de los datos de una estación convencional versus una estación automática ubicadas en el mismo lugar con la finalidad de validar la calidad y calibración del instrumento de registro de datos

---

Fecha: 20/01/2025



Tabla 17.

Ficha de validación del instrumento experto 4

VALIDACIÓN DE INSTRUMENTO

- Titulo de la Tesis : Modelo matemático ante datos erróneos de temperatura del aire para el control de calidad en la sierra del departamento de puno del Perú
- Presentado por : Olimpio Solís Cáceres

I. DATOS DEL EXPERTO:

1.1. Apellidos y Nombres: Flores Sancho, Sixto

1.2. Grado Académico: Maestro en Meteorología Aplicada, Business Innovation: Competencias y Herramientas Avanzadas para la Generación de Innovación (U.Politécnica de Valencia - España)

1.4. Instrumento de Evaluación: Encuesta sobre la recolección de datos y resultados del modelo matemático

INDICADORES	CRÉDITO	Deficiente 0 - 20 %	Regular 21 - 40 %	Bueno 41 - 80 %	Muy Bueno 61 - 80 %	Excelente 81 - 100 %
1. Claridad	Está formulado con lenguaje apropiado				X	
2. Objetividad	Está expresado en conducta observable					X
3. Actualidad	Es adecudo el avance de la ciencia y tecnologia				X	
4. Organización	Existe organización lógica				X	
5. Suficiencia	Comprende los aspecto de cantidad y calidad					X
6. Intencionalidad	Adecuado para valorar aspectos del sistema metodológico y cientifico					X
7. Consistencia	Se basa en aspectos teoricos, cientificos acorde a la tecnologia educativa					X
8. Coherencia	Entre indices, indicadores y dimenciones					X
9. Metodologia	Responde al propósito del trabajo bajo los objetivos a lograr				X	

II. OPCION DE APLICABILIDAD: Se puede Aplicar

III. PROMEDIO DE VALORACION: 90.12 **Nivel :** Excelente

IV. RECOMENDACIONES. Se recomienda realizar gráficos de indicadores de detección de anomalías, corrección y rellenos de datos para validar y comunicar el desempeño de tu modelo matemático en cada etapa del control de calidad de datos

---

Fecha: 17/01/2025



Firma del Experto

Método de evaluación de resultados validación de instrumento

Para evaluar el grado de consenso entre los expertos en la validación del instrumento, se aplicó un **análisis de consistencia porcentual**. Este método considera que existe consenso cuando la mayoría de las calificaciones de los expertos se encuentran dentro de un rango de ±1 punto respecto a la media del ítem, en una escala Likert de 1 a 5. Se desarrollo un script desarrollado en Python para desarrollar el proceso de calificación final sobre los resultados de evaluación de los Expertos, a continuación, se explica los procedimientos del método análisis de consistencia porcentual seguido:

Paso 01: Cargar los datos del resultado de evaluación de los expertos

Se carga al sistema los datos los resultados de evaluación de Expertos mediante el formato de

Tabla 18.

Resultados de evaluación de expertos

Ítem	Experto_1	Experto_2	Experto_3	Experto_4
1_Claridad	4	5	5	4
2_Objetividad	5	4	5	5
3_Actualidad	5	5	4	4
4_Organizacion	4	5	4	4
5_Suficiencia	5	4	5	5
6_Intencionalidad	4	4	5	5
7_Consistencia	5	5	3	5
8_Coherencia	5	4	5	5
9_Metodologia	4	5	5	4

Nota. Elaboración propia.

Paso 02: Procesamiento de cálculos

- **Transformación a escala porcentual**

Los puntajes originales (escala Likert de 1 a 5) se convierten a una escala de 0-100% usando la fórmula:

Formula: Valor Porcentual = Pi = (x−1) x 25

Donde: Pi: Porcentaje para el ítem i (escala 0-100%).

x: Puntaje original del experto para el ítem i (escala Likert 1-5).

Explicación:

Esta fórmula convierte los puntajes likert en una escala porcentual.

Ejemplo: Un puntaje de 5 se convierte en (5−1) x 25=100%.

Esto estandariza las evaluaciones para alinearlas con la escala de la universidad.

- **Cálculo de estadísticos descriptivos**

Por cada ítem (fila), se calculan:

- **Media porcentual**

Promedio de los puntajes transformados de los cuatro expertos.

$$\bar{x} = \frac{1}{n}\sum_{j}^{n} Pj \qquad (5)$$

Donde:

$\bar{x}$: Media porcentual de las evaluaciones para un ítem.

Pj: Porcentaje transformado del experto j.

n: Número de expertos (4 en este caso).

Explicación: calcula el promedio de los puntajes transformados para cada ítem.

- **Desviación estándar: variabilidad de los puntajes transformados.**

$$\boldsymbol{\sigma P} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{n-1}}} \sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} (\boldsymbol{Pj - P})^{\mathbf{2}} \qquad (6)$$

Donde:

$\boldsymbol{\sigma P}$: Desviación estándar de los puntajes transformados.

Pj, $\bar{\boldsymbol{x}}$, n: Como se definió anteriormente.

Explicación: Mide la variabilidad de las evaluaciones porcentuales entre los expertos.

- **Mínimo y máximo**

Valores extremos de los puntajes transformados.

Estos estadísticos describen la percepción de los expertos sobre cada ítem (claridad, objetividad, etc.) y su variabilidad.

- **Clasificación según escala personalizada**

La media porcentual se categoriza según los rangos definidos:

- ✓ Deficiente: 0-20%
- ✓ Regular: 21-40%
- ✓ Bueno: 41-60%
- ✓ Muy Bueno: 61-80%
- ✓ Excelente: 81-100%

Esto permite interpretar la calidad de cada ítem según los criterios de la universidad.

- **Cálculo de consistencia porcentual**

Para cada ítem, se calcula la consistencia porcentual como:

Rango aceptable = [$\bar{\boldsymbol{x}}\boldsymbol{i}$ −1,$\bar{\boldsymbol{x}}\boldsymbol{i}$ +1]

$$Ci = \frac{n\ con\ Xij\ \in \left[\ \bar{\boldsymbol{x}}\boldsymbol{i}\ -\mathbf{1},\ \bar{\boldsymbol{x}}\boldsymbol{i}\ +\mathbf{1}\right] x100}{n} \qquad (7)$$

Donde:

n: Numero de Expertos

Ci: Consistencia Porcentual

$\bar{x}i$: es la media de las evaluaciones del ítem i

X_{ij} es la evaluación del ítem i por el experto j

Los resultados se expresan en porcentaje y se categorizan cualitativamente como Deficiente, Regular, Bueno, Muy Bueno o Excelente, según el valor de la media porcentual. Este enfoque es especialmente útil cuando el número de expertos es reducido, permitiendo una interpretación clara y visual del grado de acuerdo alcanzado (Polit & Beck, 2006).

Tabla 19.

Método de evaluación de resultados de experto

Ítem	Media porcentaje	Desviación estándar	Mínimo	Máximo	Categoría	Consistencia (%)
2_Objetividad	93.75	12.5	75	100	Excelente	100
5_Suficiencia	93.75	12.5	75	100	Excelente	100
8_Coherencia	93.75	12.5	75	100	Excelente	100
1_Claridad	87.5	14.43	75	100	Excelente	100
3_Actualidad	87.5	14.43	75	100	Excelente	100
6_Intencionalidad	87.5	14.43	75	100	Excelente	100
7_Consistencia	87.5	25	50	100	Excelente	75
9_Metodologia	87.5	14.43	75	100	Excelente	100
4_Organizacion	81.25	12.5	75	100	Excelente	100
Promedio	88.89	14.75	72.22	100.00	Excelente	97.22

Nota. Elaboración propia.

Ficha A7: Validación y confiabilidad de instrumentos de medición de los datos meteorológicos.

Para la adquisición de los datos se realizará de la plataforma que publica SENAMH. Dado que el SENAMHI por función es responsable de mantenimiento preventivo y conectivos de los Instrumentos o sensores que capturan los datos meteorológicos.

Para verificar la validación y confiabilidad del Instrumentos y/o sensores mencionados, se ha realizado los siguientes procedimientos:

Pruebas de normalidad de la serie

Para esta prueba se ha tomado como muestra una serie de datos de temperatura del aire medias mensuales de un año, de 2 estaciones meteorológicas de un mismo lugar (Estación PUNO) donde los datos se obtienen de 2 fuentes: una fuente convencional (Estación Convencional de PUNO) y la otra automática (Estación Automática de PUNO).

Como el tamaño de muestra es 12, y este valor es menor o igual que 50 entonces tomamos la prueba de Shapiro-Wilk

Donde:

p-valor: es el valor de la probabilidad

$\alpha = 0.05$ es el nivel de significancia

entonces:

p-valor (Tmed_CON) = 0.409

p-valor (Tmed_EMA) = 0.451

Cuadro 7.2.1:

Resumen de procesamiento de casos

	Casos					
	Válido		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
Tmed_CON	12	100,0%	0	0,0%	12	100,0%
TMed_EMA	12	100,0%	0	0,0%	12	100,0%

Pruebas de normalidad

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Tmed_CON	,182	12	,200*	,933	12	,409
TMed_EMA	,202	12	,190	,936	12	,451

*. Esto es un límite inferior de la significación verdadera.

a. Corrección de significación de Lilliefors

Regla de decisión

Si p-valor <= α se rechaza la hipótesis nula

Si p-valor >= α se acepta hipótesis nula

Del cuadro 7.2.1 se tiene:

p-valor (Tmed_CON) = 0.409

p-valor (Tmed_EMA) = 0.451

Como p-valor (Tmed_CON) y p-valor (Tmed_EMA) es mayor a 0.05 entonces se acepta la hipótesis nula, es decir los datos tienen una distribución normal

Prueba de confiabilidad test - retest

La confiabilidad (o consistencia) de un test es la precisión con que el test mide lo que mide, en una población determinada y en las condiciones normales de aplicación. (Anastasi,1982; Aiken,1995).

En este caso se tomó una muestra de la población en un espacio de tiempo y se midió la temperatura del aire en el mismo tiempo.

Para realizar la prueba de confiabilidad test-retest se calculó el coeficiente de correlación de la serie, antes se realizó la prueba de normalidad.

Tabla 20.

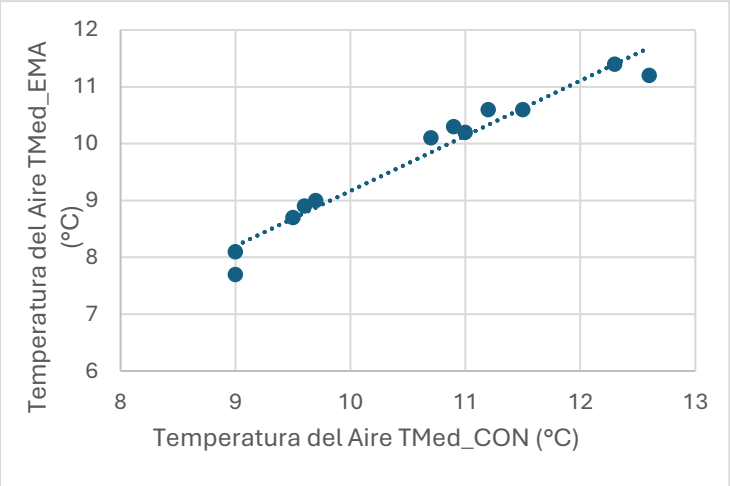
Prueba test–retest

Nro.	**Periodo**	**Estación meteorológica**	
		Test (Tmed_CON)	Retest (TMed_EMA)
1	2014-01	11	10.2
2	2014-02	11.5	10.6
3	2014-03	11.2	10.6
4	2014-04	10.7	10.1
5	2014-05	9.6	8.9
6	2014-06	9.5	8.7
7	2014-07	9	7.7
8	2014-08	9	8.1
9	2014-09	9.7	9
10	2014-10	10.9	10.3
11	2014-11	12.3	11.4
12	2014-12	12.6	11.2
Media		10.6	9.7
Desv. estándar		1.212	1.224
CON = Estación convencional			
EMA = Estación automática			
Tmed = Temperatura del aire			

Nota. Elaboración propia.

Figura 4.

Coeficiente de correlación



Nota. Elaboración propia.

Tabla 21.

Resultado de la prueba de test retest

Confiabilidad test – retest	
Estación meteorológica:	Puno
R de Pearson:	0.977
Interpretación:	Muy Alta

Nota. Elaboración propia.

Tabla 22.

Categoría de la prueba

Rango	**Interpretación**
<0.8 - 1]	Muy Alta
<0.6 - 0.8]	Alta
<0.4 - 0.6]	Moderada
<0.2 - 0.4]	Baja
[0.0 - 0.2]	Muy Baja

Nota. Elaboración propia.

Como resultado de la prueba de confiabilidad Test-Retest, se concluye que el instrumento y sensor de las estaciones convencionales y automáticas son confiable y validas dado que el coeficiente de correlación resulto de 0.977, y tomando en consideración de acuerdo a la tabla A5 se calificó de Muy Alta.

Anexo B Recolección de datos y selección de estaciones

Descripción:

Esta fase abarca la recopilación de datos históricos de temperatura del aire (2001-2020) de las estaciones meteorológicas de la Sierra de Puno y la selección de 22 estaciones (3 automáticas y 19 convencionales) basadas en criterios de altitud (>3,800 msnm), disponibilidad de datos (>75% de registros) y exclusión de provincias con influencia selvática (Carabaya y Sandia). Los anexos ilustran la distribución geográfica, características de las estaciones y el contexto climático regional.

Tabla 23.

Relación de tablas y figuras del anexo B

N.º Ítem	Descripción	Justificación	Objetivo
Tabla 24	Red de estaciones meteorológicas del SENAMHI	Lista las 22 estaciones seleccionadas, con detalles de ID, ubicación (latitud, longitud, altitud), tipo de estación y porcentaje de datos disponibles, asegurando representatividad.	Proporciona la base de datos para el modelo matemático.
Figura 5	Mapa de la red de estaciones meteorológicas del SENAMHI	Muestra la ubicación geográfica de las estaciones, diferenciando entre automáticas y convencionales, destacando la cobertura en la Sierra de Puno.	Visualiza la representatividad geográfica de las estaciones seleccionadas.
Figura 6	Mapa de distancia de estaciones	Representa las distancias geográficas entre estaciones (usando geopy. distance. geodesic), facilitando la selección de estaciones vecinas para análisis espacial.	Definir vecindad para la consistencia espacial.
Figura 7	Mapa de clasificación climática del SENAMHI - 2021	Ubicar las estaciones en su contexto climático.	Asegura que las estaciones reflejan el clima de la región sierra de estudio.

Nota. Elaboración propia.

Tabla 24.

Red de estaciones meteorológicas del SENAMHI

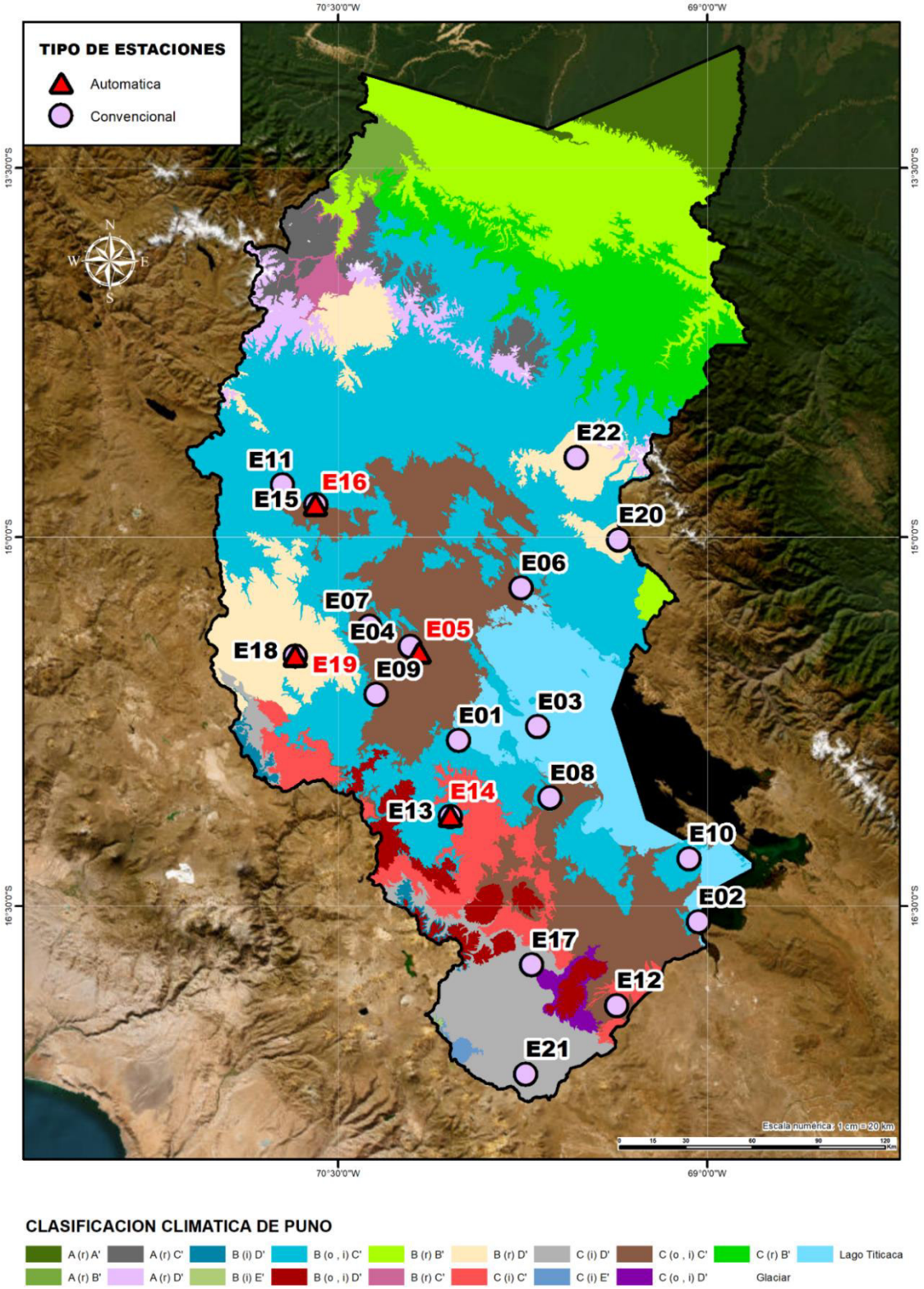
ID Estación	Estación	Longitud	Latitud	Altitud	Provincia	Distrito	Porc Tmed
E01	Puno_Con	-70.01208	-15.82625	3825	PUNO	PUNO	95.65
E02	Desaguadero_Con	-69.03682	-16.56327	3833	CHUCUITO	DESAGUADERO	95.6
E03	IslaTaquile_Con	-69.69107	-15.77019	3837	PUNO	AMANTANI	95.61
E04	Juliaca_Con	-70.2078	-15.4441	3838	SAN ROMAN	JULIACA	95.44
E05	Juliaca_EMA	-70.17122	-15.47106	3826	SAN ROMAN	JULIACA	36.28
E06	Huancane_Con	-69.7583	-15.20687	3842	HUANCANE	HUANCANE	95.59
E07	Lampa_Con	-70.37415	-15.36109	3866	LAMPA	LAMPA	98.36
E08	Ilave_Con	-69.64039	-16.05994	3874	EL COLLAO	ILAVE	97.26
E09	Cabanillas_Con	-70.34661	-15.63911	3885	SAN ROMAN	CABANILLAS	99.78
E10	Tahuaco_Con	-69.07472	-16.30783	3888	YUNGUYO	YUNGUYO	97.43
E11	Chuquibambilla_Con	-70.7285	-14.7879	3918	MELGAR	UMACHIRI	95.32
E12	Pizacoma_Con	-69.3695	-16.9043	3930	CHUCUITO	PISACOMA	98.77
E13	Laraqueri_Con	-70.04587	-16.13599	3939	PUNO	PICHACANI	99.56
E14	Laraqueri_EMA	-70.04589	-16.13592	3928	PUNO	PICHACANI	34.53
E15	Ayaviri_Con	-70.5916	-14.8688	3941	MELGAR	AYAVIRI	97.81
E16	Ayaviri_EMA	-70.59314	-14.87183	3928	MELGAR	AYAVIRI	64.83
E17	MazoCruz_Con	-69.71544	-16.73888	3980	EL COLLAO	SANTA ROSA	99.72
E18	PampaHuta_Con	-70.676	-15.4837	4316	LAMPA	PARATIA	99.99
E19	PampaHuta_EMA	-70.67578	-15.48528	4400	LAMPA	PARATIA	30.52
E20	Cojata_Con	-69.3612	-15.0118	4347	HUANCANE	COJATA	99.02
E21	Capazo_Con	-69.7391	-17.1827	4423	EL COLLAO	CAPASO	97.07
E22	Ananea_Con	-69.5339	-14.6763	4660	SAN ANTONIO DE PUTINA	ANANEA	91.68

Código Tipo	Tipo de Estación	Cantidad
Con	Estación Convencional	19
EMA	Estación Automática	3
	Total	22

Nota. Elaboración propia.

Figura 5.

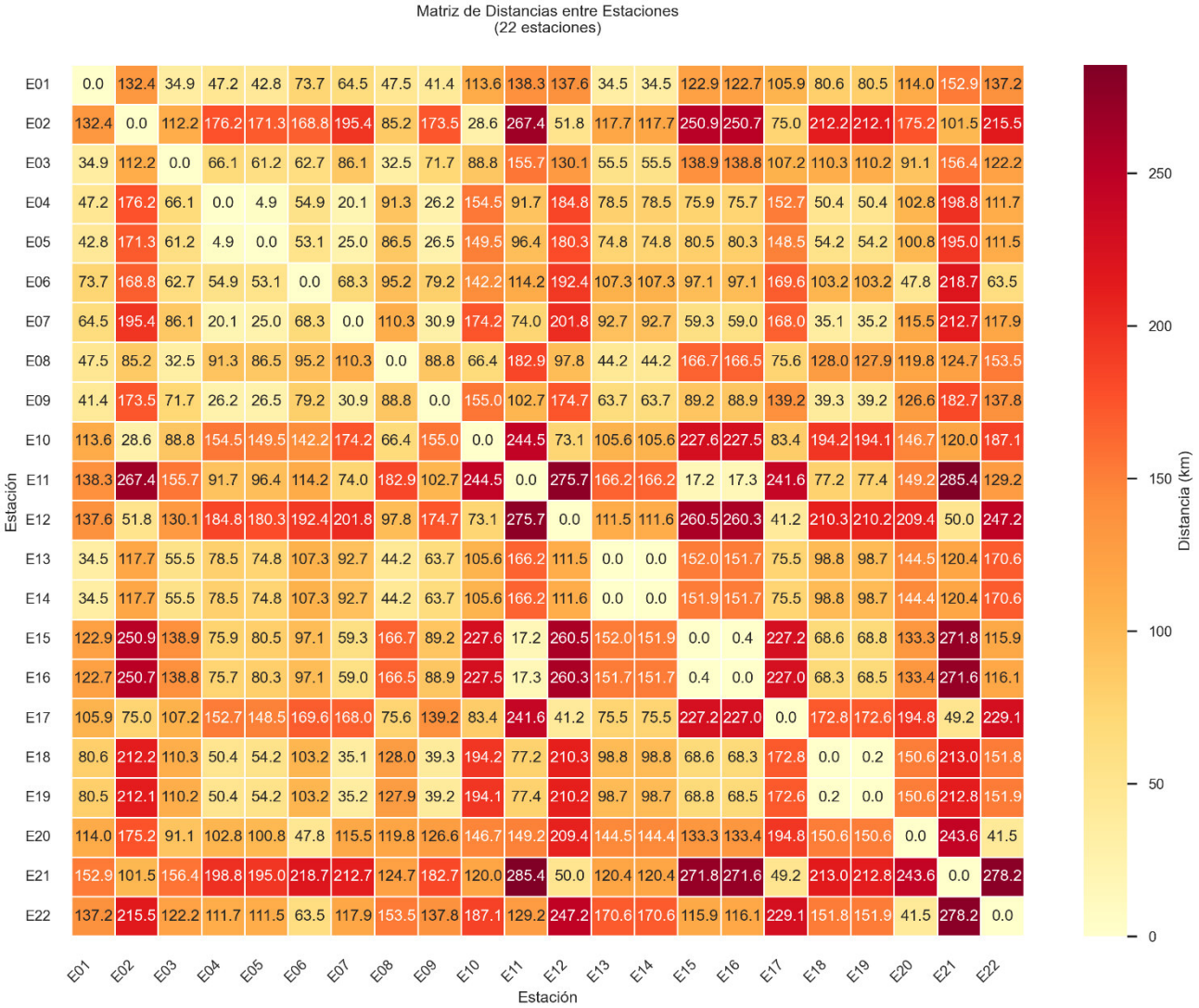
Mapa de la red de estaciones meteorológicas del SENAMHI



Nota. Elaboración propia

Figura 6.

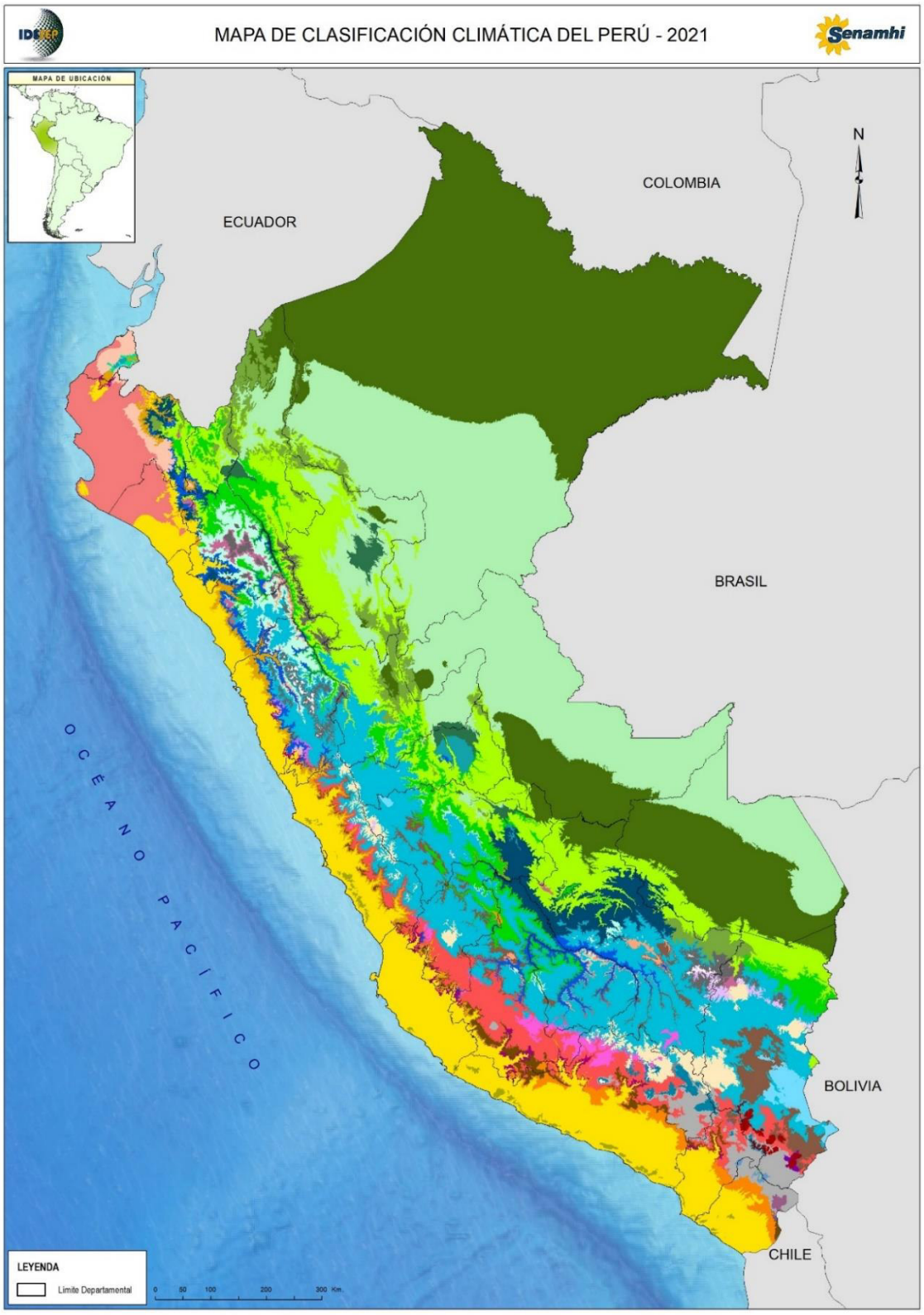
Mapa de distancia de estaciones



Nota. Elaboración propia

Figura 7.

Mapa de clasificación climática del SENAMHI – 2021



Nota. Servicio Nacional de Meteorología e Hidrología del Perú - SENAMHI (2021)

Anexo C Detección de anomalías parte 1

Descripción:

Esta fase identifica y clasifica errores en los datos de temperatura (valores faltantes, atípicos, inconsistencias) mediante reglas estadísticas y físicas (SD, LM, LN, VE, VR, FH, CE, LR, LE), siguiendo estándares de la Organización Meteorológica Mundial (OMM) y el SENAMHI. Los anexos presentan las reglas aplicadas, la distribución de anomalías y los umbrales físicos, proporcionando una base para la corrección posterior.

Tabla 25.

Relación de tablas y figuras del anexo C

N.º Ítem	Descripción	Justificación	Objetivo
Tabla 26	Reglas de control de calidad de datos de temperatura del aire	Detalla las reglas y fórmulas para detectar anomalías (rangos físicos, filtro de Hampel, consistencia espacial), asegurando un proceso sistemático.	Identifica datos erróneos para su corrección.
Figura 8	Flujo del modelo del control de calidad	Resumen de flujo de modelo del Control de Calidad	Explicar las secuencias de pasos realizado para realizar el modelo del control de calidad
Tabla 27	Umbrales de temperatura del aire media por estación	Muestra los rangos físicos de temperatura (mínimo/máximo) por estación, calculados con boxplot (IQR), para detectar valores fuera de rango.	Establece límites para identificar anomalías.
Figura 9	Gráfico de umbrales de las estaciones meteorológicas de Puno	Muestran los umbrales físicos basados en IQR, para identificar valores atípicos de temperatura de todas las estaciones ordenadas en base a la altitud.	Proporciona una representación visual de los rangos aceptables y Detectar valores atípicos.
Figura 10	Temperatura con anomalías detectadas grupo 1	Visualiza una serie de tiempo de temperatura con las anomalías por cada estación	Cuantificar visualmente las anomalías por estación para detectar patrones de errores

Figura 11	Temperatura con anomalías detectadas grupo 2	Visualiza una serie de tiempo de temperatura con las anomalías por cada estación	Cuantificar visualmente las anomalías por estación para detectar patrones de errores
Figura 12	Temperatura con anomalías detectadas grupo 3	Visualiza una serie de tiempo de temperatura con las anomalías por cada estación	Cuantificar visualmente las anomalías por estación para detectar patrones de errores
Figura 13	Temperatura con anomalías detectadas grupo 4	Visualiza una serie de tiempo de temperatura con las anomalías por cada estación	Cuantificar visualmente las anomalías por estación para detectar patrones de errores
Figura 14	Temperatura con anomalías detectadas grupo 5	Visualiza una serie de tiempo de temperatura con las anomalías por cada estación	Cuantificar visualmente las anomalías por estación para detectar patrones de errores
Figura 15	Temperatura con anomalías detectadas grupo 6	Visualiza una serie de tiempo de temperatura con las anomalías por cada estación	Cuantificar visualmente las anomalías por estación para detectar patrones de errores
Tabla 28	Indicadores de distribución de anomalías parte 1	Muestra indicadores en base a la mediana y el MAD por el análisis de anomalías por estación	Cuantificar las anomalías por estación para detectar patrones de errores
Tabla 29	Indicadores de distribución de anomalías parte 2	Muestra indicadores en base a la mediana y el MAD por el análisis de anomalías por estación	Cuantificar las anomalías por estación para detectar patrones de errores
Figura 16	Distribución de anomalías por estación	Visualiza la distribución de las anomalías en gráfico de barras	Identifica patrones de errores a nivel de estación.
Figura 17	Mapa de distribución de anomalías por estación grupo 1	Visualiza la distribución espacial de anomalías (SD, LM, LN, VE, VR, FH, CE, LR, LE) en las 22 estaciones.	Identifica patrones geográficos de errores.
Figura 18	Mapa de distribución de anomalías por estación grupo 2	Visualiza la distribución espacial de anomalías (SD, LM, LN, VE, VR, FH, CE, LR, LE) en las 22 estaciones.	Identifica patrones geográficos de errores.
Figura 19	Mapa de distribución de anomalías por estación grupo 3	Visualiza la distribución espacial de anomalías (SD, LM, LN, VE, VR, FH, CE, LR, LE) en las 22 estaciones.	Identifica patrones geográficos de errores.

Figura 20	Mapa de distribución de anomalías por estación grupo 4	Visualiza la distribución espacial de anomalías (SD, LM, LN, VE, VR, FH, CE, LR, LE) en las 22 estaciones.	Identifica patrones geográficos de errores.
Figura 21	Mapa de distribución espacial de anomalías categorizadas	Visualiza la distribución espacial de anomalías por categoría	Identifica patrones geográficos de errores por categorías.

Nota. Elaboración propia

Tabla 26.

Reglas de control de calidad de datos de temperatura del aire

Banderín	Calificación	Regla a cumplir	Descripción	Grupo banderín	Tipo de prueba
LM	Limite mundial	- 40.0 °C ≤ Ti ≤ 60.0 °C Si Cumple es "OK" de lo Contrario "LD"	Prueba de Límites duros a nivel de la Zona de Sudamérica o global. Estos controles aseguran que no existan valores considerados físicamente imposibles.	M	Umbrales físicos
LN	Limite nacional	- 32.0 °C ≤ Th ≤ 45.0 °C Si Cumple es "OK" de lo Contrario "LN"	Prueba de Limite Blando a Nivel Nacional	D	Umbrales físicos
LS	Limite regional	Linf_sect-n_T ≤ Th ≤ Lsup_sect-n_T Si Cumple es "OK" de lo Contrario "LS"	Prueba de Limite Blando a Nivel Sectorial (- 25 °C ≤ Th ≤ 10 °C)	D	Umbrales físicos
LE	Limite estación	Linf_estacion ≤ Ti ≤ Lsup_estacion Si Cumple es "OK" de lo Contrario "LE"	Linf_estacion ≤ Ti ≤ Lsup_estacion (Tukey IQR). Si Cumple es "OK" de lo Contrario "LE" Linf_estacion = Q1 – k × IQR Lsup_estacion = Q3 + k × IQR	D	Umbrales físicos
VE	Variación excesiva	|Ti -Ti-1| ≤ 9 °C Si Cumple es "OK" de lo Contrario "VE"	La absoluta de la variación de la Temperatura actual con respecto al dato anterior debe ser menor o igual a un Umbral definido, caso contrario se considerado un dato Dudoso.	D	Consistencia temporal
VR	Valor repetido	Ti = Ti-1 = Ti-2 = Ti-3 Si Cumple es "OK" de lo Contrario "VR"	Solo se consideran validos los últimos 3 valores incluyendo el dato actual, caso contrario se considerarán dato Dudoso.	D	Consistencia Temporal
FH	Filtro hampel	|Xi − mediana(X)| ≤ 3.5 * MAD * 1.4826 Si Cumple es "OK" de lo Contrario "AH"	Si el la variación Absoluta del valor actual con respecto a la Mediana de una ventana de 30 días es mayor a 3.5 * MAD (Desviación Absoluta Mediana) entonces se considera datos dudoso (3.5 * MAD * 1.4826 ≤ 3.5 * STD) 1.- Calcula la mediana de una ventana de datos de 30 días alrededor de cada punto. 2.- Calcula la MAD para la ventana tomada. 3.- Define un umbral = 3.5 * MAD * 1.4826 4.- Identifica como outliers los puntos que están fuera de este rango: |Xi − mediana(X)| ≤ 3.5 * MAD, Donde: MAD = Mediana(|Xi − mediana(X)|)	D	Filtro de hampel

CE	Consistencia espacial	Consistencia Espacial |Testacion−median (Tvecinas)| ≤ 5 Si Cumple es "OK" de lo Contrario "CE"	Validar datos comparando estaciones cercanas (radio ≤50 km): 1. Método: Comparar cada dato con la mediana de estaciones vecinas. 2. Umbral: ±5°C (ajustable según variabilidad regional). Marcar como inconsistente si: |Testacion−median (Tvecinas)| ≤ 5 Si Cumple es "OK" de lo Contrario "CE"	D	Consistencia espacial
RF	Completado		Si la predicción esta entre [percentil 1, percentil 99] se usa Random Forest (RF) caso contrario Gradiente Boosting (GB)	C	Completado
	Tabla categoría:	**C = Conforme**	**D = Dudoso**	**M = Malo**	**NA = Sin Dat**

Nota. Elaboración propia

Figura 8.

Flujo del modelo de control de calidad



Nota. Elaboración propia

Tabla 27.

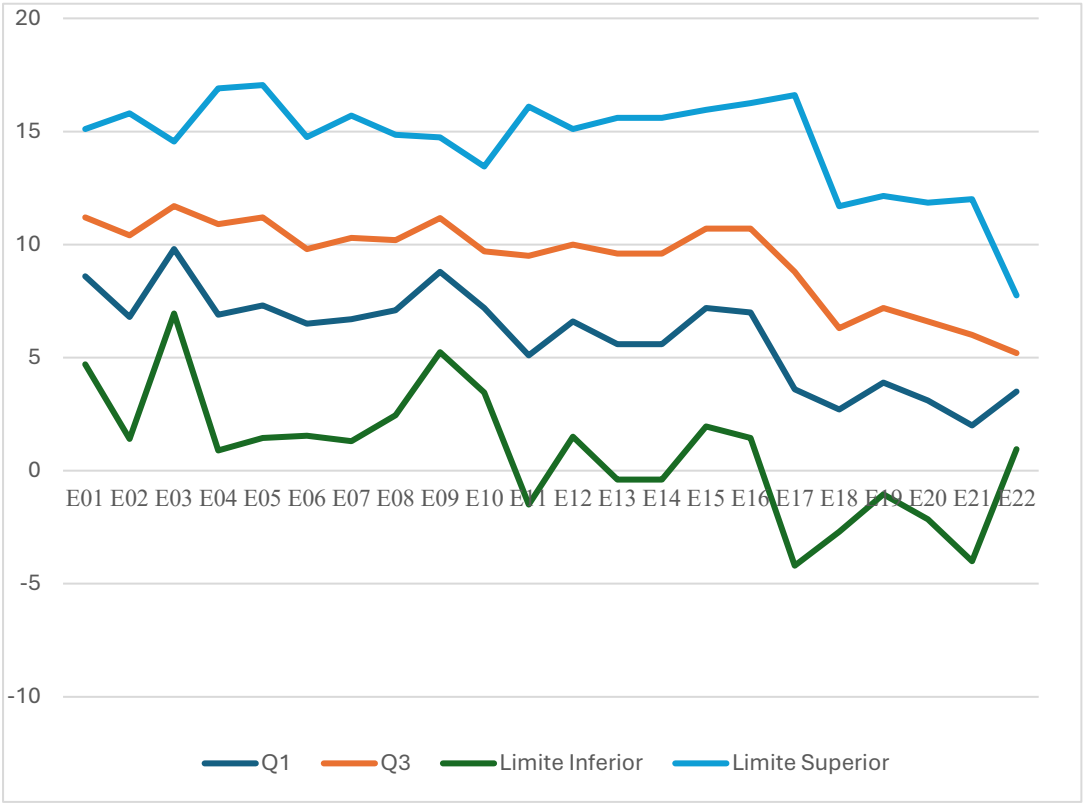
Umbrales de temperatura del aire media por estación

ID Estación	Q1	Q3	IQR	Límite Inferior	Limite Superior	Factor IQR Inferior	Factor IQR Superior
E01	8.6	11.2	2.6	4.7	15.1	1.5	1.5
E02	6.8	10.4	3.6	1.4	15.8	1.5	1.5
E03	9.8	11.7	1.9	6.95	14.55	1.5	1.5
E04	6.9	10.9	4	0.9	16.9	1.5	1.5
E05	7.3	11.2	3.9	1.45	17.05	1.5	1.5
E06	6.5	9.8	3.3	1.55	14.75	1.5	1.5
E07	6.7	10.3	3.6	1.3	15.7	1.5	1.5
E08	7.1	10.2	3.1	2.45	14.85	1.5	1.5
E09	8.8	11.17	2.37	5.24	14.74	1.5	1.5
E10	7.2	9.7	2.5	3.45	13.45	1.5	1.5
E11	5.1	9.5	4.4	-1.5	16.1	1.5	1.5
E12	6.6	10	3.4	1.5	15.1	1.5	1.5
E13	5.6	9.6	4	-0.4	15.6	1.5	1.5
E14	5.6	9.6	4	-0.4	15.6	1.5	1.5
E15	7.2	10.7	3.5	1.95	15.95	1.5	1.5
E16	7	10.7	3.7	1.45	16.25	1.5	1.5
E17	3.6	8.8	5.2	-4.2	16.6	1.5	1.5
E18	2.7	6.3	3.6	-2.7	11.7	1.5	1.5
E19	3.9	7.2	3.3	-1.05	12.15	1.5	1.5
E20	3.1	6.6	3.5	-2.15	11.85	1.5	1.5
E21	2	6	4	-4	12	1.5	1.5
E22	3.5	5.2	1.7	0.95	7.75	1.5	1.5

Nota. Elaboración propia

Figura 9.

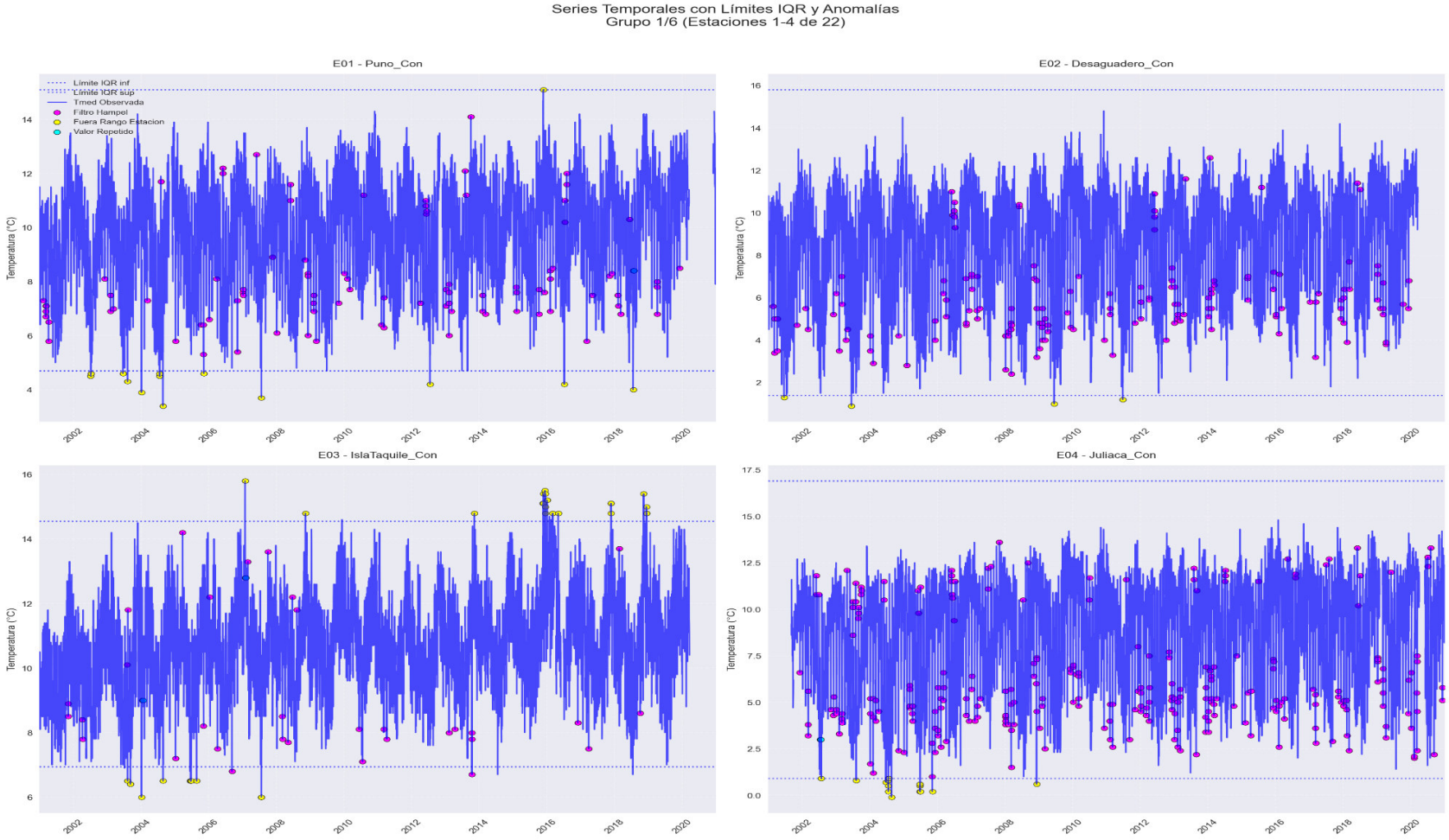
Umbrales de las estaciones meteorológicas de Puno



Nota. Elaboración propia

Figura 10.

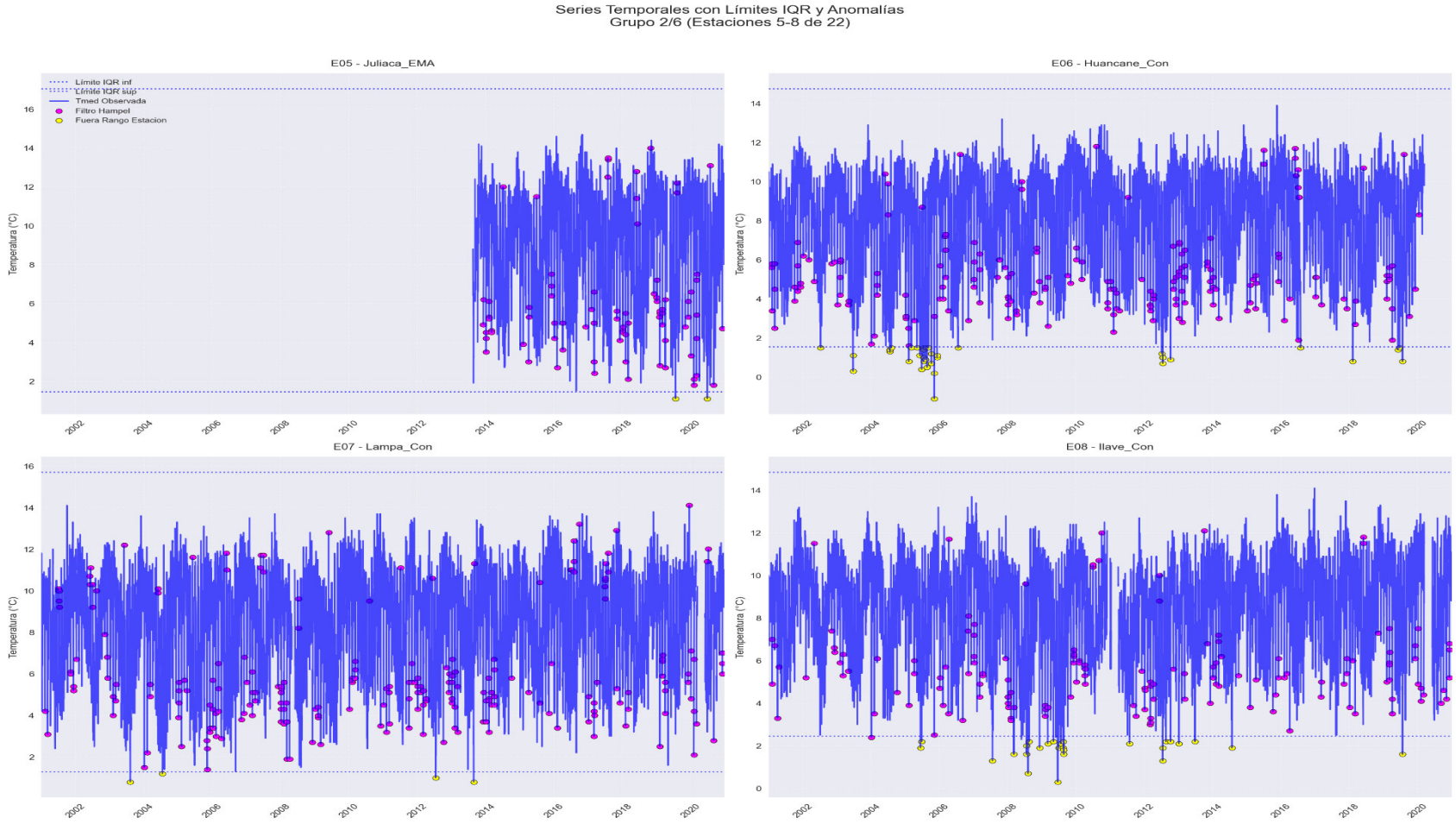
Temperatura con anomalías detectadas grupo 1



Nota. Elaboración propia

Figura 11.

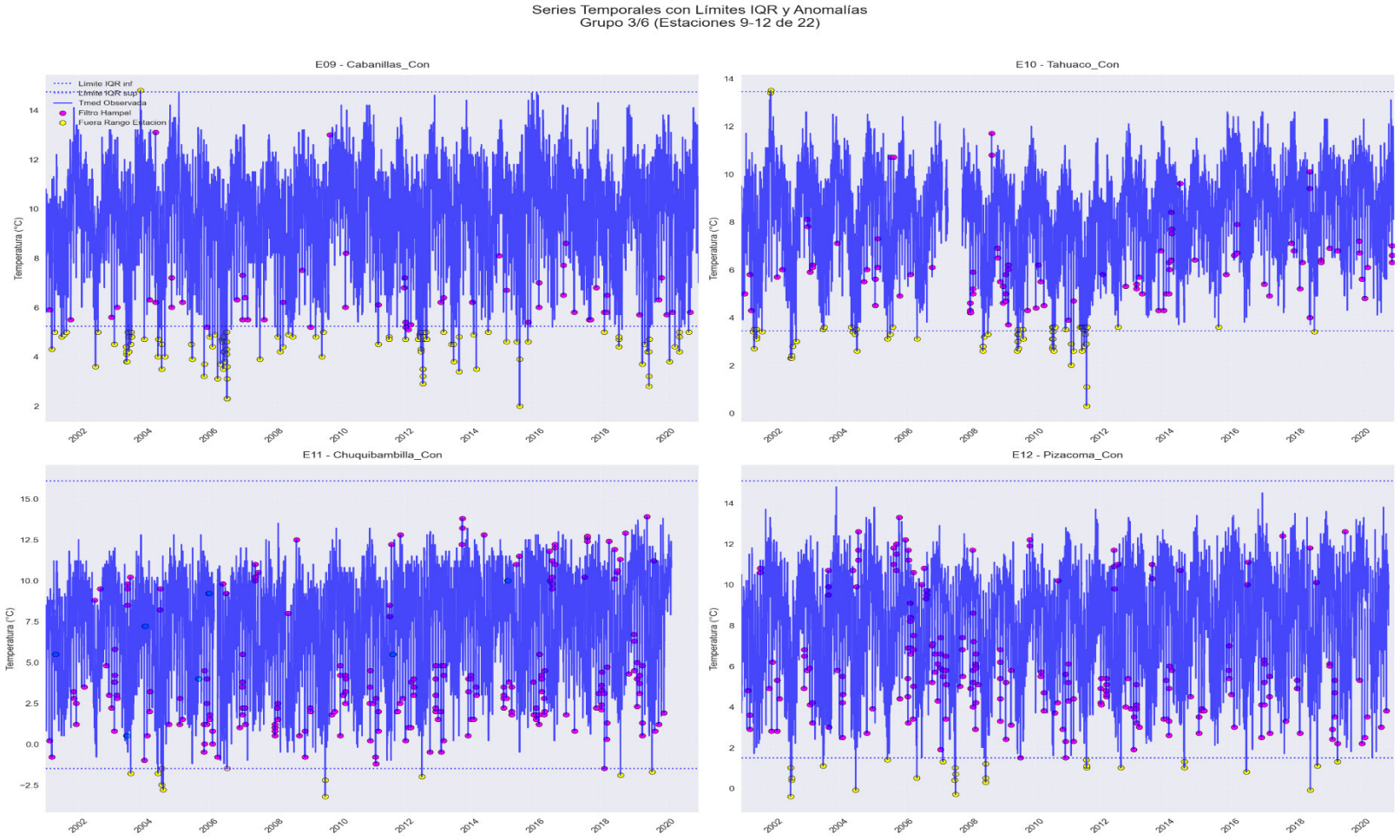
Temperatura con anomalías detectadas grupo 2



Nota. Elaboración propia

Figura 12.

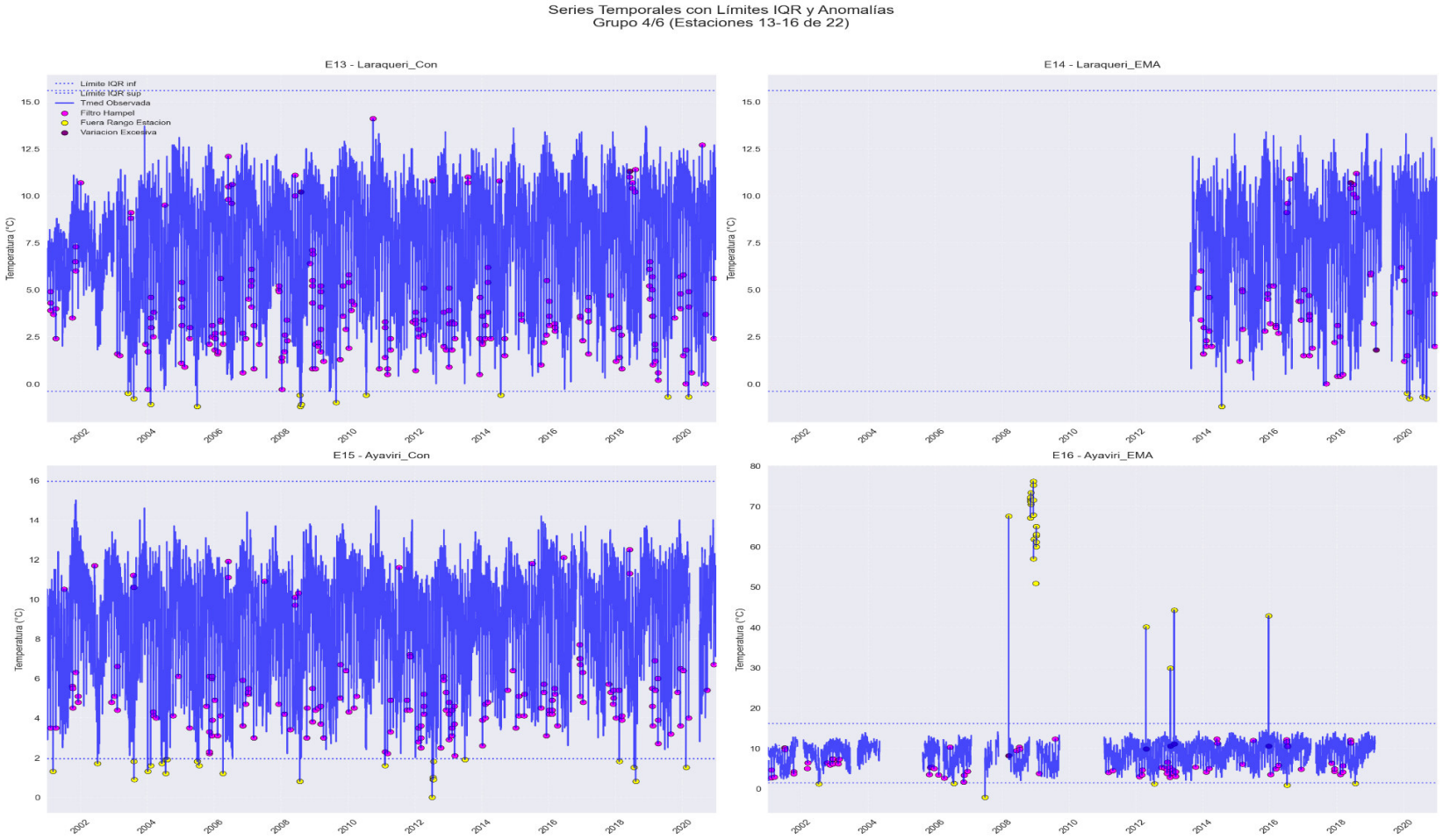
Temperatura con anomalías detectadas grupo 3



Nota. Elaboración propia

Figura 13.

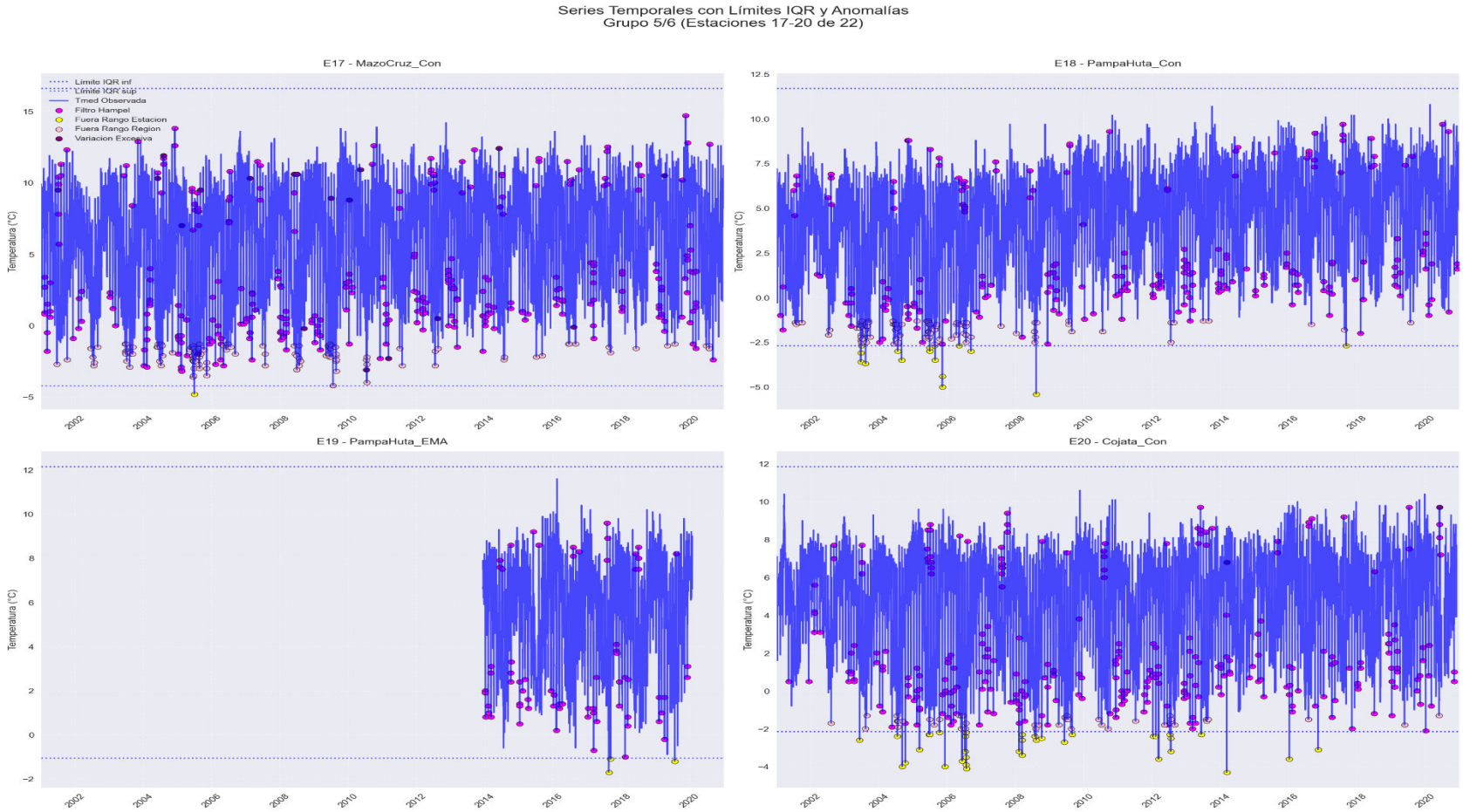
Temperatura con anomalías detectadas grupo 4



Nota. Elaboración propia

Figura 14.

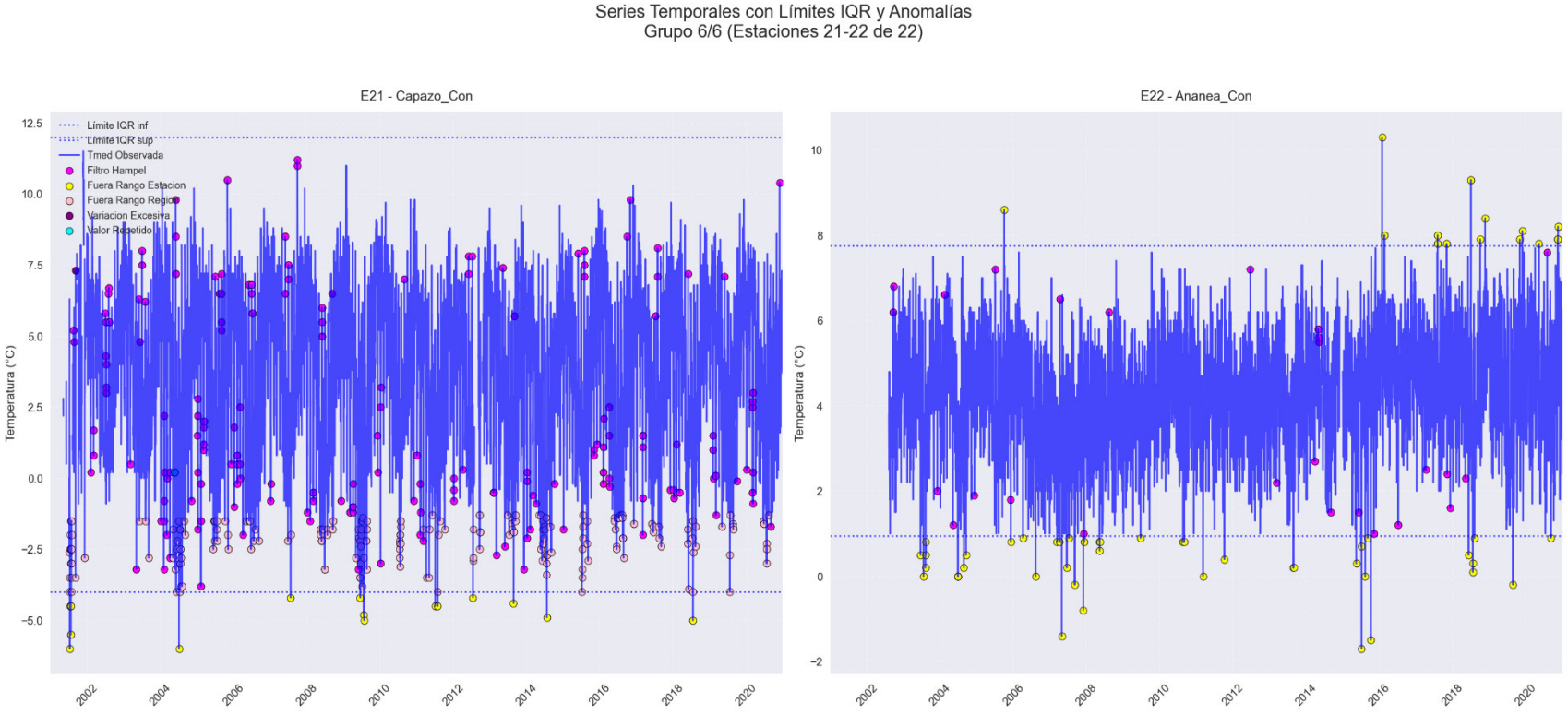
Temperatura con anomalías detectadas grupo 5



Nota. Elaboración propia

Figura 15.

Temperatura con anomalías detectadas grupo 6



Nota. Elaboración propia

Tabla 28.

Tabla de indicadores de distribución de anomalías parte 1

ID Esta	Nombre estación	Mediana OK	MAD OK	N Anomalías	Desv MAD Prom	Desv MAD Min	Desv MAD Max	Anomalía más frecuente	% Anomalías totales	Interpretación
E01	Puno Con	10	1.3	482	-1.75	-5.08	3.92	FH	6.6	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E02	Desaguadero Con	9.1	1.7	522	-1.94	-4.82	2.06	FH	7.1	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E03	IslaTaquile Con	10.7	1	433	-0.04	-4.7	5.1	FH	5.9	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E04	Juliaca Con	9.4	1.8	669	-1.96	-5.28	2.33	FH	9.2	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E05	Juliaca EMA	9.8	1.8	4785	-2.2	-4.83	2.33	FH	65.5	Alta proporción de valores anómalos con desviaciones sistemáticas. Revisar sensor o datos históricos.
E06	Huancane Con	8.5	1.6	587	-2.47	-6	2.06	FH	8	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E07	Lampa Con	8.9	1.7	378	-1.7	-4.76	3.06	FH	5.2	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E08	Ilave Con	8.9	1.5	450	-2.6	-5.73	2.13	FH	6.2	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E09	Cabanillas Con	10	1.2	183	-4.1	-6.67	4	LE	2.5	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E10	Tahuaco Con	8.6	1.2	390	-3.01	-6.92	4.08	FH	5.3	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E11	Chuquibambilla Con	7.8	2	633	-1.77	-5.5	3.05	FH	8.7	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E12	Pizacoma Con	8.7	1.6	358	-2.03	-5.69	2.87	FH	4.9	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E13	Laraqueri Con	7.9	1.9	270	-2.19	-4.79	3.26	FH	3.7	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.

Nota. Elaboración propia

Tabla 29.

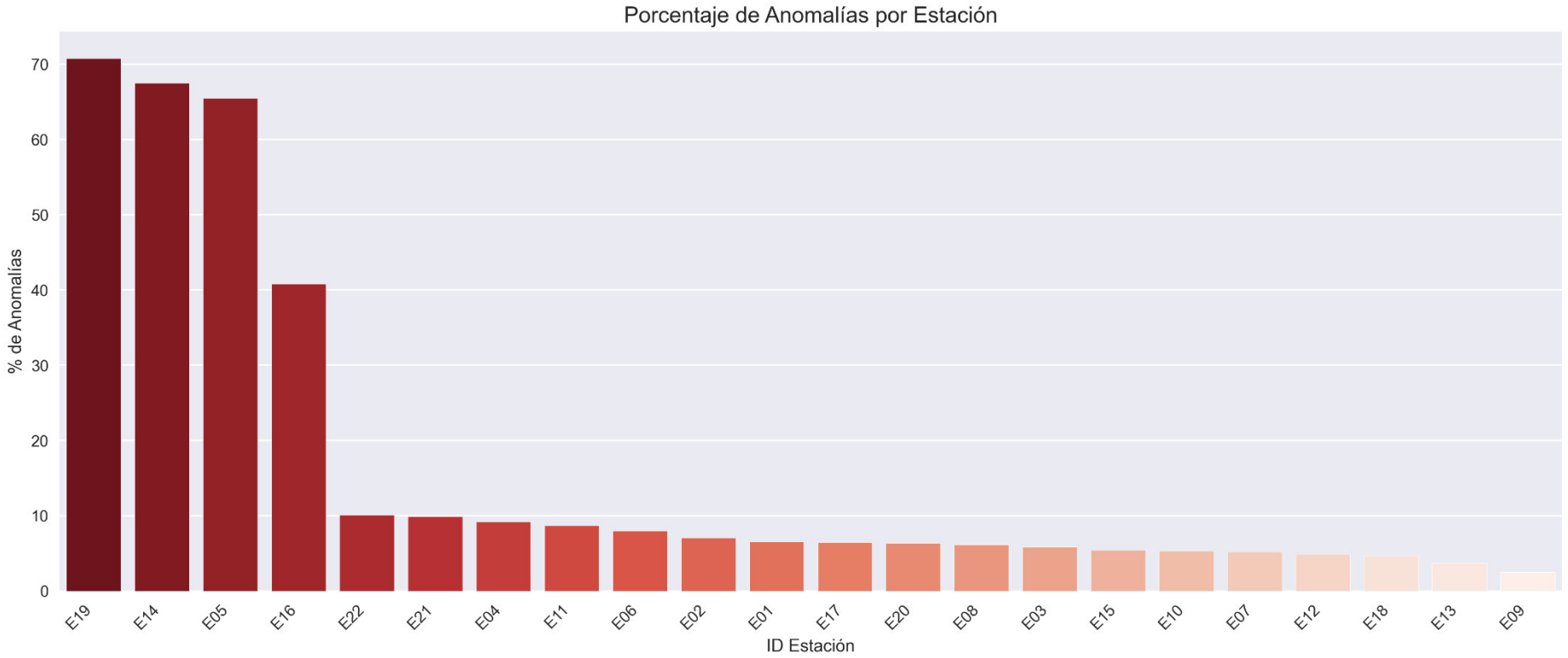
Indicadores de distribución de anomalías parte 2

ID Esta	Nombre estación	Mediana OK	MAD OK	N Anomalías	Desv MAD Prom	Desv MAD Min	Desv MAD Max	Anomalía más frecuente	% Anomalías totales	Interpretación
E14	Laraqueri EMA	8	1.8	4932	-2.26	-5.11	1.78	FH	67.5	Alta proporción de valores anómalos con desviaciones sistemáticas. Revisar sensor o datos históricos.
E15	Ayaviri Con	9.4	1.6	395	-2.93	-5.87	1.94	FH	5.4	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E16	Ayaviri EMA	9.2	1.7	2983	5.43	-6.65	39.41	FH	40.8	Alta proporción de valores anómalos con desviaciones sistemáticas. Revisar sensor o datos históricos.
E17	MazoCruz Con	6.5	2.6	475	-1.72	-4.35	3.15	FH	6.5	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
E18	PampaHuta Con	4.8	1.7	342	-2.31	-6	2.88	FH	4.7	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E19	PampaHuta EMA	5.9	1.6	5174	-1.66	-4.75	2.31	FH	70.8	Alta proporción de valores anómalos. Revisar sensor o datos históricos.
E20	Cojata Con	5.2	1.6	464	-2.55	-5.94	2.81	FH	6.4	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E21	Capazo Con	4.2	2	723	-2.22	-5.1	3.5	LR	9.9	Alta proporción de valores anómalos con desviaciones sistemáticas. Anomalías moderadas o puntuales.
E22	Ananea Con	4.3	0.8	738	-1.89	-7.5	7.5	LE	10.1	Alta proporción de valores anómalos. Anomalías moderadas o puntuales.
		Revisar Sensor								

Nota. Elaboración propia

Figura 16.

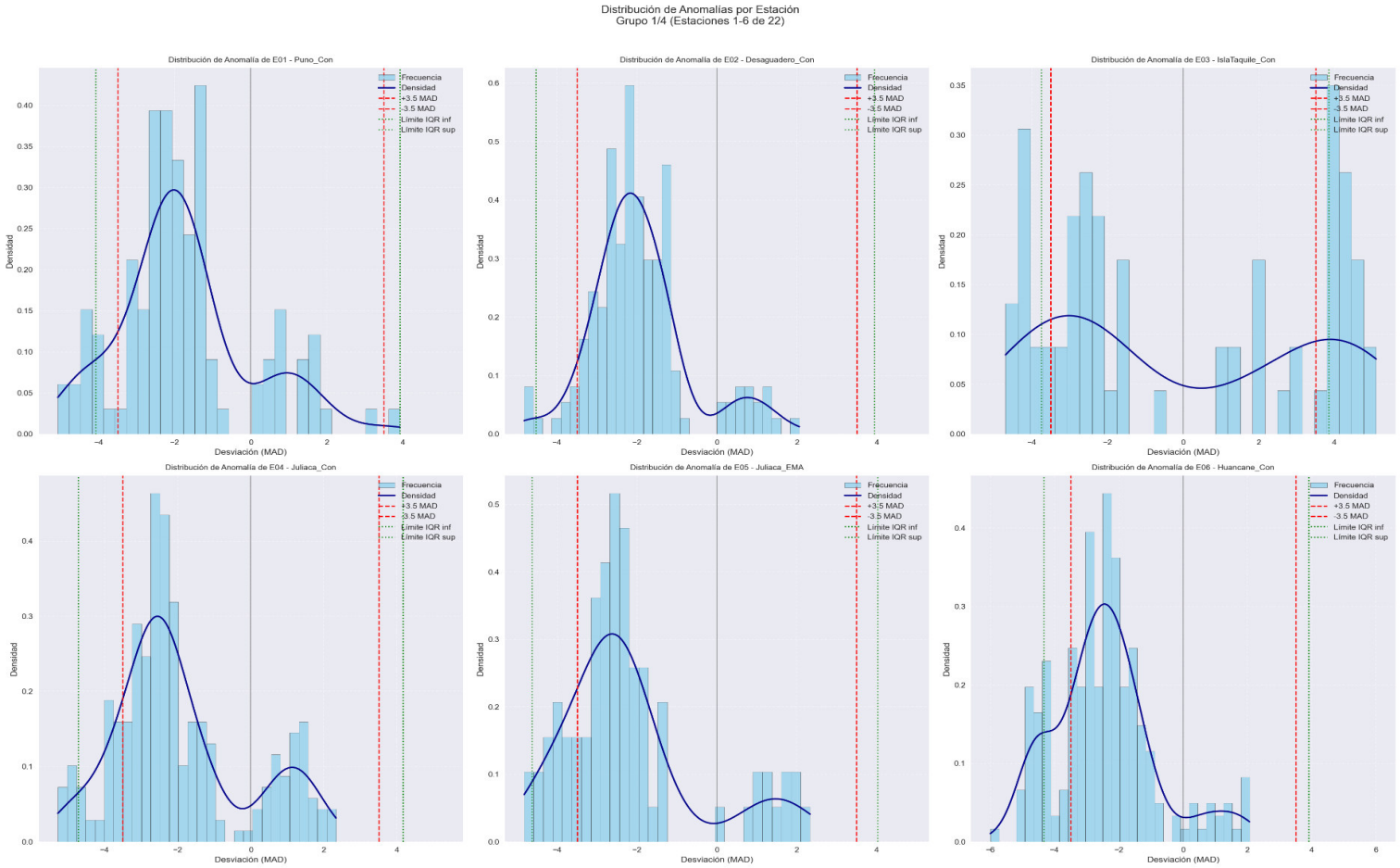
Distribución de anomalías por estación



Nota. Elaboración propia

Figura 17.

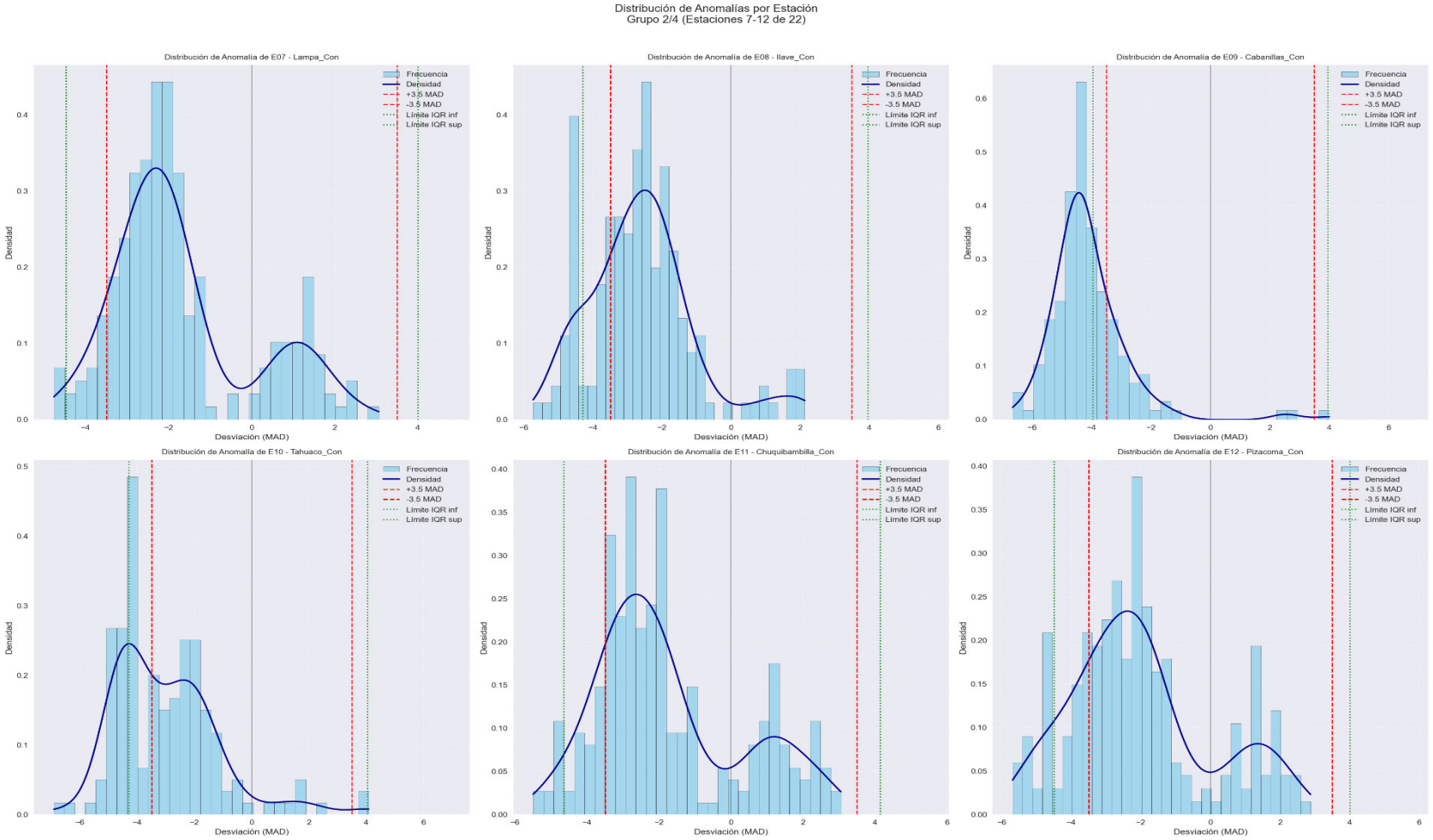
Mapa de distribución de anomalías por estación grupo 1



Nota. Elaboración propia

Figura 18.

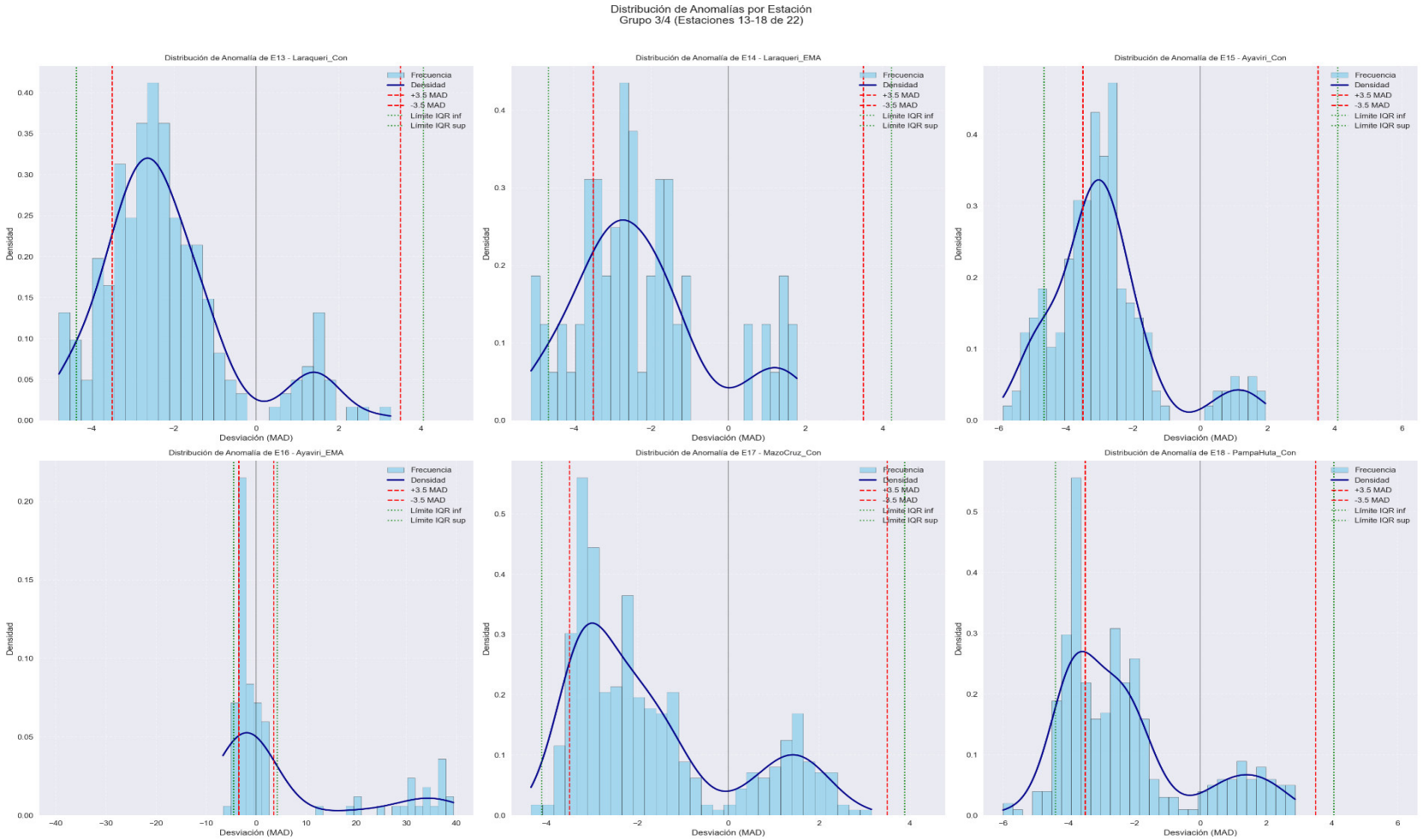
Mapa de distribución de anomalías por estación grupo 2



Nota. Elaboración propia

Figura 19.

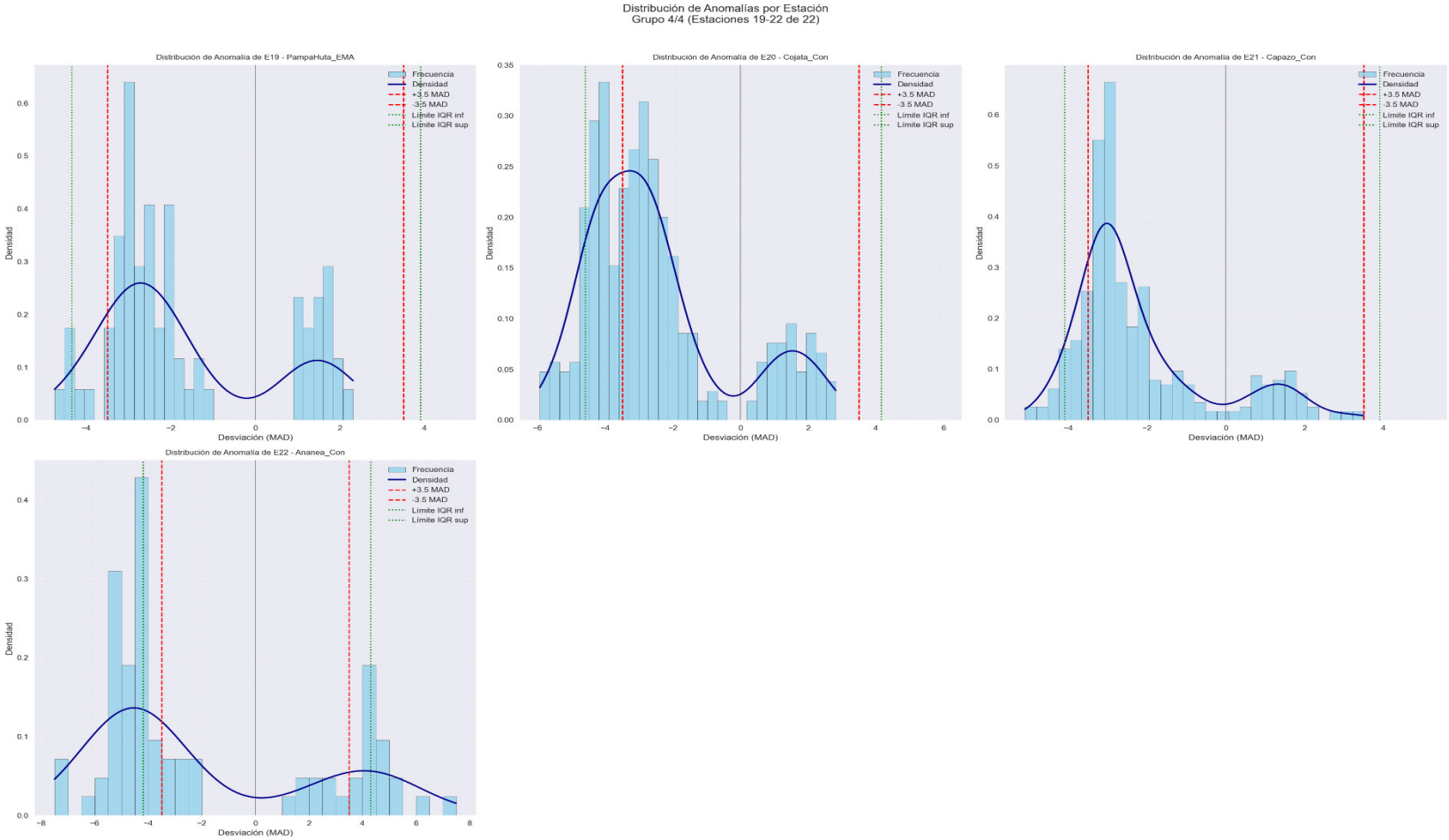
Mapa de distribución de anomalías por estación grupo 3



Nota. Elaboración propia

Figura 20.

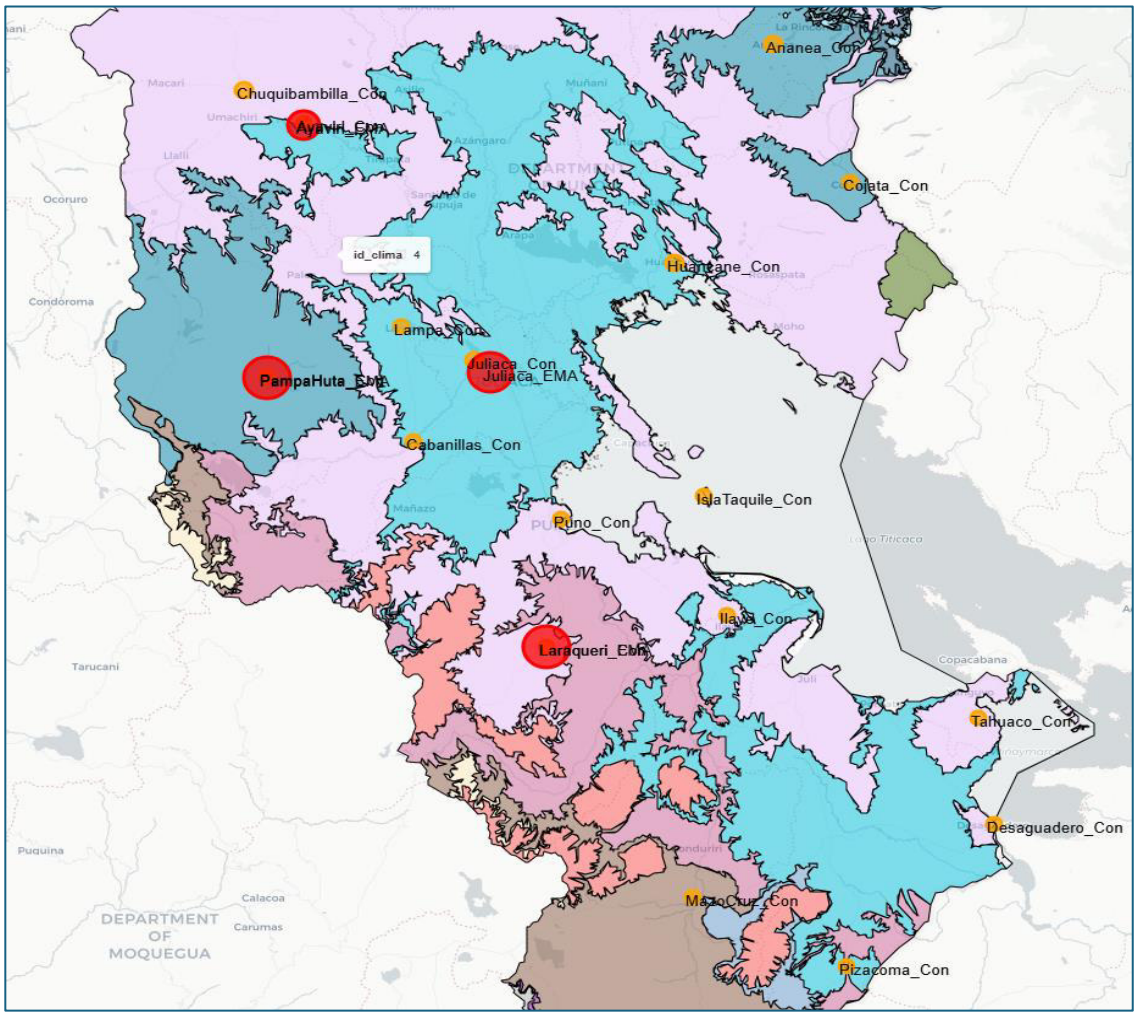
Mapa de distribución de anomalías por estación grupo 4



Nota. Elaboración propia

Figura 21.

Mapa de distribución espacial de anomalías categorizadas



Nota. Elaboración propia

Anexo D Corrección y relleno de datos

Descripción:

Esta fase aplica el modelo híbrido de Inteligencia Artificial: Random Forest y Gradient Boosting, para corregir datos erróneos y rellenar valores faltantes en las series de temperatura (Tmed). Los anexos presentan los indicadores estadísticos de la corrección, la distribución de residuales y visualizaciones que comparan los datos originales con los corregidos, destacando la reducción de incertidumbre en las series temporales.

Tabla 30.

Relación de tablas y figuras del anexo D

N.º Ítem	Descripción	Justificación	Objetivo
Tabla 31	Indicadores estadísticos de corrección y relleno	Muestra la desviación estándar original vs. corregida, p-valor (t-Student), y coeficiente de correlación (r > .9 en 95% de estaciones).	Cuantifica la reducción de incertidumbre (H1).
Figura 22	Gráfico de estadística de reducción de incertidumbre	Visualiza el porcentaje de la Reducción de la Incertidumbre	Valida la hipótesis H1 de reducción de incertidumbre.
Figura 23	Gráfico de diagrama de dispersión de temperatura diaria	Visualiza la Dispersión de los datos de Temperatura	Evalúa el Análisis de la dispersión de la Temperatura
Figura 24	Gráfico de comparación de datos del modelo	Visualiza una gráfica de correlaciona entre datos y generación de datos del modelo	Analizar la precisión de generar datos de temperatura
Figura 25	Mapas de correlación espacial original	Mapa de calor con la matriz de correlación de datos originales entre estaciones vecinas.	Proporciona la referencia para evaluar la mejora en correlación.
Figura 26	Mapas de correlación espacial corregido	Mapa de calor con la matriz de correlación de datos corregidos, mostrando un incremento en 62% de estaciones.	Valida la hipótesis H2 de mejora en consistencia espacial.
Tabla 32	Resumen de matriz de correlación de datos original vs rellenado	Resume el incremento de correlación espacial (62% en total, 72% sin estaciones automáticas).	Valida la hipótesis H2 de mejora en consistencia espacial.
Figura 27	Gráfico de porcentaje de aumento de coeficiente (r)	Visualiza el incremento de la correlación espacial de los datos de las estaciones	Mejora la consistencia espacial

Figura 28	Gráfico de predicciones versus valores reales	Gráfico de dispersión con línea de regresión, comparando Tmed original y Tmed_Rellenado, con métricas MAE, RMSE, R^2.	Demuestra la precisión y capacidad predictiva del modelo en la corrección (H1).
Figura 29	Gráfico de distribución de residuales	Visualiza la frecuencia de residuales entre el valor real y generado por el modelo	Se verifica la Distribución de Residuales

Nota. Elaboración propia

Tabla 31.

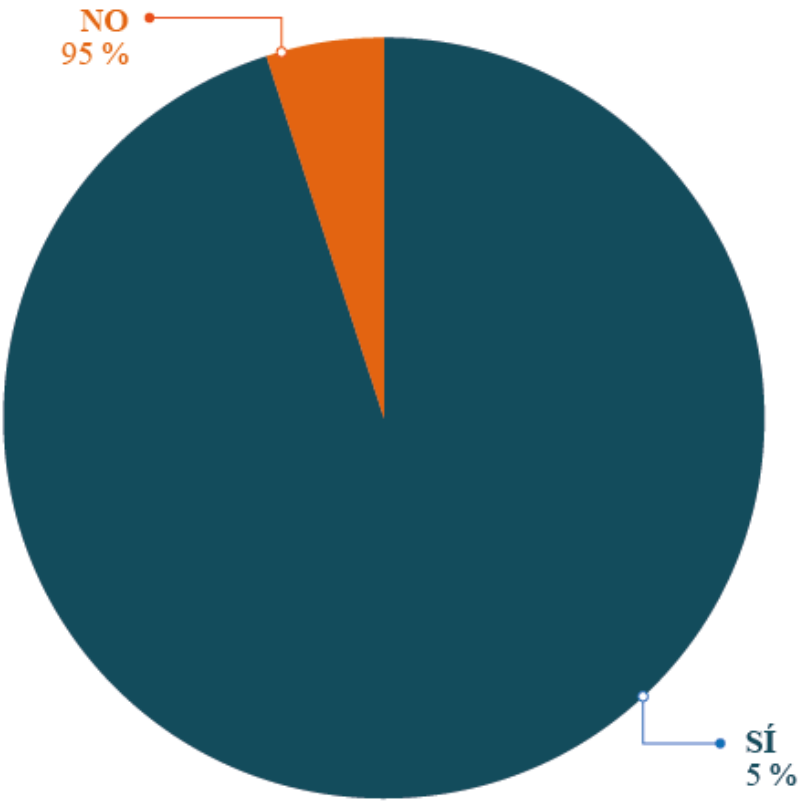
Indicadores estadísticos de corrección y relleno

Id Esta	Desviación Estándar original	Desviación Estándar corregido	t-Student (p-valor)	Reduce Incert.	Correlación Original Rellenado	RMSE Original Rellenado	Datos originales (%)	Datos generados (%)	Vecina cercana
E01	1.8176	1.7947	0	Si	0.9956	0.1706	94.9076	5.0924	E14
E02	2.4322	2.4012	0	Si	0.9959	0.2226	95.0582	4.9418	E10
E03	1.4587	1.4252	0.5123	No	0.9885	0.2211	95.0308	4.9692	E08
E04	2.7075	2.5902	0	Si	0.9705	0.6581	94.5654	5.4346	E05
E05	2.6887	2.6421	0	Si	0.9959	0.2456	35.6057	64.3943	E04
E06	2.3471	2.2784	0	Si	0.9913	0.3151	95.0582	4.9418	E20
E07	2.4186	2.3671	0	Si	0.9912	0.3213	97.9055	2.0945	E04
E08	2.181	2.1034	0	Si	0.9857	0.3727	96.1944	3.8056	E03
E09	1.909	1.7967	0	Si	0.9808	0.3837	99.7536	0.2464	E04
E10	1.82	1.7455	0	Si	0.9847	0.3226	96.8925	3.1075	E02
E11	2.9596	2.8866	0	Si	0.9935	0.3426	94.8939	5.1061	E15
E12	2.4271	2.3429	0	Si	0.9866	0.4011	98.3162	1.6838	E17
E13	2.7539	2.6919	0	Si	0.9899	0.3915	99.3977	0.6023	E14
E14	2.7678	2.7221	0	Si	0.9958	0.2563	33.4428	66.5572	E13
E15	2.4876	2.418	0	Si	0.9931	0.2978	97.1526	2.8474	E16
E16	4.7435	2.4515	0.0001	Si	0.5207	4.0564	60.6571	39.3429	E15
E17	3.4981	3.2582	0	Si	0.9741	0.8107	99.6578	0.3422	E12
E18	2.5367	2.3868	0	Si	0.9785	0.5352	99.9726	0.0274	E19
E19	2.2268	2.1863	0.0004	Si	0.9968	0.1804	30.1711	69.8289	E18
E20	2.5247	2.3393	0	Si	0.968	0.6493	98.5763	1.4237	E22
E21	2.7127	2.4989	0	Si	0.9697	0.6803	95.5647	4.4353	E17
E22	1.2826	1.219	0	Si	0.9563	0.3756	91.0472	8.9528	E20
Prd	2.49	2.3	0.02		0.96	0.56	86.36	13.64	

Nota. Elaboración propia

Figura 22.

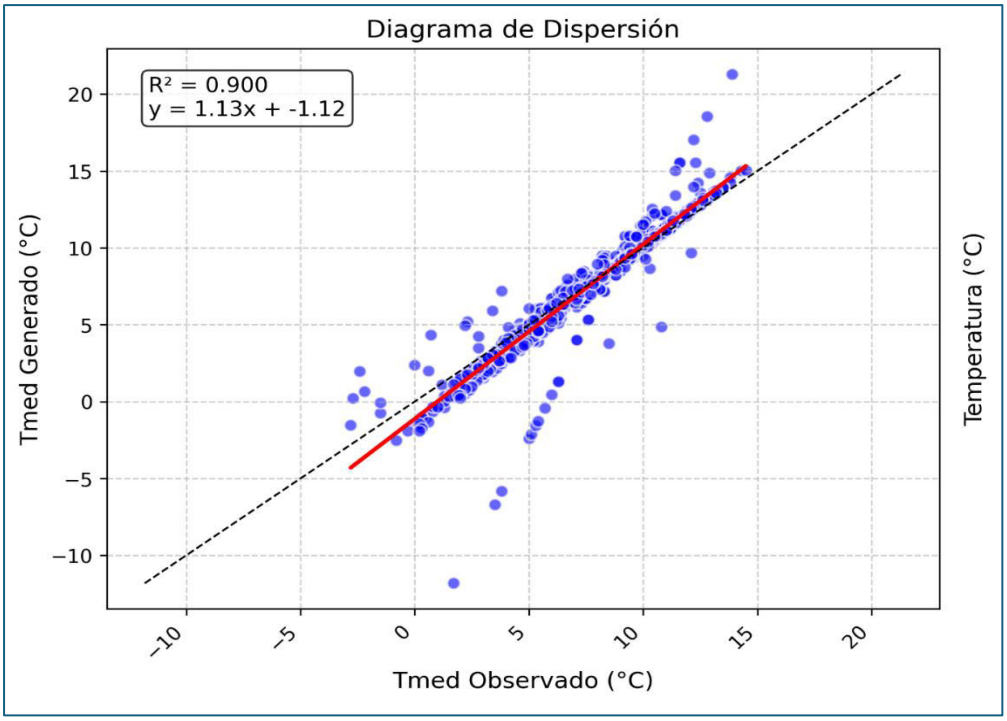
Estadística de reducción de incertidumbre



Nota. Elaboración propia

Figura 23.

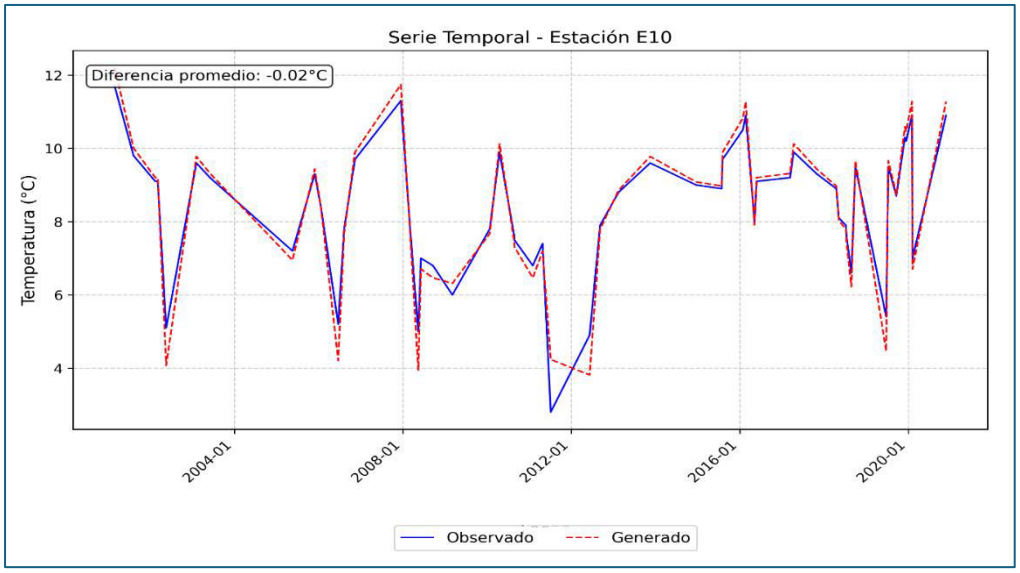
Gráfico de diagrama de dispersión de temperatura diaria



Nota. Elaboración propia

Figura 24.

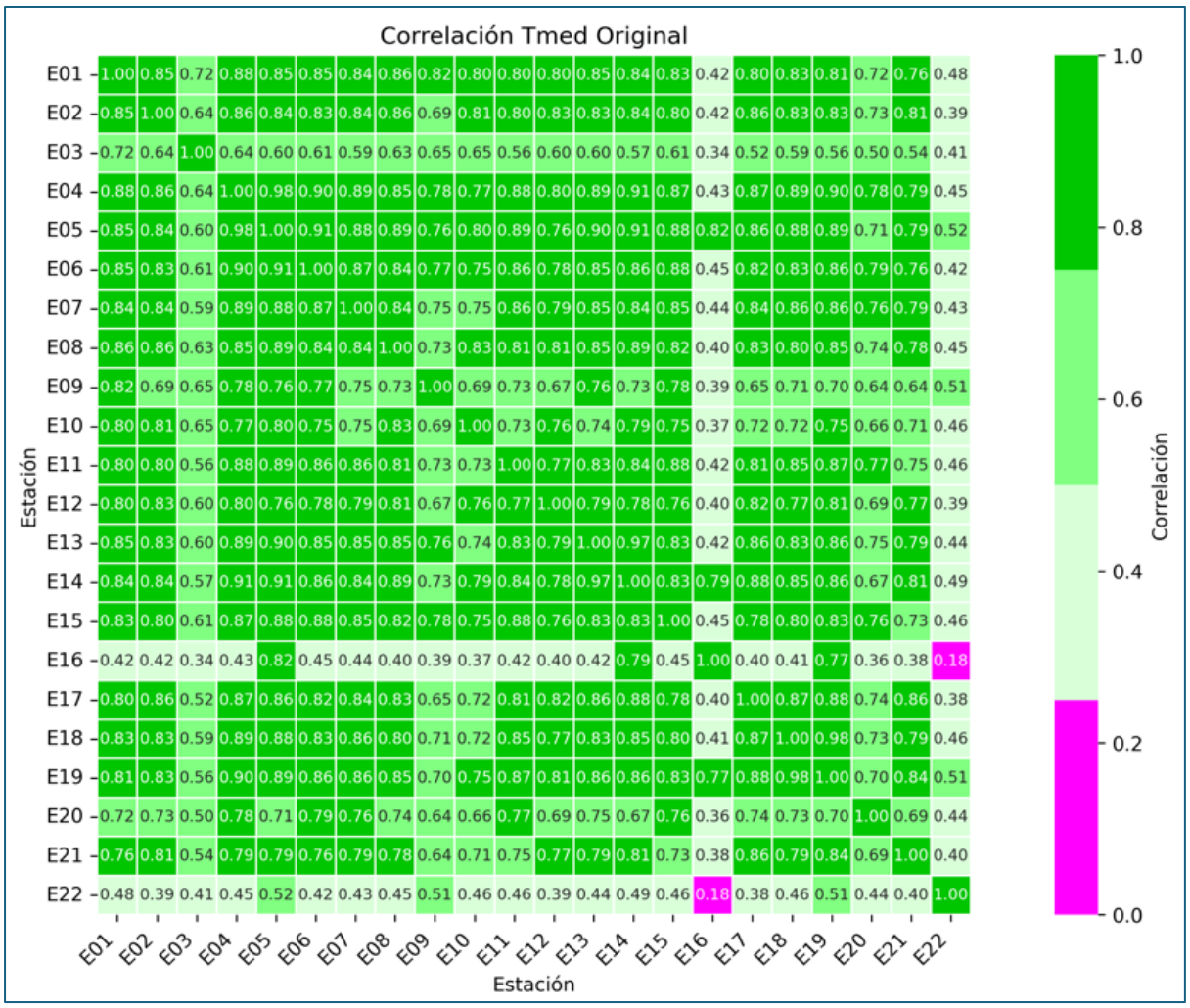
Gráfico de comparación del modelo



Nota. Elaboración propia

Figura 25.

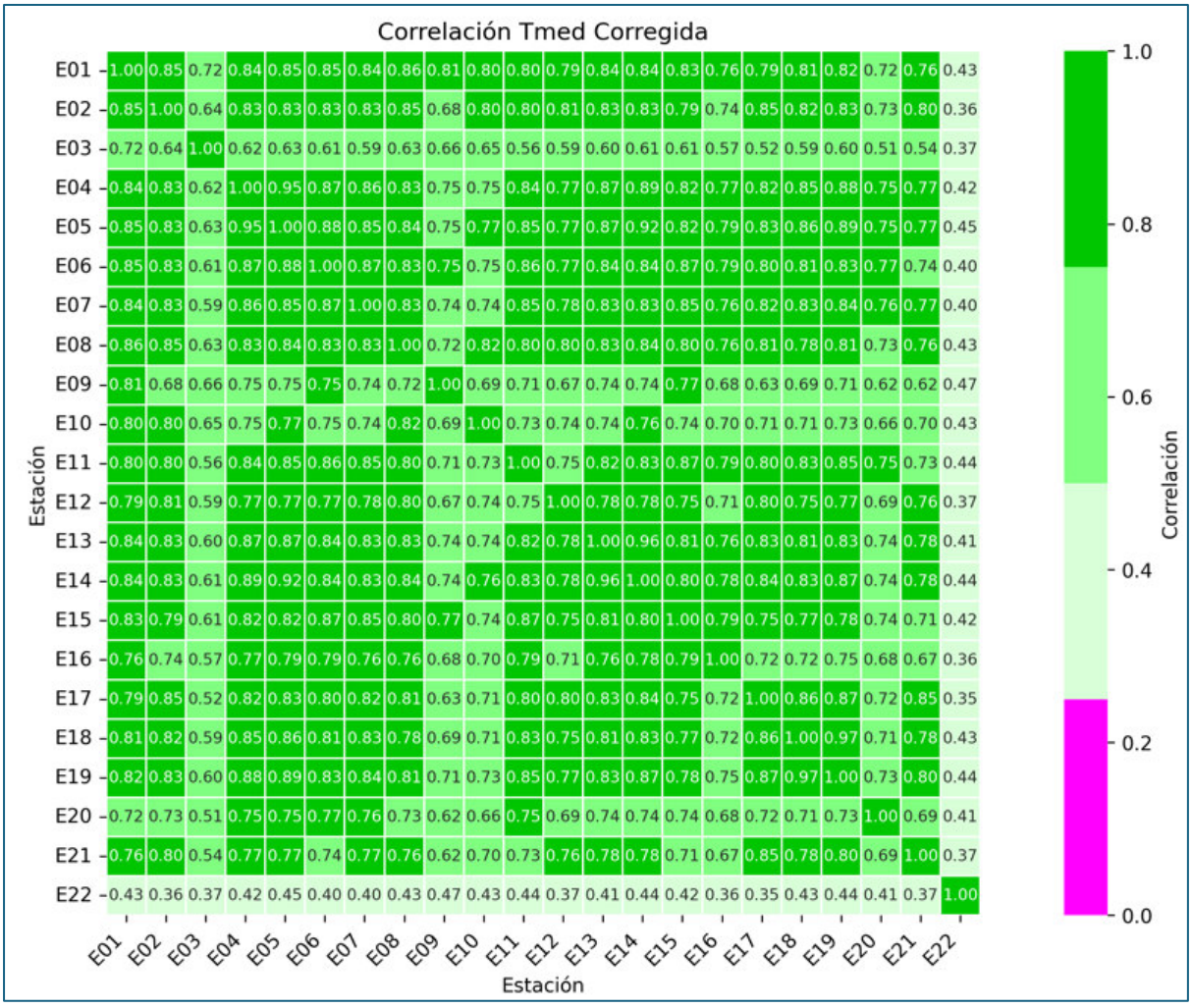
Mapa de correlación espacial original



Nota. Elaboración propia

Figura 26.

Mapa de correlación espacial corregido o rellenado



Nota. Elaboración propia

Tabla 32.

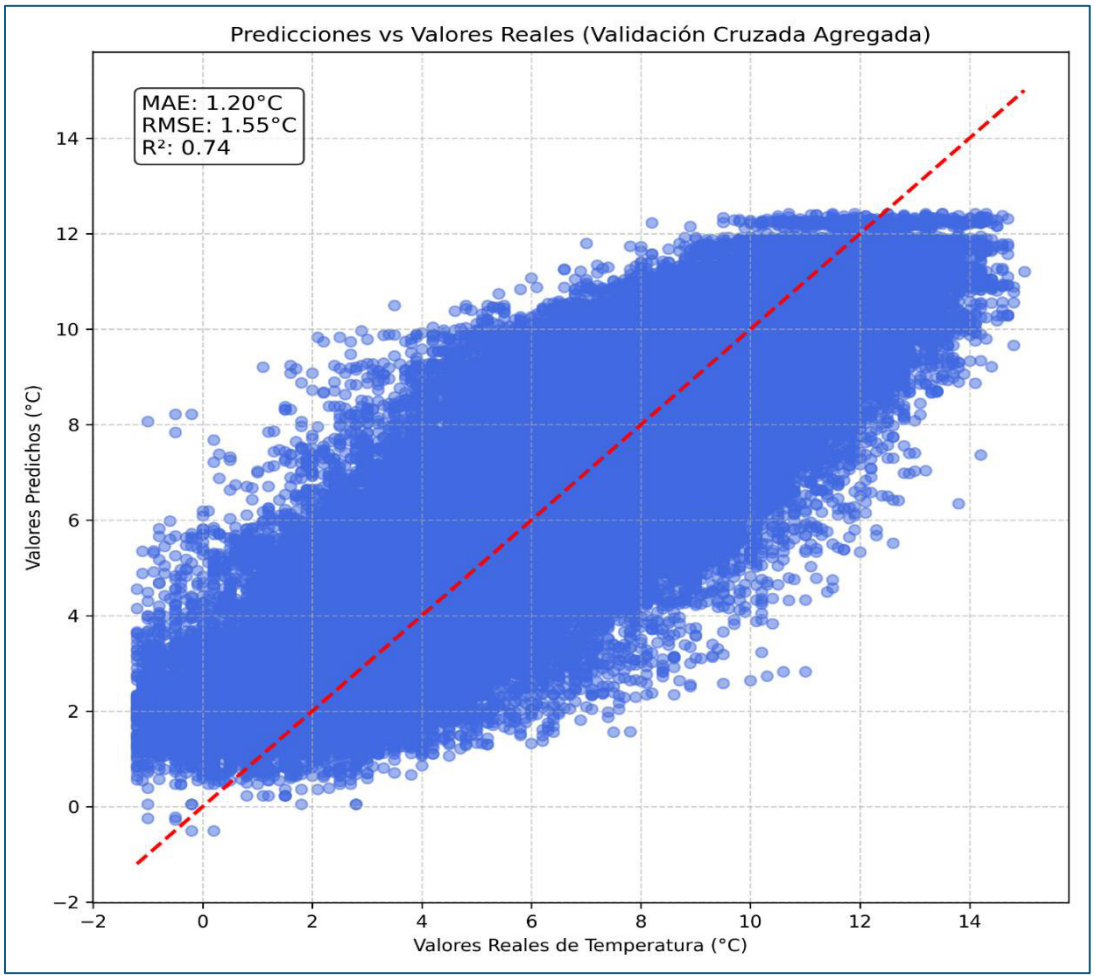
Resumen de matriz de correlación de datos original vs rellenado

Estación	Correlación Media Original	Correlación Media Rellenado	Mejor a r	Tipo de Estación
E01	0.781800133	0.789935639	Si	Con
E02	0.766018649	0.77373769	Si	Con
E03	0.566951752	0.581531965	Si	Con
E04	0.811576953	0.802433519	No	Con
E05	0.828224159	0.803040273	No	EMA
E06	0.77346754	0.783392421	Si	Con
E07	0.767381883	0.774904242	Si	Con
E08	0.756649356	0.76395397	Si	Con
E09	0.661591342	0.67352443	Si	Con
E10	0.68147223	0.696504107	Si	Con
E11	0.750742058	0.769330701	Si	Con
E12	0.698782461	0.717550872	Si	Con
E13	0.748464297	0.769606205	Si	Con
E14	0.772473542	0.760482667	No	EMA
E15	0.688390374	0.709644788	Si	Con
E16	0.417578557	0.649276572	Si	Con
E17	0.747491439	0.729125619	No	Con
E18	0.741556043	0.722876685	No	Con
E19	0.680560116	0.660423977	No	EMA
E20	0.565810384	0.549997866	No	Con
E21	0.401240985	0.368201889	No	Con
Promedio	0.70	0.71		
	Leyenda			
Con	Convencional			
EMA	Automática			

Nota. Elaboración propia

Figura 27.

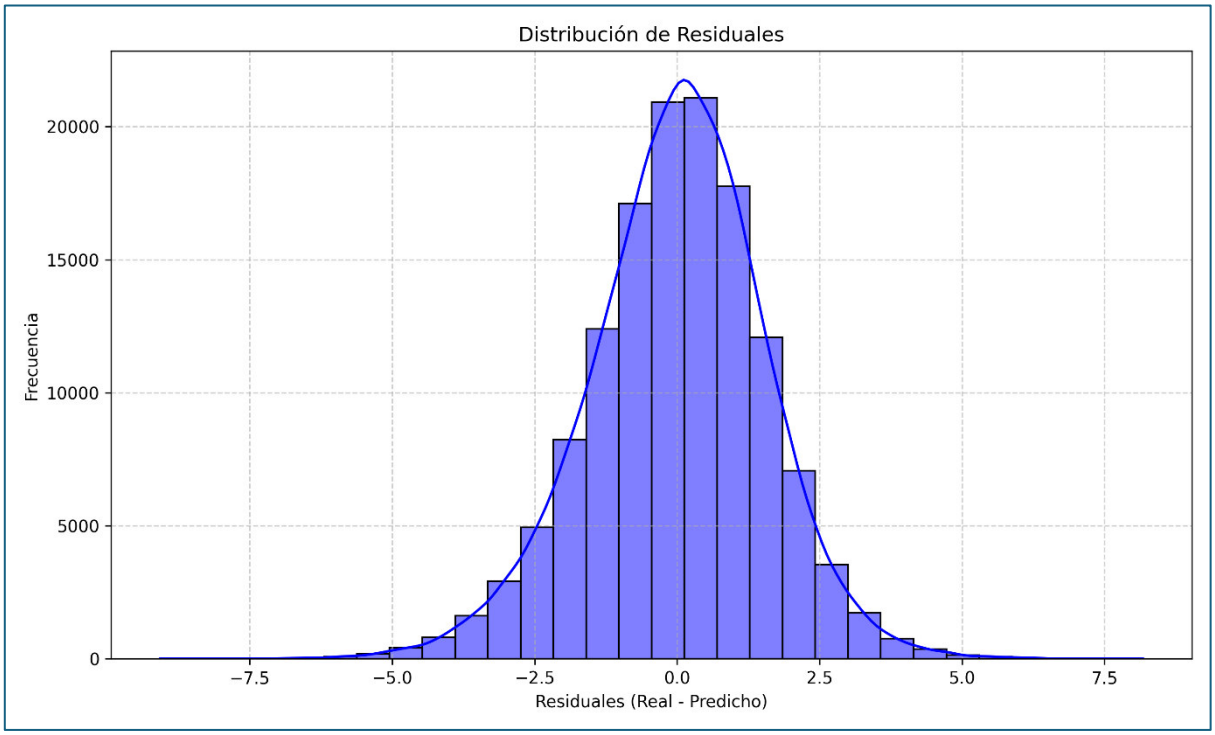
Gráfico de predicciones versus valores reales



Nota. Elaboración propia

Figura 28.

Mapa de distribución de residuales



Nota. Elaboración propia

Anexo E Validación de eficacia del modelo

Descripción:

Esta fase aplica el modelo híbrido (Random Forest + Gradient Boosting) para corregir datos erróneos y rellenar valores faltantes en las series de temperatura (Tmed). Los anexos presentan los indicadores estadísticos de la corrección, la distribución de residuales y visualizaciones que comparan los datos originales con los corregidos, destacando la reducción de incertidumbre en las series temporales.

Tabla 33.

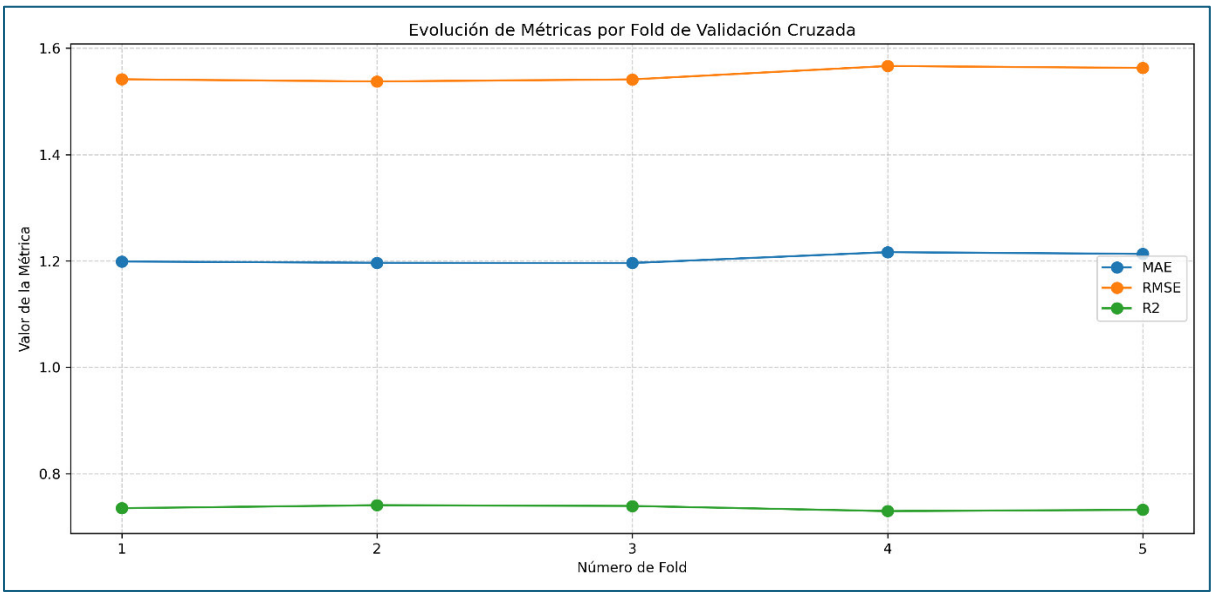
Relación de figuras del anexo E

N.º Ítem	Descripción	Justificación	Objetivo
Figura 29	Gráfico de evolución de métricas por fold de validación cruzada	Gráfico de líneas/caja que muestra la evolución de MAE, RMSE y R^2 a lo largo de los cinco folds.	Visualiza la estabilidad del modelo.
Figura 30	Gráfico de importancia de características de las variables	Gráfico de barras que muestra la contribución de variables predictoras (Day, Month, Altitud, etc.) en la predicción de Tmed.	Identifica las variables clave del modelo y Interpretar y justificar el aporte de cada variable en el modelo.
Figura 31	Diagrama del modelo híbrido (RF + GB)	Representa el flujo del modelo híbrido, incluyendo selección de características, entrenamiento y lógica de predicción.	Explica el funcionamiento del modelo matemático.

Nota. Elaboración propia

Figura 29.

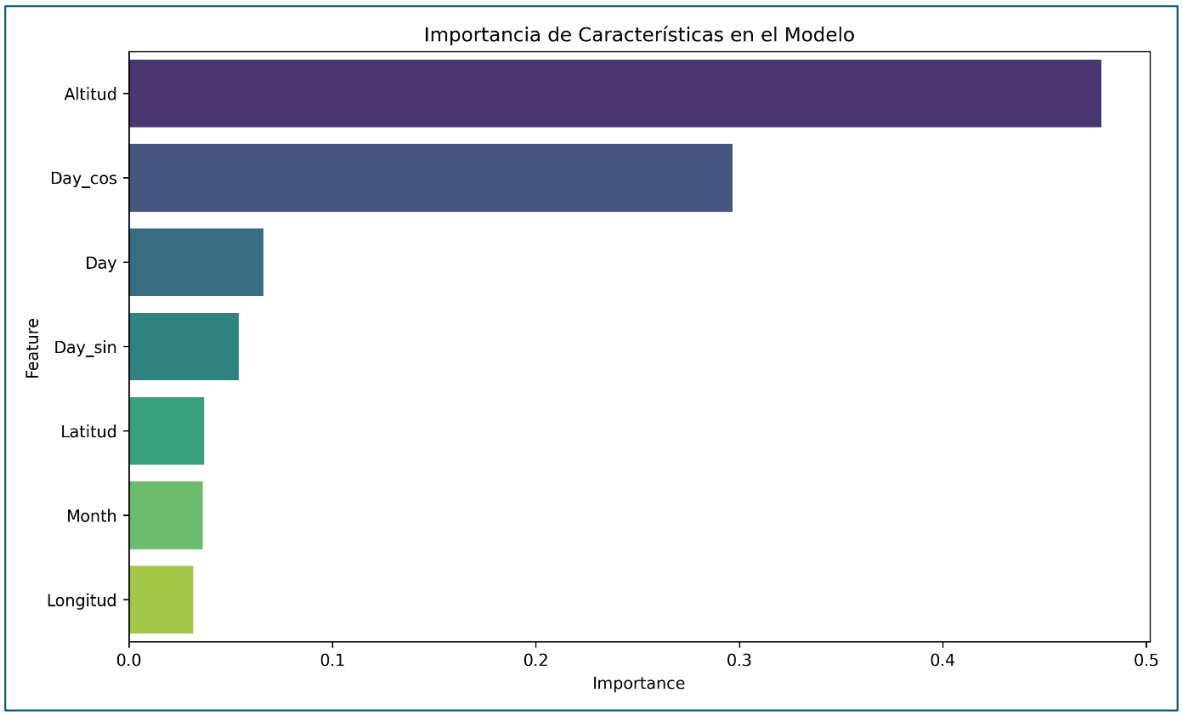
Gráfico de evolución de métricas por fold de validación cruzada



Nota. Elaboración propia

Figura 30.

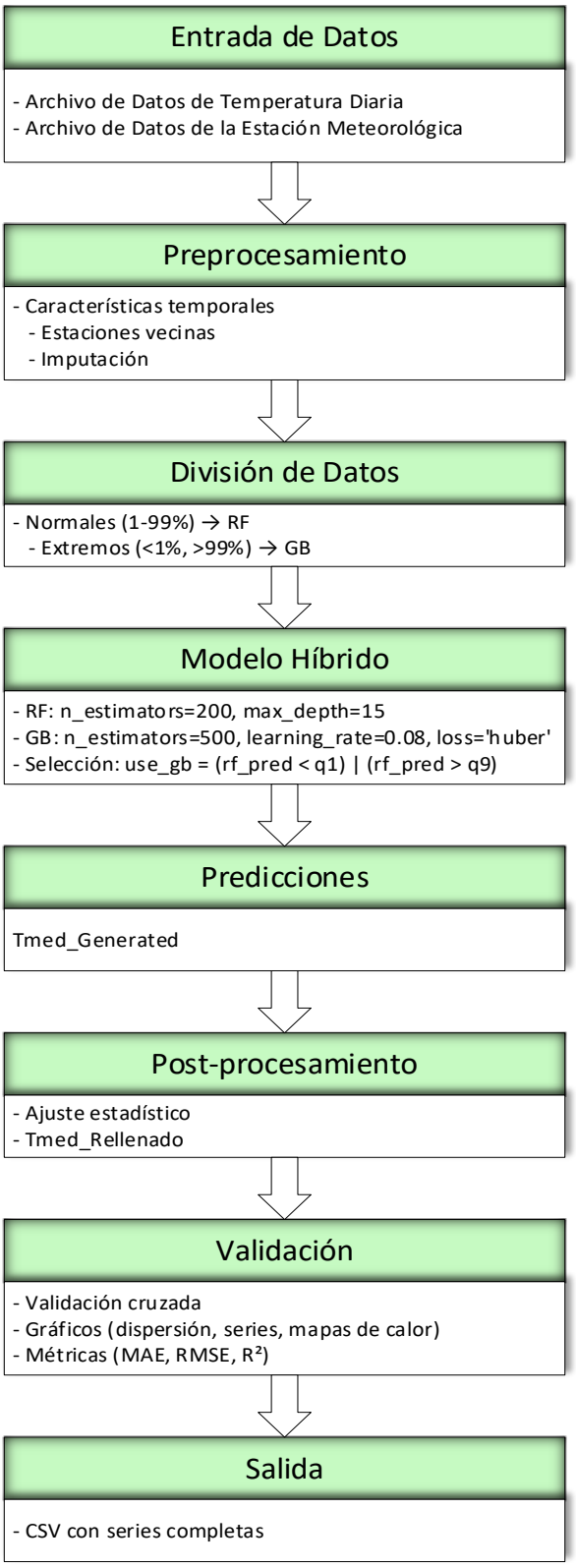
Gráfico de importancia de características de las variables



Nota. Elaboración propia

Figura 31.

Diagrama del modelo híbrido (RF + GB)



Nota. Elaboración propia