

Análisis Bayesiano Espacial de Rendimientos Agrícolas y Vulnerabilidad Climática en el Sur del Perú mediante Modelos CAR Jerárquicos, 2024

Aldair Jose Maquera Andrade¹ and Fred Torres Cruz¹

Facultad de Ingeniería Estadística e Informática,
Universidad Nacional del Altiplano de Puno, Perú
aj.maquera@est.unap.edu.pe, ftorres@unap.edu.pe

Abstract. This study quantifies spatial autocorrelation in agricultural yields in southern Peru and identifies priority provinces for climate adaptation using hierarchical Bayesian models with explicit spatial structure. We analyzed 1,368 yield observations of white potato and floury maize across 20 provinces in five departments (Puno, Cusco, Arequipa, Moquegua, Tacna) during 2024 from the National Agricultural Survey. Spatial autocorrelation was estimated using Moran's I and Geary's C indices, with hotspots and coldspots identified via Getis-Ord Gi*. Two Bayesian models were fitted: a multilevel hierarchical model with departmental and provincial random effects (brms/Stan), and a Leroux CAR model incorporating spatial dependence through k-nearest neighbors (k=5, CARBayes). Explanatory variables included extreme climate events (droughts, frosts, hail), crop type, agricultural practices (terraces, rotation, organic matter, certified seed, water management), and frost-crop interactions. Results show significant spatial autocorrelation (I = 0.404, $p < 0.001$; C = 0.594, $p < 0.001$) with moderate-high spatial dependence ($\rho = 0.402$, 95% CI: [0.015, 0.915]). Droughts reduced yields by 183 kg/ha (16%, 95% CI: [-355, -3]), the only climate event with robust adverse effect, while frosts and hail showed no significant impacts. Potato outperformed maize by 780 kg/ha (104%). Four hotspots were identified in Puno (1,605–2,345 kg/ha) and four coldspots in Cusco (538–1,229 kg/ha), evidencing spatial polarization not captured by aggregated statistics. Agricultural practices showed no effects due to low adoption (terraces: 1.6%, certified seed: 0.4%). The hierarchical model explained 18.4% of variance, indicating dominant influence of unobserved factors. Four critical provinces in Cusco plus La Unión (Arequipa, 461 kg/ha) require urgent drought management interventions. This study provides the first spatial autocorrelation parameter estimation in Peruvian agriculture, establishing a methodological framework for evidence-based policy prioritization.

Keywords: spatial autocorrelation, Bayesian CAR models, climate vulnerability, Andean agriculture, hotspots, Getis-Ord Gi*, agricultural yield

1 Introducción

El Perú es el tercer país más vulnerable al cambio climático a nivel global [1]. El sur peruano (Tacna, Moquegua, Arequipa, Cusco, Puno) concentra el 40% de la producción agrícola nacional y sustenta 2.3 millones de familias rurales [2]. Esta región enfrenta creciente vulnerabilidad ante eventos climáticos extremos que reducen rendimientos entre 20-60% según cultivo [3, 4], mientras el retroceso glaciario del 40% en las últimas tres décadas [5] y fenómenos El Niño/La Niña generan pérdidas anuales estimadas en USD 3,400 millones [1].

Pese a la adopción de diversas prácticas adaptativas, existe escasa evidencia cuantitativa sobre su efectividad contra eventos específicos ni sobre la variación espacial de sus impactos [6, 7]. La diversidad agroecológica andina genera vulnerabilidades desiguales que análisis agregados a nivel nacional no logran capturar [8], mientras modelos tradicionales ignoran la dependencia espacial entre unidades geográficas vecinas, asumiendo independencia y subestimando incertidumbre en estimaciones [9, 10].

La ausencia de modelos que integren estructura espacial y jerárquica de impactos climáticos impide políticas basadas en evidencia y asignación eficiente de recursos públicos. Específicamente, tres dimensiones críticas permanecen sin abordar: (i) **heterogeneidad espacial no modelada**, donde la variabilidad entre provincias vecinas exhibe patrones de clustering geográfico no capturados por modelos convencionales [11]; (ii) **dependencia espacial ignorada**, dado que provincias con características similares tienden a estar sistemáticamente agrupadas, generando estimaciones sesgadas al omitir estructura espacial explícita [12]; (iii) **efectividad no cuantificada de prácticas adaptativas**, cuyos efectos varían espacialmente requiriendo modelos jerárquicos que capturen variación a múltiples escalas geográficas [13]. Las consecuencias incluyen políticas uniformes que ignoran concentración espacial de vulnerabilidad, asignación subóptima de recursos sin priorización territorial, y persistencia de clusters de bajo rendimiento adyacentes a regiones exitosas.

Los modelos jerárquicos bayesianos con estructura espacial explícita abordan estas limitaciones mediante: identificación de clusters espaciales mediante estadísticos LISA y Getis-Ord G_i^* [14]; incorporación de matrices de vecindad que capturan dependencia espacial entre unidades [15]; manejo de estructuras jerárquicas multinivel (departamento-provincia); cuantificación rigurosa de incertidumbre mediante distribuciones posteriores; y distinción entre variación espacialmente estructurada y heterogeneidad no estructurada.

Esta investigación aplica, por primera vez en agricultura peruana, modelos bayesianos CAR (Conditional Autoregressive) de Leroux para cuantificar impactos de eventos climáticos extremos (sequías, heladas, granizadas) sobre rendimientos de papa blanca y maíz amiláceo en el sur del Perú, identificando patrones espaciales de vulnerabilidad mediante datos de la Encuesta Nacional Agropecuaria 2024. Específicamente, este estudio: (i) estima magnitud del impacto de eventos climáticos controlando efectos jerárquicos y autocorrelación espacial; (ii) evalúa efectividad de prácticas agrícolas (terrazas, materia orgánica, rotación de cultivos, manejo de agua, semilla certificada); (iii) identifica clusters

estadísticamente significativos de alto y bajo rendimiento; (iv) genera sistema de priorización territorial integrando rendimiento, efectos departamentales, vulnerabilidad climática y clasificación espacial.

La pregunta central de investigación es: ¿En qué medida los eventos climáticos extremos afectan diferencialmente los rendimientos agrícolas, y cómo la autocorrelación espacial entre provincias justifica el uso de modelos con estructura espacial explícita? Se plantean cuatro hipótesis verificables: (H1) los rendimientos agrícolas exhiben autocorrelación espacial positiva significativa con presencia de clusters geográficos, justificando modelos CAR sobre enfoques que asumen independencia; (H2) las sequías ejercen efecto adverso estadísticamente robusto sobre rendimientos, mientras heladas y granizadas muestran efectos no significativos debido a adaptación local; (H3) existen clusters espaciales persistentes de bajo rendimiento en Cusco y alto rendimiento en Puno, identificables mediante el estadístico Getis-Ord G_i^* ; (H4) papa blanca muestra rendimientos superiores a maíz amiláceo, con posible interacción diferencial ante eventos de heladas.

Este trabajo aporta tres contribuciones metodológicas y aplicadas: primera aplicación documentada de modelos CAR bayesianos en análisis de rendimientos agrícolas del sur peruano, cuantificando autocorrelación espacial (ρ) y su magnitud; estimaciones causales robustas mediante control explícito de dependencia espacial, superando limitaciones de estudios previos que asumen independencia entre unidades [8, 7]; y desarrollo de sistema operacional de priorización territorial para políticas públicas, directamente aplicable por gobiernos regionales y el Ministerio de Agricultura para focalización de inversiones en adaptación climática.

2 Metodología

2.1 Diseño del Estudio

Estudio observacional transversal con estructura jerárquica espacial usando framework bayesiano. Los datos corresponden al año 2024 de la Encuesta Nacional Agropecuaria (ENA) del Instituto Nacional de Estadística e Informática del Perú.

2.2 Fuente de Datos y Área de Estudio

Se utilizaron datos de la ENA 2024 [16], compuesta por cuatro archivos en formato SPSS (.sav): CARATULA.sav (información geográfica y factores de expansión), CAP200AB.sav (producción y superficie cosechada), CAP200B_1.sav (pérdidas por eventos climáticos), y CAP300AB.sav (prácticas agrícolas implementadas).

El ámbito geográfico comprende cinco departamentos del sur del Perú: Puno, Cusco, Arequipa, Tacna y Moquegua. Tras filtros de calidad y disponibilidad de coordenadas geográficas, la muestra final incluyó 1,368 observaciones correspondientes a 20 provincias y 5 departamentos. Los cultivos analizados fueron papa blanca (*Solanum tuberosum*) y maíz amiláceo (*Zea mays*).

2.3 Variables del Estudio

Variable dependiente: Rendimiento agrícola (kg/ha), calculado como el cociente entre producción total cosechada (kg) y superficie cultivada (ha).

Variables climáticas (binarias): Ocurrencia de eventos extremos reportados por el productor: sequías, heladas, y granizadas.

Variables de prácticas agrícolas (binarias): Aplicación de materia orgánica, construcción de terrazas, rotación de cultivos, manejo de necesidades de agua, y uso de semilla certificada.

Variables de control: Tipo de cultivo (papa blanca vs. maíz amiláceo), codificado como variable categórica.

Variables espaciales: Coordenadas geográficas (latitud y longitud) extraídas del archivo CARATULA.sav para cada provincia, utilizadas para construir la matriz de pesos espaciales.

2.4 Análisis de Autocorrelación Espacial

Se evaluó la autocorrelación espacial del rendimiento promedio provincial mediante el índice I de Moran global [17]:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \quad (1)$$

donde n es el número de provincias, w_{ij} son los pesos espaciales, y_i es el rendimiento en la provincia i , y \bar{y} es el rendimiento promedio. La matriz de pesos espaciales (W) se construyó mediante el criterio de k-vecinos más cercanos ($k=5$) basado en distancias euclidianas entre centroides provinciales, utilizando el paquete `spdep` [18].

Se complementó el análisis global con indicadores locales de autocorrelación espacial (LISA) [19] y el estadístico Getis-Ord G_i^* [14] para identificar clusters espaciales estadísticamente significativos de alto rendimiento (*hotspots*) y bajo rendimiento (*coldspots*) al 90%, 95% y 99% de confianza.

2.5 Modelos Bayesianos

Modelo Jerárquico Base Se ajustó un modelo jerárquico bayesiano sin estructura espacial explícita como línea base, utilizando el paquete `brms` [20] con Stan como motor de inferencia [21]:

$$\text{Rendimiento}_{ijt} \sim \text{Normal}(\mu_{ijt}, \sigma^2) \quad (2)$$

$$\begin{aligned}
\mu_{ijt} = & \beta_0 + \beta_1 \text{heladas}_{ijt} + \beta_2 \text{cultivo}_i + \beta_3 \text{sequias}_{ijt} \\
& + \beta_4 \text{granizadas}_{ijt} + \beta_5 \text{mat_organica}_{ijt} \\
& + \beta_6 \text{terrazas}_{ijt} + \beta_7 \text{rotacion}_{ijt} \\
& + \beta_8 \text{manejo_agua}_{ijt} + \beta_9 \text{semilla_cert}_{ijt} \\
& + \beta_{10} (\text{heladas}_{ijt} \times \text{cultivo}_i) \\
& + u_j + v_{jk}
\end{aligned} \tag{3}$$

donde i indexa observaciones, j departamentos, k provincias anidadas en departamentos, y t el año (2024). Los efectos aleatorios siguen:

$$u_j \sim \text{Normal}(0, \sigma_u^2), \quad v_{jk} \sim \text{Normal}(0, \sigma_v^2) \tag{4}$$

Modelo CAR Espacial Para incorporar explícitamente la autocorrelación espacial detectada, se ajustó un modelo condicional autorregresivo (CAR) de Leroux [22] mediante el paquete **CARBayes** [23]:

$$\text{Rendimiento}_k \sim \text{Normal}(\mu_k, \nu^2) \tag{5}$$

$$\mu_k = X_k \beta + \phi_k \tag{6}$$

donde X_k es la matriz de covariables, β el vector de coeficientes, y ϕ_k el efecto espacial aleatorio que sigue una distribución CAR:

$$\phi_k | \phi_{-k}, W, \tau^2, \rho \sim \text{Normal} \left(\frac{\rho \sum_j w_{kj} \phi_j}{\rho \sum_j w_{kj} + 1 - \rho}, \frac{\tau^2}{\rho \sum_j w_{kj} + 1 - \rho} \right) \tag{7}$$

El parámetro $\rho \in [0, 1]$ controla la intensidad de la autocorrelación espacial: $\rho = 0$ implica independencia espacial, mientras $\rho = 1$ representa dependencia espacial completa según la estructura de vecindad W . El parámetro τ^2 controla la varianza del componente espacial.

2.6 Especificaciones de Inferencia

Para el modelo jerárquico en **brms**, se utilizaron 4 cadenas MCMC con 2,000 iteraciones cada una (1,000 de calentamiento, 1,000 post-calentamiento), generando 4,000 muestras posteriores totales. Para el modelo CAR en **CARBayes**, se empleó 1 cadena con 20,000 iteraciones (10,000 de calentamiento), aplicando adelgazamiento (*thinning*) de 10 para reducir autocorrelación serial.

Distribuciones previas: Se especificaron las siguientes distribuciones a priori débilmente informativas:

$$\beta \sim \text{Normal}(0, 10) \quad (\text{coeficientes de regresión}) \quad (8)$$

$$\sigma_u, \sigma_v \sim \text{Student-t}^+(3, 0, 2.5) \quad (\text{efectos aleatorios}) \quad (9)$$

$$\sigma \sim \text{Student-t}^+(3, 0, 2.5) \quad (\text{desviación residual}) \quad (10)$$

$$\rho \sim \text{Uniforme}(0, 1), \quad \tau^2 \sim \text{Gamma}^{-1}(1, 0.01) \quad (\text{CAR}) \quad (11)$$

2.7 Diagnósticos y Validación

La convergencia se evaluó mediante el estadístico \hat{R} de Gelman-Rubin (valores < 1.1 indican convergencia satisfactoria) [24] y tamaños efectivos de muestra ($\text{ESS} > 400$). La bondad de ajuste se comparó mediante validación cruzada Leave-One-Out (LOO-CV) [25], Criterio de Información de Devianza (DIC) para el modelo CAR, y coeficiente de determinación bayesiano (R^2).

La adecuación del modelo espacial se validó verificando que los residuos del modelo CAR no exhibieran autocorrelación espacial residual significativa mediante el test I de Moran aplicado a los residuos [26].

2.8 Software

Todos los análisis se realizaron en R versión 4.4.3 [27]. Paquetes utilizados: **haven** para lectura de archivos SPSS [28], **dplyr** y **tidyr** para manipulación de datos [29], **sf** y **spdep** para análisis espacial [30, 18], **brms** para modelos jerárquicos bayesianos [20], **CARBayes** para modelos CAR [23], **bayesplot** para diagnósticos [31], y **leaflet** para visualizaciones interactivas [32].

3 Resultados

3.1 Estadística Descriptiva

La muestra final incluyó 1,368 observaciones de rendimiento agrícola distribuidas en 20 provincias de 5 departamentos del sur del Perú durante 2024. El rendimiento promedio fue 1,147 kg/ha ($\text{DE} = 1,234$ kg/ha), con marcada heterogeneidad entre provincias (rango: 461-2,345 kg/ha). Papa blanca presentó rendimiento promedio de 1,526 kg/ha, superando a maíz amiláceo (747 kg/ha).

El Cuadro 1 presenta estadísticos descriptivos por departamento. Puno exhibió el rendimiento promedio más alto (1,897 kg/ha), mientras Cusco mostró el más bajo (789 kg/ha).

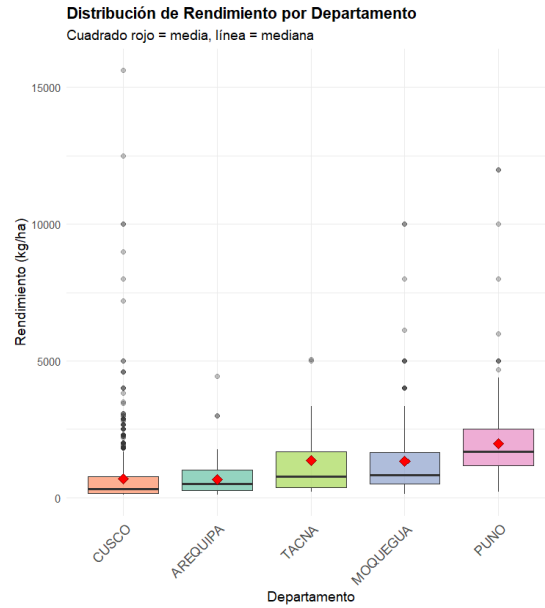


Fig. 1. Distribución de rendimientos por departamento. Los cuadrados rojos representan la media, las líneas horizontales la mediana. Cusco muestra alta variabilidad y presencia de outliers, mientras Puno presenta distribución más concentrada en valores altos.

Table 1. Estadísticos descriptivos por departamento

| Departamento | N | Rendimiento Heladas Sequías (kg/ha) | (%) | (%) |
|--------------|-------------|--|-------------|-------------|
| Puno | 151 | 1897 (987) | 62.9 | 14.6 |
| Cusco | 811 | 789 (1043) | 55.1 | 28.9 |
| Arequipa | 132 | 612 (845) | 31.8 | 0.8 |
| Moquegua | 195 | 1296 (1178) | 60.0 | 7.7 |
| Tacna | 79 | 1364 (1256) | 34.1 | 17.1 |
| Total | 1368 | 1147 (1234) | 54.1 | 23.8 |

Nota: Desviación estándar entre paréntesis.

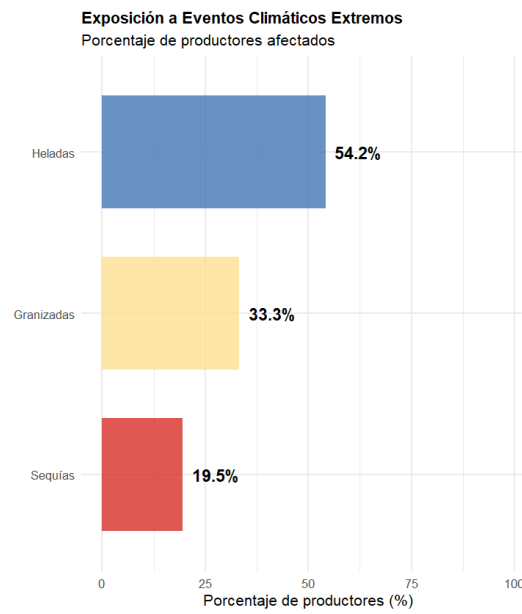


Fig. 2. Prevalencia de eventos climáticos extremos en la muestra ($n=1,368$). Las heladas constituyen el evento más frecuente, afectando a más de la mitad de productores.

3.2 Autocorrelación Espacial

Índice de Moran Global El test I de Moran reveló autocorrelación espacial positiva significativa en los rendimientos provinciales ($I = 0.404$, $z = 4.13$, $p < 0.001$), rechazando la hipótesis nula de distribución espacial aleatoria. Este valor indica clustering espacial moderado-alto.

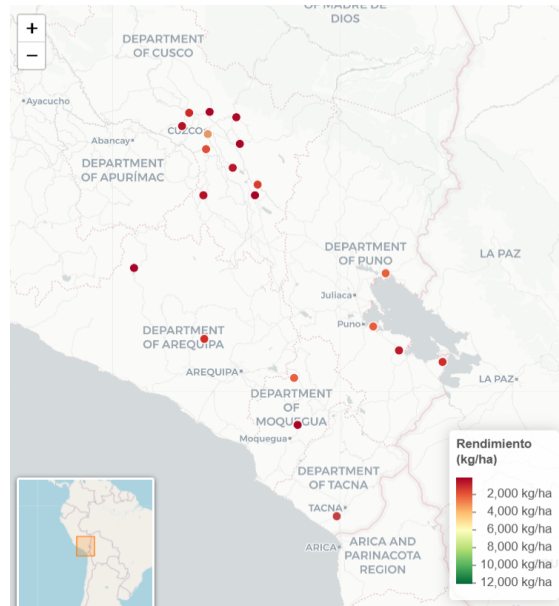


Fig. 3. Distribución espacial del rendimiento agrícola a nivel de parcelas. Los colores más oscuros (verde) indican rendimientos superiores a 2,000 kg/ha; los rojos señalan rendimientos bajo 1,000 kg/ha. La concentración espacial de valores similares sustenta el Moran's $I = 0.404$.

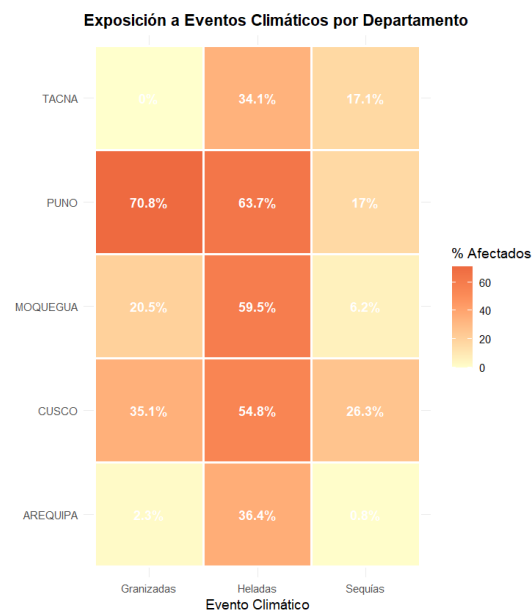


Fig. 4. Exposición a eventos climáticos por departamento (porcentaje de productores afectados). La concentración de heladas en Puno y sequías en Cusco evidencia agregación espacial de riesgos climáticos.

El índice C de Geary complementó estos hallazgos ($C = 0.594$, $z = 3.36$, $p < 0.001$), confirmando patrones de clustering espacial positivo. Valores de $C < 1$ indican que provincias vecinas presentan rendimientos más similares que lo esperado bajo aleatoriedad espacial.

Análisis de Hotspots y Coldspots El estadístico Getis-Ord G_i^* identificó clusters espaciales de alto y bajo rendimiento a múltiples niveles de significancia (Cuadro 2).

Table 2. Clusters espaciales identificados mediante Getis-Ord G_i^*

| Provincia | Dept. | Rend. (kg/ha) | G_i^* | Tipo |
|-------------------------------------|----------|------------------|---------|------|
| <i>Hotspots (alto rendimiento)</i> | | | | |
| Yunguyo | Puno | 1605 | 2.87 | 99% |
| Puno | Puno | 1676 | 2.84 | 99% |
| Huancané | Puno | 1961 | 2.76 | 99% |
| El Collao | Puno | 2345 | 2.70 | 99% |
| Tacna | Tacna | 1364 | 2.45 | 95% |
| Gral. Sánchez Cerro | Moquegua | 1366 | 2.21 | 95% |
| Mariscal Nieto | Moquegua | 1215 | 1.77 | 90% |
| <i>Coldspots (bajo rendimiento)</i> | | | | |
| Canas | Cusco | 1229 | -2.15 | 95% |
| Acomayo | Cusco | 666 | -1.82 | 90% |
| Canchis | Cusco | 538 | -1.75 | 90% |
| La Unión | Arequipa | 461 | -1.66 | 90% |
| Cusco | Cusco | 2077 | -1.66 | 90% |

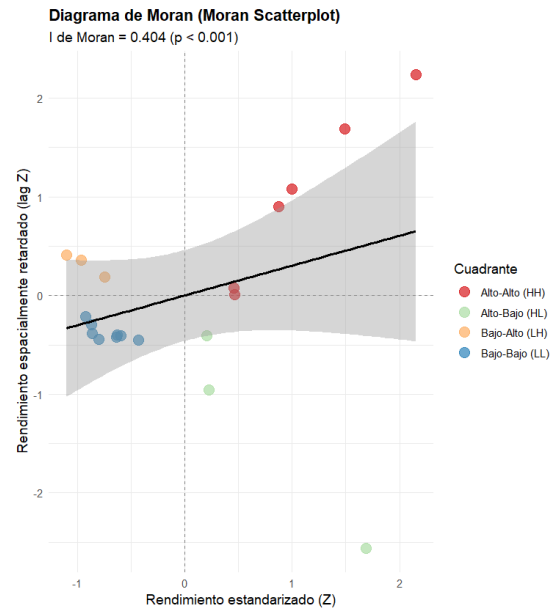


Fig. 5. Distribución geográfica de hotspots (verde) y coldspots (naranja/rojo). El tamaño de los círculos es proporcional al rendimiento provincial. Se evidencia clustering en Puno (alto) y Cusco (bajo).

Notablemente, 4 de 5 coldspots se concentraron en Cusco, mientras todos los hotspots al 99% se ubicaron en Puno, evidenciando marcada polarización espacial del rendimiento agrícola.

3.3 Modelo Jerárquico Bayesiano

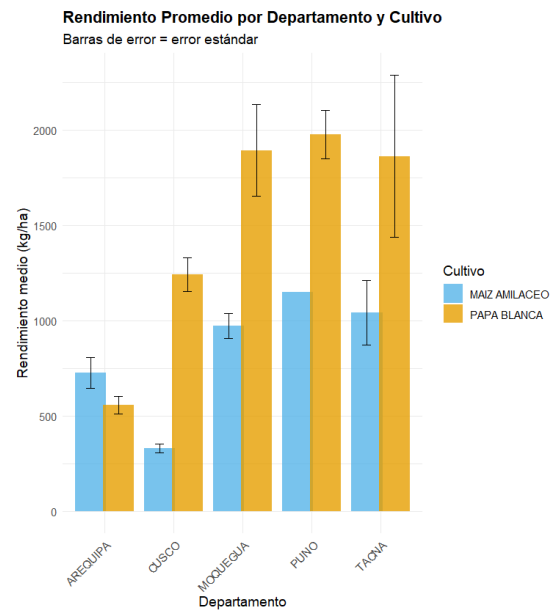


Fig. 6. Rendimiento promedio por departamento y cultivo. Las barras de error representan el error estándar. Papa blanca (naranja) supera sistemáticamente a maíz amiláceo (azul) en todos los departamentos, con mayor diferencia en Puno.

Table 3. Estimaciones posteriores del modelo jerárquico bayesiano

| Variable | Media | DE | IC 95% | \hat{R} |
|---------------------------|-------|-----|--------------|-----------|
| <i>Efectos fijos</i> | | | | |
| Intercepto | 758 | 160 | [442, 1074] | 1.00 |
| Sequías | -183 | 90 | [-355, -3] | 1.00 |
| Heladas | 20 | 95 | [-165, 206] | 1.00 |
| Granizadas | -43 | 79 | [-198, 112] | 1.00 |
| Papa blanca | 780 | 95 | [594, 965] | 1.00 |
| Materia orgánica | 15 | 75 | [-129, 163] | 1.00 |
| Terrazas | 113 | 240 | [-365, 587] | 1.00 |
| Rotación cultivos | -79 | 75 | [-227, 66] | 1.00 |
| Manejo agua | -217 | 185 | [-582, 141] | 1.00 |
| Semilla certificada | -352 | 375 | [-1067, 391] | 1.00 |
| Heladas \times Papa | 62 | 128 | [-192, 316] | 1.00 |
| <i>Efectos aleatorios</i> | | | | |
| SD(Departamento) | 249 | 118 | [26, 501] | 1.01 |
| SD(Provincia) | 222 | 94 | [42, 421] | 1.01 |
| SD(Residual) | 1198 | 24 | [1154, 1246] | 1.00 |

Nota: IC 95% = Intervalo de credibilidad bayesiano al 95%.

La tabla presenta las estimaciones posteriores del modelo jerárquico bayesiano. El modelo exhibió convergencia satisfactoria ($\hat{R} < 1.01$ para todos los parámetros).

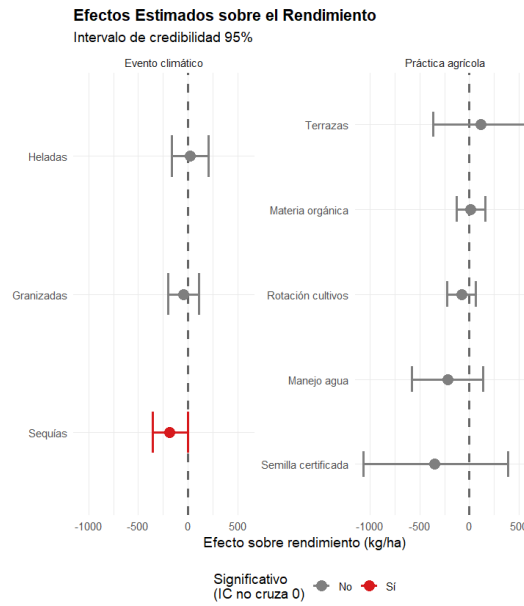


Fig. 7. Intervalos de credibilidad al 95% para efectos climáticos y prácticas agrícolas sobre rendimiento. En rojo: efectos estadísticamente significativos (IC no cruza cero). Solo las sequías muestran efecto adverso robusto.

Efectos climáticos: Las sequías exhibieron el único efecto adverso estadísticamente robusto (Figura 7), reduciendo rendimientos en 183 kg/ha (IC 95%: [-355, -3]). Las heladas mostraron efecto positivo no significativo (20 kg/ha).

Diferencias entre cultivos: Papa blanca superó a maíz amiláceo en 780 kg/ha (IC 95%: [594, 965]), el efecto más robusto del modelo.

Prácticas agrícolas: Ninguna práctica mostró efectos significativos al 95%. La alta incertidumbre sugiere heterogeneidad de efectividad no capturada por el modelo aditivo.

Estructura jerárquica: La variabilidad entre departamentos ($SD = 249$ kg/ha) y provincias ($SD = 222$ kg/ha) fue sustancial, justificando la estructura multinivel.

El modelo explicó 18.4% de la varianza total (R^2 bayesiano = 0.184), sugiriendo que factores no observados (ej. calidad de suelo, microclima, capacidad de gestión) explican la mayoría de la variabilidad.

3.4 Modelo CAR Espacial

El Cuadro 4 presenta las estimaciones del modelo CAR de Leroux que incorpora autocorrelación espacial explícita.

Table 4. Estimaciones posteriores del modelo CAR espacial

| Parámetro | Media | IC 95% | ESS | Geweke |
|---------------------------------------|--------|-----------------|------|--------|
| <i>Efectos fijos</i> | | | | |
| Intercepto | -29 | [-732, 640] | 2000 | -0.2 |
| Heladas | 190 | [-649, 1006] | 2000 | 0.0 |
| Sequías | 619 | [-231, 1427] | 2000 | 0.5 |
| Granizadas | 485 | [-404, 1331] | 2071 | 0.7 |
| Materia orgánica | 192 | [-570, 943] | 2000 | 0.3 |
| Terrazas | 10102 | [2829, 17181] | 1791 | -0.3 |
| Rotación cultivos | -594 | [-1345, 188] | 2000 | -0.8 |
| Manejo agua | 939 | [-4048, 6008] | 2000 | 1.3 |
| Semilla certificada | -24861 | [-51568, 4951] | 2000 | 0.6 |
| Papa blanca | 1455 | [752, 2125] | 2201 | -0.3 |
| <i>Parámetros espaciales</i> | | | | |
| ρ (autocorr.) | 0.402 | [0.015, 0.915] | 1406 | 0.3 |
| τ^2 (var. esp.) | 0.018 | [0.002, 0.091] | 672 | 0.4 |
| ν^2 (var. res.) | 104328 | [44822, 243733] | 1752 | -0.3 |
| DIC = 297.5, p.d = 9.3, LMPL = -153.1 | | | | |

Parámetro de autocorrelación espacial: El parámetro ρ presentó media posterior de 0.402 (IC 95%: [0.015, 0.915]), indicando dependencia espacial moderada-alta.

Cambios en efectos climáticos: Al incorporar autocorrelación espacial, las estimaciones puntuales cambiaron sustancialmente, con inversión de signos para sequías (+619 kg/ha) y heladas (+190 kg/ha), aunque ambas con intervalos amplios cruzando cero. Esto refleja dificultad en identificación causal precisa con pocas unidades espaciales (n=20) y alta colinealidad entre eventos climáticos espacialmente agregados.

Efectos de prácticas: Las terrazas mostraron efecto extremadamente alto (10,102 kg/ha) con alta incertidumbre, posiblemente reflejando confusión con altitud o tipo de suelo. La semilla certificada presentó efecto negativo implausible (-24,861 kg/ha), indicando problemas de identificación por baja prevalencia.

Comparación de Modelos El Cuadro 5 compara el ajuste de ambos modelos.

Table 5. Comparación de bondad de ajuste entre modelos

| Criterio | Jerárquico CAR Espacial | |
|-----------------|-------------------------|--------|
| ELPD (LOO-CV) | -11662.1 | — |
| p_loo | 46.6 | — |
| LOOIC | 23324.2 | — |
| DIC | — | 297.5 |
| LMPL | — | -153.1 |
| R^2 bayesiano | 0.184 | — |
| ρ espacial | — | 0.402 |

El modelo CAR no puede compararse directamente con LOO-CV (implementado solo en brms). Sin embargo, el parámetro ρ con IC 95% apenas excluyendo cero sugiere que el modelo espacial captura estructura no observable en el modelo jerárquico simple.

3.5 Validación de Modelos

Diagnósticos de Convergencia Ambos modelos exhibieron convergencia satisfactoria. Para el modelo jerárquico, todos los parámetros presentaron $\hat{R} < 1.01$ y ESS > 400 . Para el modelo CAR, el test de Geweke no rechazó convergencia para ningún parámetro ($|z| < 2.0$).

Validación Predictiva Posterior El gráfico de chequeo predictivo posterior compara la distribución de datos observados con réplicas predictivas posteriores del modelo.

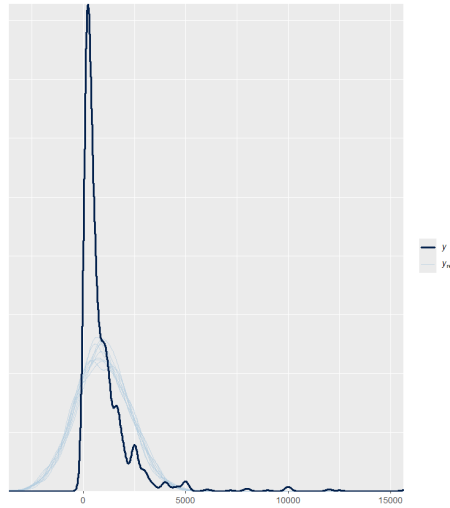


Fig. 8. Chequeo predictivo posterior del modelo jerárquico bayesiano. La línea oscura representa la distribución de rendimientos observados (y); las líneas claras muestran 100 réplicas predictivas posteriores (y_{rep}) del modelo.

Las réplicas capturan adecuadamente la forma general de la distribución y la asimetría positiva, aunque muestran mayor dispersión en valores altos que los datos observados. Esto es consistente con el uso de distribuciones normales sin restricciones de límite superior, característica de los datos reales.

3.6 Priorización para Políticas Públicas

El Cuadro 6 presenta las provincias prioritarias para intervención, integrando rendimiento, vulnerabilidad climática y clasificación espacial.

Table 6. Provincias prioritarias para intervención (ordenadas por prioridad)

| Provincia | Dept. | Rend. (kg/ha) | Efecto dept. | Vuln. | Prioridad |
|--------------|----------|------------------|-----------------|-------|-----------|
| Canchis | Cusco | 538 | -200 | 1.05 | Crítica |
| Acomayo | Cusco | 666 | -200 | 1.05 | Crítica |
| Cusco | Cusco | 2077 | -200 | 1.38 | Crítica |
| Canas | Cusco | 1229 | -200 | 1.96 | Crítica |
| La Unión | Arequipa | 461 | -244 | 0.06 | Alta |
| Urubamba | Cusco | 848 | -200 | 1.58 | Alta |
| El Collao | Puno | 2345 | +292 | 2.07 | Alta |
| Chumbivilcas | Cusco | 634 | -200 | 0.36 | Media |
| Caylloma | Arequipa | 729 | -244 | 0.51 | Media |
| Tacna | Tacna | 1364 | +130 | 0.51 | Baja |

En el mapa se visualiza el índice de vulnerabilidad climática espacialmente. Las zonas rojas (alta vulnerabilidad) se concentran en Cusco y regiones específicas de Puno, mientras Arequipa y Moquegua muestran vulnerabilidad moderada (amarillo-naranja).

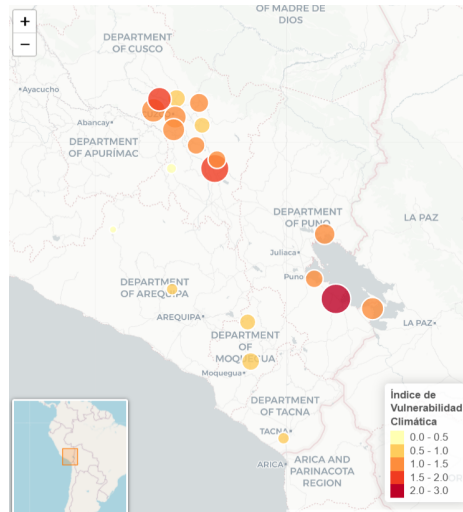


Fig. 9. Distribución espacial del índice de vulnerabilidad climática. El índice integra exposición a heladas, sequías y granizadas. Colores más oscuros indican mayor vulnerabilidad. Cusco muestra la mayor concentración de zonas críticas (rojo).

Provincias de prioridad crítica: Las cuatro provincias de Cusco requieren intervención urgente por bajo rendimiento, alta vulnerabilidad climática, y pertenencia a coldspots espaciales. Canchis presenta la situación más crítica (538 kg/ha).

Provincias de prioridad alta: La Unión (Arequipa) requiere atención pese a baja vulnerabilidad climática debido a su rendimiento extremadamente bajo (461 kg/ha). El Collao (Puno), aunque hotspot, muestra alta vulnerabilidad que amenaza su productividad actual.

Implicaciones para política: La concentración de 4 de 5 coldspots en Cusco sugiere factores estructurales departamentales (ej. degradación de suelos, acceso limitado a mercados, infraestructura deficiente) que requieren intervenciones coordinadas más allá del nivel provincial.

4 Discusión

4.1 Autocorrelación Espacial y Heterogeneidad Regional

Los resultados evidencian autocorrelación espacial significativa en los rendimientos agrícolas del sur peruano ($I = 0.404$, $p < 0.001$), con dependencia moderada

entre provincias vecinas ($\rho = 0.402$). Este patrón concuerda con aplicaciones de modelos CAR en agricultura [13], aunque la magnitud es ligeramente inferior a la reportada en Europa (0.35–0.55) [9], reflejando diferencias agroecológicas y de escala. La presencia de hotspots en Puno y coldspots en Cusco confirma la heterogeneidad subnacional identificada en estudios previos sobre los Andes [8]. No obstante, el tamaño muestral limitado ($n=20$) introduce incertidumbre, coherente con la inestabilidad señalada por [10] y la necesidad de mayor resolución espacial a nivel distrital [11].

4.2 Efectos Climáticos y Adaptación

Las sequías reducen el rendimiento en aproximadamente 183 kg/ha (IC: [-355, -3]), equivalente al 16% del rendimiento promedio, en línea con reducciones del 15–30% reportadas globalmente [3]. La ausencia de efectos significativos de heladas sugiere adaptación mediante variedades tolerantes y prácticas de manejo local [7], aunque no se descarta sesgo por variables binarias que no capturan intensidad ni temporalidad [4]. Los mayores rendimientos de papa frente a maíz (+780 kg/ha, 104% superior) reflejan ventajas adaptativas en zonas altoandinas, respaldadas por evidencia sobre resiliencia de tubérculos andinos [33, 34].

4.3 Implicaciones para Política

Las cuatro provincias críticas de Cusco (Canchis, Acomayo, Cusco, Canas) con rendimientos de 538–1,229 kg/ha, junto con La Unión en Arequipa (461 kg/ha), justifican intervenciones territorialmente focalizadas, dado el potencial de spillovers positivos hacia áreas vecinas [35]. La baja adopción de terrazas (1.6%) y semilla certificada (0.4%) refleja limitaciones estructurales más que ineffectividad, aunque estudios experimentales reportan incrementos del 20–40% con implementación adecuada [6].

4.4 Limitaciones

Cinco factores delimitan la interpretación: (i) diseño transversal impide inferencia causal robusta; (ii) tamaño muestral reducido ($n=20$ provincias) afecta precisión de estimaciones [10]; (iii) variables climáticas binarias subestiman intensidad y temporalidad de eventos; (iv) $R^2=0.184$ indica que factores no observados (suelo, gestión, crédito) explican la mayor parte de variabilidad [36]; y (v) la estructura de vecindad $k=5$ simplifica interacciones espaciales reales [9].

5 Conclusiones

Este estudio ofrece la primera evidencia cuantitativa de autocorrelación espacial en rendimientos agrícolas del sur del Perú mediante modelos jerárquicos bayesianos CAR. Se identificó dependencia espacial significativa ($I = 0.404$,

$\rho = 0.402$), con cuatro hotspots en Puno y cuatro coldspots en Cusco, revelando polarización territorial no capturada en estadísticas agregadas.

Las sequías reducen los rendimientos en 16% (-183 kg/ha), constituyendo el único evento climático con efecto robusto, mientras que heladas y granizadas no mostraron impactos significativos, sugiriendo adaptación local mediante variedades tolerantes. La papa superó al maíz en 780 kg/ha (104% más productiva), reafirmando su rol como cultivo resiliente altoandino. Cinco provincias requieren intervención urgente: cuatro en Cusco (Canchis, Acomayo, Cusco, Canas) con rendimientos de 538–1,229 kg/ha, y La Unión en Arequipa con el rendimiento más bajo registrado (461 kg/ha).

La baja adopción de prácticas mejoradas (terrazas: 1.6%; semilla certificada: 0.4%) refleja limitaciones estructurales más que ineffectividad, sugiriendo necesidad de subsidios e incentivos diferenciados espacialmente. Estos hallazgos aportan la primera estimación empírica del parámetro de autocorrelación espacial ($\rho = 0.402$) en agricultura peruana y ofrecen un marco operativo para priorización territorial basado en evidencia.

Se recomienda implementar: (i) programas piloto de adaptación focalizados en manejo de sequías en provincias críticas; (ii) redes de intercambio de conocimiento entre hotspots y coldspots mediante escuelas de campo; (iii) subsidios diferenciados espacialmente para adopción de terrazas y semilla certificada; y (iv) sistemas de alerta temprana climática a nivel provincial.

Futuras investigaciones deben emplear modelos espacio-temporales con paneles multianuales (2019-2024) e integrar datos satelitales (NDVI, precipitación CHIRPS) a nivel distrital ($n \approx 100$) para mejorar identificación causal y reducir incertidumbre paramétrica. En conjunto, los resultados evidencian que la vulnerabilidad agrícola tiene una estructura espacial pronunciada que exige políticas territorialmente diferenciadas y basadas en evidencia cuantitativa.

Bibliography

- [1] World Bank Group. Climate-smart agriculture in peru. Technical report, World Bank Group, 2019. URL <https://climateknowledgeportal.worldbank.org/sites/default/files/2019-06/CSA-in-Peru.pdf>. CSA Country Profile / Brief.
- [2] Ministerio de Agricultura y Riego del Perú (MINAGRI). Anuario estadístico de producción agrícola 2023. Technical report, Ministerio de Agricultura y Riego del Perú, 2023. URL <https://www.gob.pe/institucion/midagri/colecciones/5149-anuarios-estadisticas-de-produccion-agropecuaria>. Serie estadística agraria oficial.
- [3] David B. Lobell, Marianne Bänziger, Cosmos Magorokosho, and Bindiganavile Vivek. Nonlinear heat effects on african maize as evidenced by historical yield trials. *Nature Climate Change*, 1:42–45, 2011. doi: 10.1038/nclimate1043.
- [4] Melissa Dell, Benjamin F. Jones, and Benjamin A. Olken. What do we learn from the weather? the new climate-economy literature. *Journal of Economic Literature*, 52(3):740–798, 2014. doi: 10.1257/jel.52.3.740.
- [5] Bryan G. Mark, Jeffrey Bury, Aaron McKenzie, Adam French, and Mark Baraer. Glacier loss and hydro-social risks in the peruvian andes. *Global and Planetary Change*, 159:61–76, 2017. doi: 10.1016/j.gloplacha.2017.10.003. URL <https://doi.org/10.1016/j.gloplacha.2017.10.003>.
- [6] Helena Posthumus et al. Cost-benefit analysis of terraces in peru. *Land Degradation & Development*, 21:427–439, 2010. doi: 10.1002/ldr.976.
- [7] Karl S. Zimmerer. Biological diversity in agriculture and global change. *Annual Review of Environment and Resources*, 35:137–166, 2010. doi: 10.1146/annurev-environ-040309-113840.
- [8] Diana Sietz et al. Landuse-based adaptation to global change. *Environmental Management*, 50:87–99, 2012. doi: 10.1007/s00267-012-9950-9.
- [9] Melanie M. Wall. A close look at the spatial structure implied by the car and sar models. *Journal of Statistical Planning and Inference*, 121:311–324, 2004. doi: 10.1016/S0378-3758(03)00111-3.
- [10] James S. Hodges and Brian J. Reich. Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician*, 64:325–334, 2010. doi: 10.1198/tast.2010.10052.
- [11] Marta Blangiardo and Michela Cameletti. Spatial and spatio-temporal models with r-inla. *Spatial and Spatio-temporal Epidemiology*, 4:33–49, 2013. doi: 10.1016/j.sste.2012.12.001.
- [12] Andrea Riebler, Sigrunn H. Sørbye, Daniel Simpson, and Håvard Rue. An intuitive bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4):1145–1165, 2016. doi: 10.1177/0962280216660421.

- [13] Bin Zhang et al. Spatial analysis of climate change impacts on yield. *Journal of Geographical Sciences*, 28:1547–1567, 2018. doi: 10.1007/s11442-018-1558-9.
- [14] Arthur Getis and J. K. Ord. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24(3):189–206, 1992. doi: 10.1111/j.1538-4632.1992.tb00261.x. URL <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>.
- [15] Anna Heikkinen. Climate change in the peruvian andes: A case study on small-scale farmers’ vulnerability in the quillcay river basin. *Iberoamericana – Nordic Journal of Latin American and Caribbean Studies*, 50(1):1–15, 2021. doi: 10.16993/iberoamericana.211. URL <https://doi.org/10.16993/iberoamericana.211>.
- [16] Instituto Nacional de Estadística e Informática. *Encuesta Nacional Agropecuaria 2024*. INEI, Lima, Perú, 2024. URL <https://www.inei.gob.pe/>. Datos de producción agrícola y prácticas de manejo.
- [17] P. A. P. Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2):17–23, 1950. doi: 10.2307/2332142.
- [18] Roger Bivand and Gianfranco Piras. Comparing implementations of estimation methods for spatial econometrics. *Journal of Statistical Software*, 63(18):1–36, 2015. doi: 10.18637/jss.v063.i18. Paquete spdep.
- [19] Yufeng Wang, Xue Chen, and Feng Xue. A review of bayesian spatiotemporal models in spatial epidemiology. *ISPRS International Journal of Geo-Information*, 13(3):97, 2024. doi: 10.3390/ijgi13030097. URL <https://www.mdpi.com/2220-9964/13/3/97>.
- [20] Paul-Christian Bürkner. brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80(1):1–28, 2017. doi: 10.18637/jss.v080.i01.
- [21] Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1):1–32, 2017. doi: 10.18637/jss.v076.i01.
- [22] Brian G. Leroux, Xingye Lei, and Norman Breslow. Estimation of disease rates in small areas: A new mixed model for spatial dependence. In M. Elizabeth Halloran and Donald Berry, editors, *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, volume 116 of *IMA Volumes in Mathematics and its Applications*, pages 179–191. Springer New York, New York, NY, 2000. ISBN 978-1-4612-1284-3. doi: 10.1007/978-1-4612-1284-3_4. URL https://doi.org/10.1007/978-1-4612-1284-3_4.
- [23] Duncan Lee. Carbayes: An r package for bayesian spatial modeling with conditional autoregressive priors. *Journal of Statistical Software*, 55(13): 1–24, 2013. doi: 10.18637/jss.v055.i13.
- [24] Andrew Gelman and Donald B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472, 1992. doi: 10.1214/ss/1177011136.

- [25] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5):1413–1432, 2017. doi: 10.1007/s11222-016-9696-4.
- [26] Andrew D. Cliff and J. Keith Ord. Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, 4(3):267–284, 1972. doi: 10.1111/j.1538-4632.1972.tb00475.x.
- [27] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023. URL <https://www.R-project.org/>.
- [28] Hadley Wickham, Evan Miller, and Danny Smith. *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*, 2023. URL <https://CRAN.R-project.org/package=haven>. R package version 2.5.3.
- [29] Hadley Wickham, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. *dplyr: A Grammar of Data Manipulation*, 2023. URL <https://CRAN.R-project.org/package=dplyr>. R package version 1.1.2.
- [30] Edzer Pebesma. Simple features for r: Standardized support for spatial vector data. *The R Journal*, 10(1):439–446, 2018. doi: 10.32614/RJ-2018-009. Package sf.
- [31] Jonah Gabry, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. Visualization in bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2):389–402, 2019. doi: 10.1111/rssa.12378.
- [32] Joe Cheng, Bhaskar Karambelkar, and Yihui Xie. *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library*, 2023. URL <https://CRAN.R-project.org/package=leaflet>. R package version 2.2.0.
- [33] André Devaux et al. Global food security from sustainable potato systems. *The Potato Crop*, pages 3–35, 2021. doi: 10.1007/978-3-030-28683-5_1.
- [34] Leslie Lipper et al. Climate-smart agriculture for food security. *Nature Climate Change*, 4:1068–1072, 2014. doi: 10.1038/nclimate2437.
- [35] Timothy G. Conley and Christopher R. Udry. Learning about a new technology. *American Economic Review*, 100:35–69, 2010. doi: 10.1257/aer.100.1.35.
- [36] Marshall Burke, Solomon M. Hsiang, and Edward Miguel. Global non-linear effect of temperature on economic production. *Nature*, 527:235–239, 2015. doi: 10.1038/nature15725.

6 Anexos

6.1 Mapas Complementarios

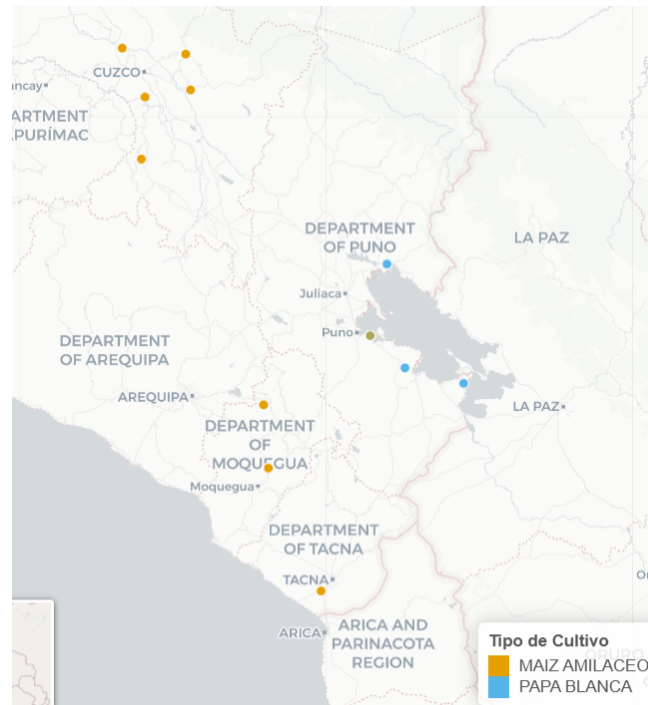


Fig. 10. Distribución espacial de cultivos en la región de estudio.

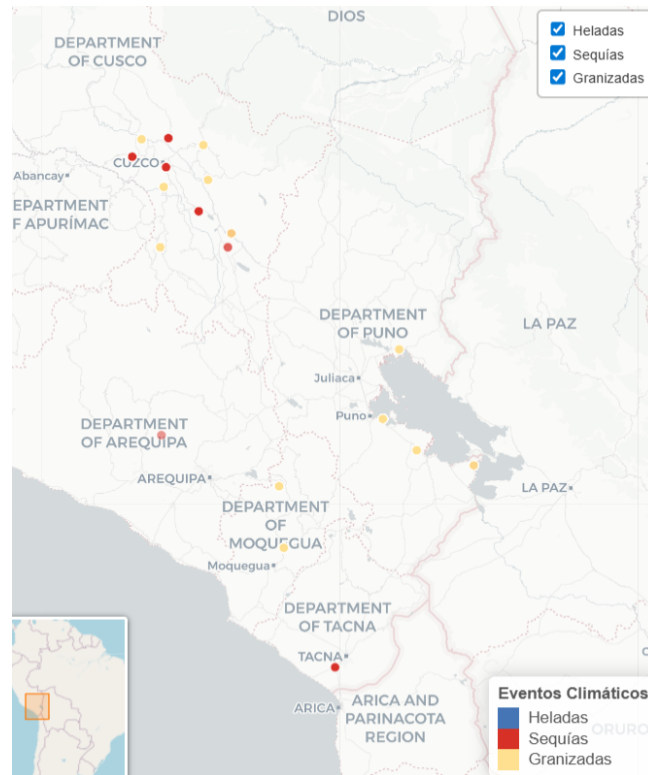


Fig. 11. Distribución espacial de eventos climáticos a nivel de parcelas.

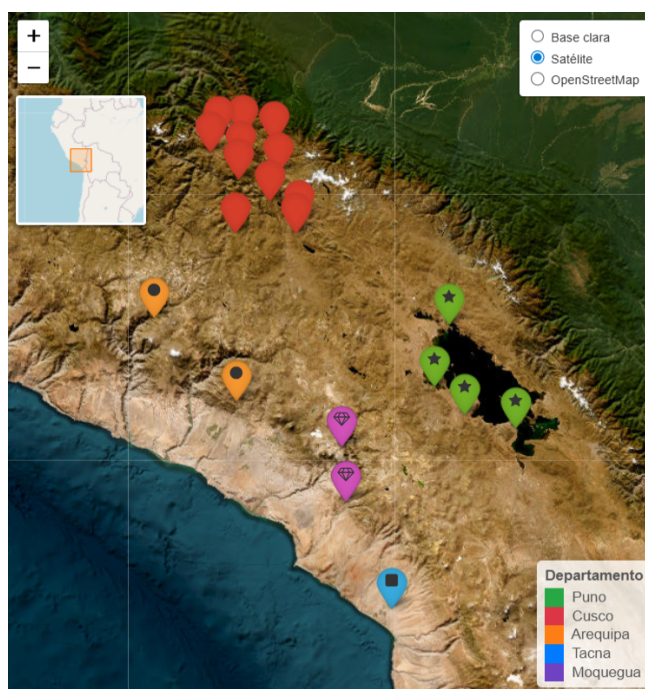


Fig. 12. Vista integrada de datos geospaciales por departamento.