

# Modelado de Patrones Agrícolas Utilizando Procesos de Cox Espacio-Temporales con Datos de la ENA 2024

Yhack Bryan Aycaya Paco<sup>1,†</sup>

<sup>1</sup>Facultad de Estadística e Informática, Universidad Nacional del Altiplano, Puno, Perú

<sup>†</sup>Autor Principal

## Resumen

Los procesos de Cox espacio-temporales son herramientas estadísticas avanzadas para modelar patrones de puntos con intensidades variables en el espacio y el tiempo. Este estudio utiliza datos de la Encuesta Nacional Agrícola (ENA) 2024 para analizar eventos agrícolas, como la ubicación de parcelas y actividades de producción, en el contexto peruano. Integrando variables espaciales (latitud, longitud, región) y temporales (meses de siembra y cosecha), junto con covariables como el tamaño de la parcela y el tipo de cultivo, se propone un proceso de Cox log-Gaussiano (LGCP) localizado para modelar la agrupación de eventos agrícolas. La metodología aborda desafíos como coordenadas a nivel de segmento y granularidad temporal mensual, incorporando heterogeneidad ambiental para predecir puntos críticos de producción y riesgos hidrológicos. Este enfoque contribuye a la planificación agrícola sostenible en Perú, apoyando la toma de decisiones en un contexto de cambio climático y riesgos naturales.

**Keywords:** Procesos de puntos espacio-temporales, Procesos de Cox log-Gaussianos, Modelado de datos agrícolas, ENA 2024, Estadística espacial

**Autor de correspondencia:** Yhack Bryan Aycaya Paco, Facultad de Estadística e Informática, Universidad Nacional del Altiplano, Puno 21001, Perú. *Correo electrónico:* yaycaya@est.unap.edu.pe

## 1. Introducción y Motivación

Este trabajo se centra en el modelado de patrones agrícolas en Perú utilizando procesos de Cox espacio-temporales, con énfasis en la aplicación de procesos log-Gaussianos (LGCP) a datos de la Encuesta Nacional Agrícola (ENA) 2024. Este tema es relevante debido a su impacto social en la seguridad alimentaria, su relevancia académica en estadística espacial y su aplicación profesional en la planificación agrícola sostenible. La agricultura peruana enfrenta desafíos significativos, como la variabilidad climática, la intensificación del uso de la tierra y los riesgos hidrológicos como inundaciones [1, 2], que afectan regiones clave como la Costa, la Sierra y la Selva. La ENA 2024 proporciona datos detallados sobre coordenadas espaciales (LATITUD, LONGITUD), fechas de actividades agrícolas (P124\_MES, P124\_AÑO) y covariables como el tamaño de la parcela (P105\_SUP\_ha), lo que permite un análisis espacio-temporal robusto [3, 4].

El interés en este tema surge de la necesidad de abordar lagunas en la literatura respecto a la aplicación de modelos espacio-temporales avanzados a datos agrícolas, un área menos explorada en comparación con disciplinas como la sismología [4], la epidemiología [5, 6], el análisis de crímenes [7], la dinámica de incendios [8] y la predicción de eventos urbanos [9]. La motivación también proviene de la experiencia adquirida en cursos de estadística espacial, donde se identificó el potencial de los procesos de Cox para capturar patrones complejos en contextos agrícolas. Este estudio busca analizar cómo las variables espaciales, temporales y ambientales de la ENA 2024 pueden modelarse para predecir puntos críticos de producción y riesgos, contribuyendo al entendimiento de la dinámica agrícola en un contexto de cambio climático y creciente presión sobre los recursos hídricos [2].

## 2. Planteamiento del Problema

El fenómeno de los patrones agrícolas ha sido abordado principalmente desde perspectivas agronómicas y de modelado hidrológico [1, 2]; sin embargo, existen limitaciones significativas en la aplicación de procesos de puntos espacio-temporales a datos agrícolas. Los datos de la ENA 2024, aunque ricos en información, presentan desafíos como coordenadas a nivel de segmento (no individuales por parcela) y una granularidad temporal mensual, lo que introduce incertidumbre en la estimación de la intensidad [10]. Además, la heterogeneidad ambiental y antropogénica, como la variabilidad climática y las prácticas agrícolas, requiere modelos que capturen interacciones no separables [8]. Estudios previos en sismología [4], epidemiología [5] y análisis de crímenes [7] han utilizado con éxito procesos LGCP y Hawkes, pero su adaptación al contexto agrícola peruano es limitada, particularmente en la integración de covariables como el tamaño de la parcela (P105\_SUP\_ha) y la región natural (REGIÓN).

Esta brecha de conocimiento justifica la necesidad de un estudio que explore cómo los procesos de Cox espacio-temporales pueden modelar patrones de puntos agrícolas, considerando la heterogeneidad espacial (por ejemplo, regiones de Costa, Sierra y Selva) y temporal (ciclos de siembra y cosecha). El objetivo es desarrollar un marco que integre los datos de la ENA 2024 para predecir puntos críticos de producción y riesgos hidrológicos, apoyando la planificación sostenible en un contexto de creciente vulnerabilidad climática [1, 2].

## 3. Objetivos y Preguntas de Investigación

### 3.1. Objetivo General

Desarrollar un marco de modelado basado en procesos de Cox espacio-temporales para analizar patrones de puntos agrícolas en los datos de la ENA 2024, integrando covariables espaciales, temporales y ambientales para predecir puntos críticos de producción y riesgos en el contexto agrícola peruano.

### 3.2. Objetivos Específicos

- Identificar y analizar variables clave de la ENA 2024 (por ejemplo, LATITUD, LONGITUD, P124\_MES, P105\_SUP\_ha) para modelar la intensidad de eventos agrícolas, considerando la heterogeneidad espacial y temporal.
- Proponer un modelo LGCP localizado y extensiones de procesos de Hawkes para capturar patrones de agrupación no separables, adaptando enfoques de [3, 11].
- Evaluar el impacto de covariables como el tamaño de la parcela y la región natural en la predicción de puntos críticos, utilizando métodos de inferencia variacional [12] y transformadas de Fourier [13].
- Generar pronósticos de flujos de producción y riesgos hidrológicos, integrando modelos híbridos como ARNO/-VIC+RAPID [1] y considerando datos externos como imágenes satelitales [14].

### 3.3. Preguntas de Investigación

- ¿Cómo influyen las covariables espaciales (LATITUD, LONGITUD, REGIÓN) y temporales (P124\_MES, P124\_AÑO) de la ENA 2024 en la intensidad de los procesos de Cox, considerando limitaciones como coordenadas segmentadas?
- ¿Pueden los modelos LGCP localizados [3] mejorar la detección de puntos críticos agrícolas en comparación con enfoques globales en regiones heterogéneas como la Selva?
- ¿Qué papel juegan los factores ambientales, como se describe en [2, 6], en la predicción de riesgos hidrológicos y derrames virales en las regiones agrícolas peruanas?
- ¿Cómo se pueden integrar diversas fuentes de datos, como big data de redes sociales [15] o imágenes satelitales [14], para mejorar las estimaciones en datos agrícolas agregados [10]?

## 4. Metodología Propuesta

La investigación se desarrollará bajo un enfoque cuantitativo, utilizando procesos de Cox espacio-temporales, específicamente modelos LGCP localizados, para modelar patrones de puntos agrícolas

### 4.1. Descripción de la Metodología

- **Procesamiento de Datos:** Los datos de la ENA 2024 se organizarán por variables espaciales (LATITUD, LONGITUD, REGIÓN), variables temporales (P124\_MES, P124\_AÑO, P125\_MES) y covariables (P105\_SUP\_ha, P115\_COD). La granularidad mensual se corregirá añadiendo ruido controlado para simular precisión diaria, inspirado en [10].
- **Modelado de la Intensidad:** Se estimará la intensidad separable con un componente temporal utilizando regresión de Poisson con covariables calendáricas (por ejemplo, día de la semana, estación) [16] y un componente espacial utilizando estimación de kernel con selección de ancho de banda mediante K-means [16]. Para modelos no separables, se utilizarán enfoques de aprendizaje profundo probabilístico [9, 17] y mezclas profundas. La metodología se aplicará a los datos de la ENA 2024, integrando referencias de hidrología [1, 2], sismología [4],

epidemiología [5, 6], incendios [8], crímenes [7], accidentes [18] y patrones de embarcaciones [14], para un análisis integral. También se considerarán enfoques de aprendizaje profundo para modelar estructuras complejas

## 5. Impacto y Proyección

Se espera que este trabajo contribuya al avance académico al llenar la brecha en la aplicación de procesos de Cox espacio-temporales a datos agrícolas, fomentando la discusión sobre el uso de LGCP localizados en contextos heterogéneos [3]. Prácticamente, el modelo permitirá identificar puntos críticos de producción y riesgos hidrológicos, apoyando intervenciones dirigidas en regiones vulnerables como la Costa y la transición Andes-Amazonas [1]. En el marco del curso, esta investigación fortalecerá competencias en estadística espacial y modelado de procesos de puntos, preparando al autor para futuras investigaciones en agricultura de precisión y gestión de recursos hídricos. A largo plazo, se proyecta extender el modelo a pronósticos en tiempo real, integrando datos satelitales

## ■ Referencias

- [1] Llauca H, Leon K, Lavado-Casimiro W. Construction of a daily streamflow dataset for Peru using a similarity-based regionalization approach and a hybrid hydrological modeling framework. *Journal of Hydrology: Regional Studies*. 2023;47:101381. Available from: <https://www.sciencedirect.com/science/article/pii/S221458182300068X>.
- [2] Wei X, Garcia-Chevesich P, Alejo F, García V, Martínez G, Daneshvar F, et al. Hydrologic Analysis of an Intensively Irrigated Area in Southern Peru Using a Crop-Field Scale Framework. *Water*. 2021;13(3):318. Available from: <https://www.mdpi.com/2073-4441/13/3/318>.
- [3] D'Angelo N, Adelfio G, Mateu J. Locally weighted minimum contrast estimation for spatio-temporal log-Gaussian Cox processes. *Computational Statistics Data Analysis*. 2023;180:107679. Available from: <https://www.sciencedirect.com/science/article/pii/S0167947322002596>.
- [4] D'Angelo N, Siino M, D'Alessandro A, Adelfio G. Local spatial log-Gaussian Cox processes for seismic data. *AStA Advances in Statistical Analysis*. 2022 Dec;106(4):633-71. Available from: <https://link.springer.com/article/10.1007/s10182-022-00444-w>.
- [5] Huang CC, Trevisi L, Becerra MC, Calderón RI, Contreras CC, Jimenez J, et al. Spatial scale of tuberculosis transmission in Lima, Peru. *Proceedings of the National Academy of Sciences*. 2022;119(45):e2207022119. Available from: <https://www.pnas.org/doi/10.1073/pnas.2207022119>.
- [6] Rita R, Matthiopoulou J, Lindgren F, Tello C, Zariquiey CM, Valderrama W, et al. Incorporating environmental heterogeneity and observation effort to predict host distribution and viral spillover from a bat reservoir. *Proceedings of the Royal Society B*. 2023;290:20231739. Available from: <https://royalsocietypublishing.org/doi/10.1098/rspb.2023.1739>.

- [7] Escudero I, Angulo JM, Mateu J, Choiruddin A. Crime risk assessment through Cox and self-exciting spatio-temporal point processes. *Stochastic Environmental Research and Risk Assessment*. 2025 Jan;39(1):181-203. Available from: <https://link.springer.com/article/10.1007/s00477-024-02857-2>.
- [8] D'Angelo N, Albano A, Gilardi A, Adelfio G. Non-separable spatio-temporal Poisson point process models for fire occurrences. *Environmental and Ecological Statistics*. 2025 Mar;32(1):347-81. Available from: <https://link.springer.com/article/10.1007/s10651-025-00645-x>.
- [9] Okawa M, Iwata T, Kurashima T, Tanaka Y, Toda H, Ueda N. Deep Mixture Point Processes: Spatio-temporal Event Prediction with Rich Contextual Information. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; 2019. p. 373-83. Available from: <https://dl.acm.org/doi/10.1145/3292500.3330937>.
- [10] Asfaw ZG, Brown PE, Stafford J. The root-Gaussian Cox Process for spatial-temporal disease mapping with aggregated data. *Computational Statistics*. 2025 Mar;40(3):1171-84. Available from: <https://link.springer.com/article/10.1007/s00180-024-01532-y>.
- [11] Clark NJ, Watts K. Identification of latent structure in spatio-temporal models of violence. In: *Proceedings of the Winter Simulation Conference. WSC '21*. IEEE Press; 2022. Available from: <https://dl.acm.org/doi/10.5555/3522802.3522945>.
- [12] Aglietti V, Bonilla EV, Damoulas T, Cripps S. Structured variational inference in continuous cox process models. 2019. Available from: <https://papers.nips.cc/paper/2019/hash/5c1e1c7f6a63e2e56dd5d688e8d7d3b2-Abstract.html>.
- [13] Rajala TA, Olhede SC, Grainger JP, Murrell DJ. What is the Fourier Transform of a Spatial Point Process? *IEEE Transactions on Information Theory*. 2023;69(8):5219-52. Available from: <https://ieeexplore.ieee.org/document/10107467>.
- [14] Nakayama S, Dong W, Corroero RG, Selig ER, Wabnitz CCC, Hastie TJ, et al. Comparing spatial patterns of marine vessels between vessel-tracking data and satellite imagery. *Frontiers in Marine Science*. 2023;9. Available from: <https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2022.1076775>.
- [15] Cheng Z, Wang J, Yen N, Wu Y. Allocation of Resources after Disaster Based on Big Data from SNS and Spatial Scan. In: *2016 International Conference on Advanced Cloud and Big Data (CBD)*; 2016. p. 334-4. Available from: <https://ieeexplore.ieee.org/document/7815232>.
- [16] Bayisa FL, Ådahl M, Rydén P, Cronie O. Large-scale modelling and forecasting of ambulance calls in northern Sweden using spatio-temporal log-Gaussian Cox processes. *Spatial Statistics*. 2020;39:100471. Available from: <https://www.sciencedirect.com/science/article/pii/S2211675320300658>.
- [17] Choiruddin A, Rino Fajar Sakti E, Dwi Ary Widhianningsih T, Mateu J, Fithriasari K. Probabilistic Deep Learning for Highly Multivariate Spatio-Temporal Log-Gaussian Cox Processes. *IEEE Access*. 2025;13:94761-76. Available from: <https://ieeexplore.ieee.org/document/11005449>.
- [18] Spychala C, Dombry C, Goga C. Variable selection methods for Log-Gaussian Cox processes: A case-study on accident data. *Spatial Statistics*. 2024;61:100831. Available from: <https://www.sciencedirect.com/science/article/pii/S2211675324000228>.