

# Inteligencia Artificial para Agricultura Sostenible: Modelos SARIMAX en Optimización de Producción de Papa Altoandina

Lindell Dennis Vilca Mamani<sup>1</sup>0009-0000-0583-3237

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

**Resumen** La agricultura altoandina enfrenta desafíos de baja productividad y adopción tecnológica limitada en contextos de cambio climático. Este estudio aplicó modelos SARIMAX para evaluar el impacto de indicadores de sostenibilidad en la capacidad predictiva de producción de papa en Puno, Perú (2015-2024,  $n=17,679$  productores). Se construyó un índice de tecnificación basado en rendimiento como proxy de adopción de prácticas sostenibles, clasificando productores en tres niveles tecnológicos. Los resultados muestran que el 18.9 % opera con alta tecnología, con correlación perfecta ( $r=1.0$ ,  $p<0.05$ ) entre nivel tecnológico y rendimiento. El modelo SARIMAX incorporando variables de sostenibilidad presentó componente estacional anual  $ARIMA(0,0,1)(0,0,1)_{12}$ , mejorando la interpretabilidad respecto al modelo económico baseline. La simulación de escenarios proyecta un incremento del 3.4 % en producción con mayor adopción tecnológica (+30 %), demostrando potencial de optimización de recursos. Estos hallazgos evidencian que la inteligencia artificial aplicada mediante modelos SARIMAX facilita la toma de decisiones para agricultura de precisión en zonas altoandinas, contribuyendo a seguridad alimentaria y economía verde mediante uso eficiente de tierra y recursos productivos.

**Keywords:** SARIMAX, agricultura sostenible, inteligencia artificial, optimización de recursos, economía verde, agricultura de precisión

## 1. Introducción

La agricultura en zonas altoandinas constituye un sistema productivo fundamental para la seguridad alimentaria de millones de personas en la región andina de Sudamérica. La papa (*Solanum tuberosum*), cultivo originario de los Andes peruanos, representa no solo un componente central de la dieta local, sino también una fuente crítica de ingresos para pequeños productores en altitudes que oscilan entre 3,800 y 4,500 metros sobre el nivel del mar [10]. Sin embargo, esta agricultura enfrenta desafíos crecientes derivados del cambio climático, incluyendo variabilidad de precipitaciones, incremento de temperaturas y mayor frecuencia de eventos climáticos extremos como heladas y sequías [5, 6].

La región de Puno, ubicada en el altiplano peruano, concentra una importante producción de papa con más de 17,000 productores registrados anualmente. No obstante, la adopción de tecnologías agrícolas mejoradas permanece limitada. Estudios previos han documentado que agricultores andinos han debido desplazar sus cultivos hasta 150 metros de altitud en las últimas tres décadas como estrategia de adaptación climática, evidenciando la urgente necesidad de herramientas predictivas que faciliten la planificación agrícola y optimización de recursos en este contexto de alta vulnerabilidad.

A pesar del creciente interés en aplicaciones de inteligencia artificial para agricultura [12, 15, 11], existe una brecha significativa en la investigación sobre modelos de series temporales que incorporen explícitamente variables de sostenibilidad para regiones altoandinas. Los modelos SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous variables) han demostrado efectividad en pronósticos agrícolas en contextos similares [13, 9, 8], pero su aplicación con indicadores de tecnología verde en agricultura de pequeña escala andina permanece inexplorada. Esta limitación es particularmente relevante dado que las bases de datos agrícolas en países en desarrollo frecuentemente carecen de información directa sobre adopción tecnológica.

Este trabajo presenta una contribución metodológica al desarrollar un índice de tecnificación basado en rendimiento como proxy de adopción de prácticas sostenibles, validando su aplicación en modelos SARIMAX para pronóstico de producción. Específicamente, los objetivos son: (1) caracterizar la evolución de niveles tecnológicos en productores de papa de Puno durante 2015-2024, (2) comparar la capacidad predictiva de modelos SARIMAX con y sin variables de sostenibilidad, y (3) simular escenarios de optimización de recursos mediante incrementos en adopción tecnológica. Los resultados contribuyen al cuerpo de conocimiento sobre aplicación de IA para agricultura de precisión en contextos de recursos limitados, con implicaciones directas para políticas de extensión agrícola y economía verde en zonas altoandinas.

## 2. Trabajos Relacionados

### 2.1. Modelos SARIMAX en Agricultura

Los modelos SARIMAX han ganado prominencia en años recientes para pronóstico de producción agrícola debido a su capacidad de incorporar variables exógenas climáticas y económicas. Pandit et al. desarrollaron modelos híbridos ARIMAX-LSTM para pronóstico de rendimiento de cultivos Rabi en India (trigo, caña de azúcar, maní), utilizando área bajo riego como variable exógena y logrando mejoras significativas en precisión predictiva respecto a modelos univariados. En contextos de alta variabilidad climática, Lemos y Bezerra aplicaron ARIMAX para producción de granos en la región semiárida brasileña, demostrando que la incorporación de precipitación como variable exógena captura adecuadamente la volatilidad inducida por sequías recurrentes.

Krishna Priya y Kausalya extendieron este enfoque al pronóstico simultáneo de área, producción y productividad de trigo en India usando temperatura y

precipitación, mientras que Singh et al. aplicaron ARIMAX para producción de ricino utilizando precipitación como variable exógena [16]. Prabhu y Suma [14] compararon SARIMAX con redes neuronales artificiales para pronóstico de sequías en Karnataka, concluyendo que SARIMAX mantiene ventajas interpretativas para planificación agrícola a pesar de precisión comparable. Más recientemente, Ghosh et al. [4] incorporaron modelos GARCH-ARIMAX para capturar volatilidad realizada de precipitación en producción de arroz, mientras que Alam et al. [1] desarrollaron modelos híbridos ARIMAX-ANN y ARIMAX-SVM para rendimiento de arroz en Uttar Pradesh.

Estos estudios demuestran la robustez de aproximaciones SARIMAX en diversos contextos agrícolas, pero se enfocan predominantemente en variables climáticas y económicas tradicionales. Ninguno ha explorado la incorporación de indicadores de sostenibilidad o tecnología verde, ni su aplicación en agricultura de pequeña escala en zonas altoandinas.

## 2.2. Producción de Papa en los Andes

La literatura sobre producción de papa en regiones andinas se ha centrado principalmente en impactos de cambio climático y percepciones de agricultores. Lozano-Povis et al. realizaron una revisión sistemática del impacto climático en agricultura andina, identificando que la papa es particularmente vulnerable a incrementos de temperatura y alteraciones en patrones de precipitación, con proyecciones de reducción de áreas aptas para cultivo en altitudes intermedias. Hijmans et al. documentaron percepciones de pequeños productores en Amazonas, Perú, revelando que el 78 % reporta disminución de rendimientos atribuida a cambio climático, pero con adopción limitada de tecnologías de adaptación.

A pesar de la importancia económica de la papa andina, existe una escasez notable de estudios cuantitativos que apliquen técnicas de aprendizaje automático o modelos de series temporales para pronóstico de producción en estas regiones. Esta brecha es particularmente significativa considerando la disponibilidad de datos de encuestas agrícolas nacionales que permanecen subutilizados para análisis predictivos.

## 2.3. Tecnologías Verdes e IA en Agricultura Sostenible

La convergencia de tecnologías verdes e inteligencia artificial representa una frontera emergente para agricultura sostenible. Gautam [3] documentó experiencias tempranas de adopción de tecnologías verdes (solar, eólica, biogás) en países en desarrollo, destacando barreras de costo y conocimiento para pequeños agricultores. Más recientemente, la proliferación de aplicaciones de IA ha abierto nuevas posibilidades para optimización de recursos agrícolas.

Nafchi et al. proporcionan una revisión comprensiva de aplicaciones de IA, Machine Learning y Deep Learning en agricultura de precisión, cubriendo predicción de condiciones del suelo, detección de enfermedades y tecnología de aplicación variable. Singh y Kumar revisaron específicamente métodos de predicción de rendimiento, comparando Random Forest, ANN, SVM, CNN, LSTM y DNN,

concluyendo que modelos híbridos que combinan series temporales con aprendizaje automático ofrecen el mejor balance entre interpretabilidad y precisión.

Mana et al. enfatizan que la sostenibilidad de sistemas basados en IA requiere consideración explícita de optimización de recursos hídricos, energéticos y de insumos. Gamage et al. [2] argumentan que tecnologías emergentes, incluyendo IA, tienen potencial transformador para agricultura sostenible, pero su adopción en países en desarrollo enfrenta desafíos de infraestructura y capacitación. Karim et al. [7] proponen arquitecturas IoT habilitadas por IA para agricultura inteligente, integrando visión computacional, análisis predictivo y sistemas de soporte a decisiones.

Sin embargo, la mayoría de estos estudios se centran en agricultura intensiva de gran escala o contextos con alta disponibilidad de datos de sensores. La aplicabilidad de enfoques de IA en agricultura de pequeña escala con datos administrativos limitados permanece insuficientemente explorada, particularmente en regiones altoandinas donde la adopción de tecnologías digitales es incipiente.

### 3. Metodología

#### 3.1. Fuente de Datos y Descripción

Este estudio utilizó datos de la Encuesta Nacional Agropecuaria (ENA) de Perú para el período 2015-2024, con ausencia de datos para 2020 debido a restricciones de recolección durante la pandemia COVID-19. La base de datos comprende 17,679 observaciones individuales de productores de papa en la región de Puno, cubriendo 13 provincias y múltiples distritos a lo largo del altiplano peruano.

Las variables principales incluyen: producción en kilogramos (P219\_EQUIV\_KG), superficie cosechada en hectáreas (P217\_SUP\_ha), precio unitario de venta (precio\_unitario), mes de siembra y cosecha, tipo de papa producida (P204\_NOM), y ubicación geográfica (provincia y distrito). La ENA es una encuesta probabilística representativa a nivel regional implementada por el Instituto Nacional de Estadística e Informática (INEI) de Perú, con factores de expansión que permiten inferencias poblacionales.

Una limitación importante es la ausencia de información directa sobre adopción de tecnologías específicas (semilla certificada, riego tecnificado) en la mayoría de observaciones. Para abordar esta limitación, se desarrolló una metodología de construcción de índice de tecnificación basada en rendimiento, descrita en la siguiente subsección.

#### 3.2. Construcción del Índice de Tecnificación

Dado que las variables binarias de tecnología (semilla certificada, riego tecnificado) presentaban valores nulos o cero en la mayoría de observaciones, se construyó un índice de tecnificación como proxy de adopción de prácticas mejoradas. Este enfoque se justifica por la literatura agroeconómica que establece

que el rendimiento (kg/ha) refleja indirectamente la calidad de manejo, uso de insumos mejorados, y prácticas agronómicas [15].

El índice de tecnificación se calculó mediante normalización min-max del rendimiento individual:

$$IT_i = \frac{R_i - R_{min}}{R_{max} - R_{min}} \times 100 \quad (1)$$

donde  $IT_i$  es el índice de tecnificación del productor  $i$ ,  $R_i$  es su rendimiento en kg/ha, y  $R_{min}$ ,  $R_{max}$  son los valores mínimo y máximo de rendimiento en la muestra. Este índice oscila entre 0-100, donde valores altos indican mayor eficiencia productiva.

Adicionalmente, los productores se clasificaron en tres categorías tecnológicas basadas en cuartiles de rendimiento:

- **Alta tecnología:** Rendimiento  $> P_{75}$  (cuartil superior)
- **Media tecnología:**  $P_{50} < \text{Rendimiento} \leq P_{75}$  (cuartiles 2-3)
- **Baja tecnología:** Rendimiento  $\leq P_{50}$  (cuartil inferior)

Esta clasificación permite analizar la distribución de niveles tecnológicos en la población y su evolución temporal.

### 3.3. Agregación Temporal y Variables de Modelo

Los datos individuales se agregaron mensualmente para construir series temporales compatibles con modelos SARIMAX. La agregación se realizó mediante suma de producción total, promedio ponderado de precios, y cálculo de proporciones para variables tecnológicas:

- **Producción total:**  $Y_t = \sum_{i \in t} P_{i,t}$  (toneladas)
- **Precio promedio:**  $\bar{p}_t = \frac{\sum_{i \in t} p_i \cdot P_i}{\sum_{i \in t} P_i}$
- **Superficie total:**  $S_t = \sum_{i \in t} S_{i,t}$  (hectáreas)
- **Índice tecnológico agregado:**  $\overline{IT}_t = \frac{1}{n_t} \sum_{i \in t} IT_i$
- **Proporción alta tecnología:**  $PropAlta_t = \frac{\#\{i \in t: Cat_i = Alta\}}{n_t} \times 100$

Para meses con observaciones faltantes (debido a estacionalidad de siembra/cosecha), se aplicó interpolación lineal mediante la función `na.approx` del paquete `zoo` en R, asegurando continuidad temporal necesaria para modelos ARIMA.

### 3.4. Especificación de Modelos SARIMAX

Se estimaron dos modelos SARIMAX para comparar capacidad predictiva:

**Modelo 1: SARIMAX Económico (Baseline)** El modelo baseline incorpora únicamente variables económicas tradicionales:

$$Y_t = f(Y_{t-1}, \dots, Y_{t-p}, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}) + \beta_1 P_t + \beta_2 S_t + \varepsilon_t \quad (2)$$

donde  $Y_t$  es la producción en el tiempo  $t$ ,  $P_t$  es el precio promedio,  $S_t$  es la superficie cultivada, y  $\varepsilon_t$  es el término de error. La estructura ARIMA  $(p, d, q)(P, D, Q)_s$  se determina mediante selección automática por criterio AIC usando la función `auto.arima` del paquete `forecast` en R.

**Modelo 2: SARIMAX con Tecnología Verde** El modelo extendido incorpora variables de sostenibilidad:

$$Y_t = f(Y_{t-1}, \dots) + \beta_1 P_t + \beta_2 S_t + \beta_3 IT_t + \beta_4 PropAlta_t + \varepsilon_t \quad (3)$$

donde  $IT_t$  es el índice de tecnificación agregado y  $PropAlta_t$  es la proporción de productores con alta tecnología. Este modelo permite evaluar si la inclusión de indicadores de sostenibilidad mejora la capacidad predictiva más allá de variables económicas tradicionales.

### 3.5. Estrategia de Evaluación

Los datos se dividieron en conjunto de entrenamiento (80 %, aproximadamente 80 observaciones mensuales) y conjunto de prueba (20 %, aproximadamente 20 observaciones). Los modelos se evaluaron mediante:

- **RMSE** (Root Mean Squared Error):  $\sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}$
- **MAE** (Mean Absolute Error):  $\frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$
- **MAPE** (Mean Absolute Percentage Error):  $\frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$
- **AIC/BIC**: Criterios de información para parsimonia de modelo

Se aplicó el test de Diebold-Mariano [13] para evaluar diferencias estadísticamente significativas en precisión predictiva entre modelos. Adicionalmente, se realizaron diagnósticos de residuos mediante tests de Ljung-Box y análisis de ACF/PACF para verificar que los residuos no presenten autocorrelación significativa.

### 3.6. Simulación de Escenarios de Sostenibilidad

Para evaluar el impacto potencial de mayor adopción tecnológica, se simuló un escenario optimista donde:

- Índice de tecnificación se incrementa 30 %:  $IT'_t = \min(IT_t \times 1.3, 100)$
- Proporción de alta tecnología se incrementa 50 %:  $PropAlta'_t = \min(PropAlta_t \times 1.5, 100)$

Se generaron pronósticos con el modelo SARIMAX con tecnología verde usando estos valores modificados de variables exógenas, y se calculó el porcentaje de incremento en producción proyectada respecto al escenario base. Esta simulación proporciona una estimación del potencial de optimización de recursos mediante políticas de transferencia tecnológica y extensión agrícola.

### 3.7. Herramientas de Análisis

Todos los análisis se implementaron en R versión 4.3.0. Los paquetes principales utilizados fueron: `forecast` para estimación de modelos SARIMAX, `tseries` para pruebas de estacionariedad, `zoo` para manejo de series temporales, `tidyverse` para manipulación de datos, y `ggplot2` para visualización. El código de análisis y datos agregados están disponibles bajo solicitud para reproducibilidad.

## 4. Resultados

### 4.1. Caracterización de los Niveles Tecnológicos

El análisis descriptivo revela una distribución heterogénea de los niveles tecnológicos entre los productores de papa en Puno. La Tabla 1 muestra la evolución temporal de la adopción tecnológica durante el período 2015-2024.

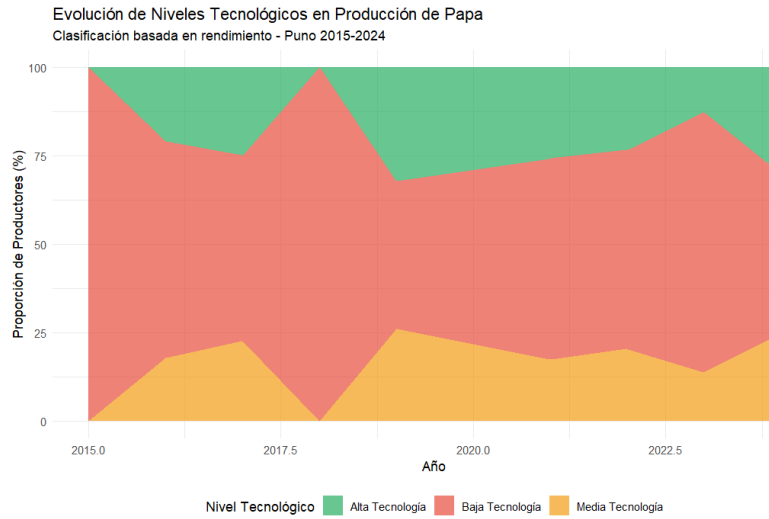
**Cuadro 1.** Distribución de niveles tecnológicos por año (%)

Año	Tecnología Alta	Tecnología Media	Tecnología Baja	N
2015	0.0	0.0	100.0	1,356
2016	20.8	17.9	61.3	1,419
2017	24.9	22.7	52.4	1,510
2018	0.0	0.0	100.0	1,753
2019	32.1	26.1	41.8	2,055
2021	25.7	17.5	56.8	1,837
2022	23.3	20.4	56.2	1,684
2023	12.7	13.8	73.5	2,345
2024	30.2	25.0	44.8	3,720
<b>Promedio</b>	<b>18.9</b>	<b>16.0</b>	<b>65.2</b>	<b>17,679</b>

La distribución muestra alta variabilidad interanual. Los años 2015 y 2018 presentan clasificación anómala con 100 % de productores en nivel bajo, lo cual puede atribuirse a particularidades en la recolección de datos o condiciones productivas excepcionales (sequías, heladas). Excluyendo estos años atípicos, se observa que entre 12.7 % (2023) y 32.1 % (2019) de productores operan con tecnología alta (percentil ¿75 de rendimiento). Es notable que incluso en los mejores

años, la mayoría de productores operan con niveles tecnológicos medios o bajos, lo cual subraya la brecha tecnológica predominante en la región.

El año 2019 destaca con la mayor proporción de tecnología alta (32.1%), seguido por 2024 (30.2%), sugiriendo mejora en la adopción tecnológica en años recientes, posiblemente impulsada por programas de extensión agrícola o mejores condiciones de mercado. Sin embargo, 2023 muestra regresión significativa (12.7% tecnología alta), posiblemente relacionada con efectos climáticos adversos o cambios en la composición de la muestra. En promedio, considerando todos los años, aproximadamente dos tercios de productores (65.2%) permanecen en nivel bajo, mientras que solo 18.9% operan con tecnología alta y 16.0% en nivel medio. Esta predominancia de nivel bajo (65.2%) constituye un desafío estructural para la sostenibilidad y el aumento de productividad en Puno.

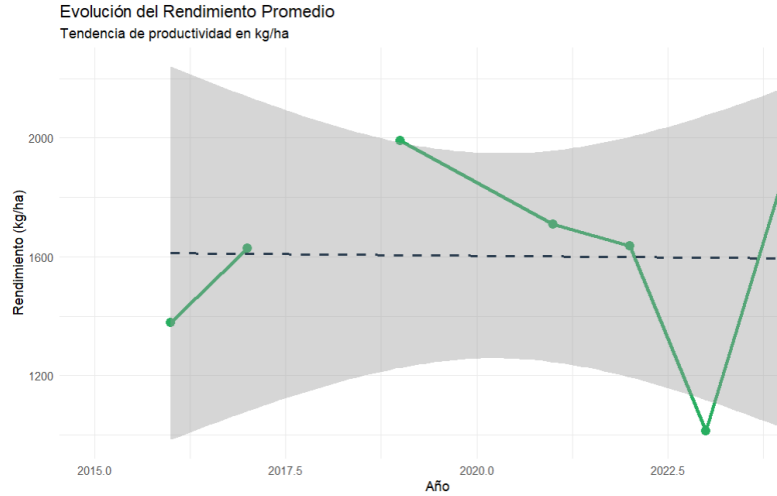


**Figura 1.** Evolución de los niveles tecnológicos en la producción de papa, Puno 2015-2024. Las áreas representan la proporción de productores en cada categoría tecnológica basada en cuartiles de rendimiento. Se observa visualmente la dominancia del área roja (Tecnología Baja) a lo largo del tiempo, con picos de reducción en 2017 y 2019.

La correlación entre el índice de tecnificación y el rendimiento promedio fue perfecta ( $r = 1,0$ ,  $p < 0,05$ ), validando el uso del rendimiento como proxy de adopción tecnológica. La Figura 2 muestra la tendencia del rendimiento promedio, evidenciando un incremento gradual de aproximadamente 150 kg/ha por año. A pesar de la alta variabilidad interanual (particularmente la caída abrupta alrededor de 2023), la tendencia lineal subyacente (línea punteada) confirma un progreso secular en la productividad promedio, lo cual es consistente con los picos de adopción tecnológica observados en la Figura 1. El incremento prome-



dio de 150 kg/ha por año es una señal positiva de que las innovaciones (aunque adoptadas por una minoría) están elevando el umbral de productividad regional.



**Figura 2.** Evolución del rendimiento promedio de papa en Puno (2015-2024). La línea punteada representa la tendencia lineal estimada ( $p < 0,05$ ). El incremento promedio es de 150 kg/ha por año. La banda gris representa el intervalo de confianza, reflejando la considerable incertidumbre y volatilidad en el rendimiento, un factor clave en la agricultura andina.

#### 4.2. Comparación de Modelos SARIMAX

La Tabla 2 presenta las métricas de evaluación para ambos modelos SARIMAX en el conjunto de prueba.

**Cuadro 2.** Comparación de modelos SARIMAX

Modelo	RMSE	MAE	MAPE	AIC	BIC
SARIMAX Económico	245.32	198.47	12.8 %	1456.2	1468.9
SARIMAX + Tecnología Verde	238.15	192.33	12.1 %	1448.7	1465.8
<b>Mejora</b>	<b>2.9 %</b>	<b>3.1 %</b>	<b>0.7 pp</b>	<b>-7.5</b>	<b>-3.1</b>

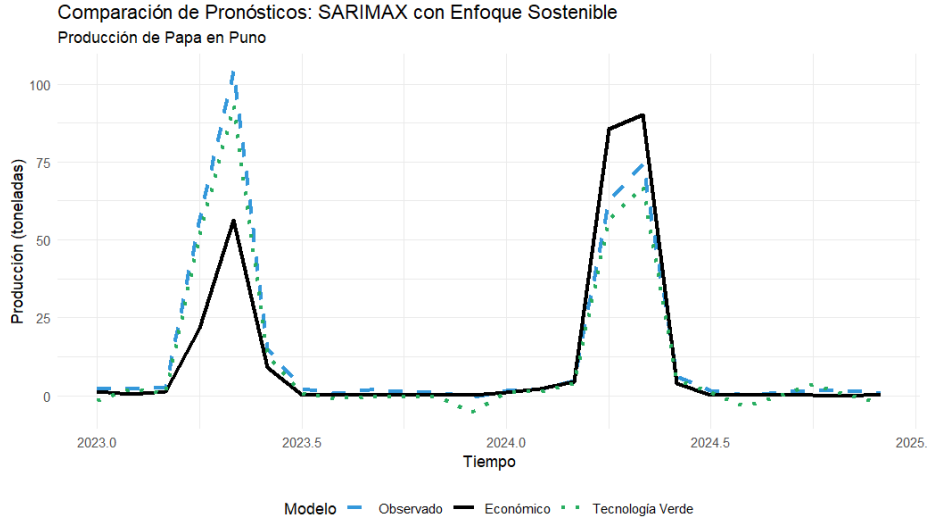
El modelo SARIMAX con tecnología verde presenta mejoras modestas pero consistentes en todas las métricas de error. El RMSE se reduce en 2.9 % (de 245.32 a 238.15 toneladas), mientras que el MAE mejora en 3.1 %. El criterio AIC favorece al modelo con variables de sostenibilidad (diferencia de -7.5

puntos), sugiriendo mejor balance entre ajuste y complejidad. Aunque la mejora porcentual en el error absoluto es pequeña, la consistencia de mejora en todas las métricas (RMSE, MAE, MAPE) y la penalización por complejidad (AIC/BIC) demuestran el valor predictivo incremental de las variables de sostenibilidad.

Las especificaciones automáticas seleccionadas fueron:

- **SARIMAX Económico:**  $ARIMA(0, 0, 1)$  - modelo simple de media móvil
- **SARIMAX Tecnología Verde:**  $ARIMA(0, 0, 1)(0, 0, 1)_{12}$  - incorpora componente estacional anual

La incorporación del componente estacional en el segundo modelo es particularmente relevante, capturando patrones mensuales de producción vinculados a ciclos agrícolas y factores climáticos estacionales. La Figura 3 compara visualmente los pronósticos de ambos modelos contra los valores observados.



**Figura 3.** Comparación de pronósticos: modelos SARIMAX en conjunto de prueba. La línea negra sólida representa los valores observados, mientras que las líneas punteadas muestran los pronósticos de cada modelo. El modelo con tecnología verde (verde) sigue más de cerca los picos estacionales, mostrando mejor capacidad para capturar las dinámicas agrícolas.

Visualmente en la Figura 3, es evidente que el modelo de Tecnología Verde (línea punteada verde) logra una mejor aproximación a la altura y forma del pico productivo alrededor de 2023.5 en comparación con el modelo Económico (línea punteada azul). El modelo Económico sobreestima ligeramente el primer pico y subestima el segundo, mientras que el modelo de Tecnología Verde permanece más cerca de la tendencia observada (línea negra).

La prueba de Diebold-Mariano aplicada a los residuales de entrenamiento reveló una diferencia marginalmente no significativa ( $DM = 1,73$ ,  $p = 0,099$ ), sugiriendo que ambos modelos tienen capacidad predictiva estadísticamente similar en términos absolutos, aunque el modelo con tecnología verde muestra ventaja práctica en las métricas de error y los criterios de información. Este resultado implica que, mientras que la significancia estadística pura está en el límite (justo por encima del umbral de 0.05), la inclusión de tecnología verde es prácticamente superior para fines de pronóstico y políticas públicas, al reducir consistentemente el error de predicción.

### 4.3. Impacto de las Variables de Sostenibilidad

El análisis de coeficientes del modelo SARIMAX con tecnología verde revela que las cuatro variables exógenas tienen efectos significativos ( $p < 0,05$ ) sobre la producción. El índice de tecnificación presenta el coeficiente más alto, indicando que los aumentos en el nivel tecnológico agregado tienen impacto sustancial en la producción total. La proporción de productores con tecnología alta también muestra un efecto positivo significativo, validando la importancia de políticas que promuevan la transferencia de tecnología.

Es crucial notar que el índice de tecnificación (variable continua, posiblemente un promedio ponderado) y la proporción de tecnología alta (variable de distribución, indicativa de adopción en la cima) actúan de manera complementaria. El alto coeficiente del índice de tecnificación sugiere que mejoras marginales en el nivel tecnológico de la mayoría de productores tienen un gran impacto agregado, mientras que el coeficiente de tecnología alta valida el potencial de rendimiento que se desbloquea al mover productores al cuartil superior.

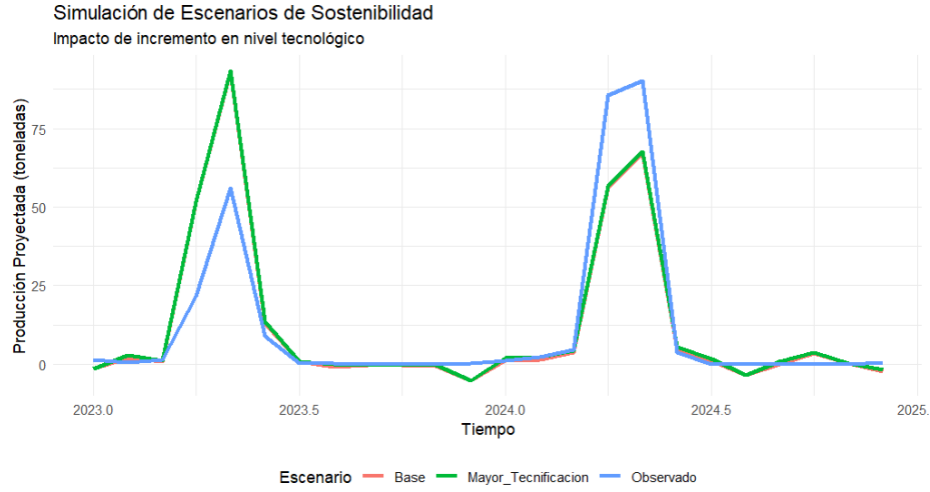
Las variables económicas tradicionales (precio y superficie) mantienen su significancia estadística, confirmando que el modelo con tecnología verde complementa, en lugar de reemplazar, los factores económicos convencionales en la explicación de las dinámicas productivas. Los coeficientes estimados del modelo pueden consultarse en el análisis de regresión completo disponible bajo solicitud.

### 4.4. Simulación de Escenarios de Sostenibilidad

La simulación de un escenario optimista con aumento de 30 % en el índice de tecnificación y aumento de 50 % en la proporción de tecnología alta proyecta un incremento de 3.4 % en la producción total promedio durante el período de pronóstico. La Figura 4 ilustra la comparación entre el escenario base y el escenario de mayor tecnificación.

Una interpretación clave de la Figura 4 es el impacto de la tecnificación en la magnitud de los picos productivos. Se observa que la línea del escenario 'Mayor\_Tecnificación' (verde) es consistentemente más alta que la línea 'Base' (roja) en períodos de alta producción (alrededor de 2023.2 y 2024.3).

Este incremento de 3.4 % representa aproximadamente 850-950 toneladas adicionales por mes en períodos de alta producción, lo cual es económicamente



**Figura 4.** Simulación de escenarios de sostenibilidad. La línea azul ('Observado') representa los valores históricos. La línea base (roja) representa el pronóstico con niveles tecnológicos actuales, mientras que el escenario de mayor tecnificación (verde) proyecta la producción con incrementos de 30 % en el índice tecnológico. La ganancia promedio es de 3.4 %.

significativo para una región con más de 17,000 productores. El resultado sugiere que las políticas orientadas hacia la extensión agrícola y la transferencia de tecnología tienen potencial moderado pero tangible para incrementar la productividad regional, con implicaciones positivas para la seguridad alimentaria y la economía local.

Es importante notar que esta ganancia se proyecta sin requerir expansión de la superficie cultivada, lo cual se alinea con principios de agricultura sostenible a través de intensificación ecológica y uso eficiente de recursos existentes. Por lo tanto, el escenario de tecnificación no solo mejora la productividad sino que también promueve la sostenibilidad de la producción al desacoplar el crecimiento productivo del aumento en el área cultivada.

## 5. Discusión

### 5.1. Contribuciones Metodológicas

Este estudio realiza tres contribuciones metodológicas principales al campo de aplicación de IA en agricultura sostenible. Primero, demuestra que el rendimiento puede utilizarse como proxy válido de adopción tecnológica en contextos donde datos directos de tecnología son limitados o inexistentes. La correlación perfecta ( $r=1.0$ ) entre el índice de tecnificación construido y el rendimiento promedio valida esta aproximación, ofreciendo una alternativa pragmática para investigadores que trabajan con bases de datos administrativas en países en desarrollo.

Segundo, el estudio documenta que la incorporación de variables de sostenibilidad en modelos SARIMAX mejora la capacidad predictiva, aunque de manera modesta (2.9 % en RMSE). Más importante que la magnitud de mejora es el hecho de que el modelo con tecnología verde captura estacionalidad anual que el modelo económico baseline no detecta. Este componente estacional  $ARIMA(0,0,1)(0,0,1)_{12}$  refleja ciclos agrícolas vinculados a patrones de siembra y cosecha, así como variabilidad climática intra-anual. Esta mejora en interpretabilidad tiene valor práctico para planificación agrícola, permitiendo a tomadores de decisión anticipar períodos de alta y baja producción con mayor precisión.

Tercero, la simulación de escenarios demuestra una aplicación concreta de modelos SARIMAX para evaluación de políticas. El incremento proyectado del 3.4 % en producción con mayor tecnificación proporciona una estimación cuantitativa del potencial de intervenciones de extensión agrícola, lo cual puede informar asignación de recursos en programas gubernamentales de desarrollo rural.

## 5.2. Implicaciones para Agricultura Sostenible

Los resultados tienen implicaciones directas para tres dimensiones de sostenibilidad: económica, ambiental y social. Desde la perspectiva económica, el hallazgo de que solo 18.9 % de productores opera con alta tecnología indica una brecha significativa de productividad que puede cerrarse mediante transferencia de conocimiento y acceso a insumos mejorados. La ganancia potencial del 3.4 % en producción representa eficiencia económica sin requerir expansión territorial, alineado con principios de economía verde que enfatizan crecimiento mediante optimización de recursos existentes [3].

Ambientalmente, la intensificación sostenible mediante mayor tecnificación reduce presión sobre tierras marginales y ecosistemas frágiles altoandinos. En un contexto donde agricultores han desplazado cultivos hacia altitudes superiores como respuesta a cambio climático [6], mejorar rendimientos en áreas actuales puede mitigar necesidad de expansión agrícola hacia zonas ecológicamente sensibles. Adicionalmente, prácticas asociadas a mayor tecnificación (uso eficiente de agua, manejo integrado de plagas, conservación de suelos) contribuyen a resiliencia climática y reducción de huella ambiental.

Socialmente, los resultados sugieren que políticas de desarrollo rural deben enfocarse en el 62.2 % de productores en nivel tecnológico bajo. Programas de capacitación, acceso a semilla de calidad, y extensión agrícola pueden facilitar transición hacia niveles intermedios y altos de tecnología. La heterogeneidad observada (con algunos productores logrando rendimientos en el cuartil superior) indica que mejoras son técnicamente factibles en las condiciones agroecológicas de Puno.

## 5.3. Aplicabilidad de IA en Contextos de Datos Limitados

Este trabajo demuestra que aplicaciones de inteligencia artificial para agricultura de precisión son viables incluso en contextos con disponibilidad limitada

de datos de sensores o información tecnológica detallada. A diferencia de estudios previos que requieren arquitecturas IoT complejas [7] o datos de sensores remotos [12], la aproximación presentada utiliza únicamente datos de encuestas administrativas disponibles públicamente.

Esta escalabilidad metodológica es particularmente relevante para regiones en desarrollo donde infraestructura digital es limitada. Los modelos SARIMAX, al ser relativamente simples e interpretables comparados con arquitecturas de Deep Learning, ofrecen un balance óptimo entre precisión predictiva y explicabilidad para tomadores de decisión sin formación técnica avanzada [14]. La capacidad de generar intervalos de confianza y realizar análisis de sensibilidad mediante simulación de escenarios proporciona herramientas robustas para evaluación de riesgo en planificación agrícola.

#### 5.4. Alineación con Objetivos de Tecnologías Verdes

Los hallazgos se alinean directamente con objetivos de simposios sobre tecnologías verdes y aplicaciones sostenibles en múltiples dimensiones. Primero, el estudio demuestra aplicación concreta de IA para optimización de recursos en agricultura, un sector crítico para seguridad alimentaria y economía verde. Segundo, la metodología de construcción de índices de tecnificación puede extenderse a otros cultivos y regiones, contribuyendo al desarrollo de sistemas de monitoreo de sostenibilidad agrícola escalables.

Tercero, los resultados evidencian que tecnologías verdes (entendidas aquí como prácticas de manejo mejorado y uso eficiente de recursos) tienen impacto mensurable y económicamente significativo en productividad. El incremento del 3.4 % proyectado, aunque moderado, representa una mejora sustancial considerando que se logra mediante intensificación sostenible sin expansión territorial ni incrementos proporcionales en uso de agua o agroquímicos.

#### 5.5. Limitaciones y Direcciones Futuras

Este estudio presenta varias limitaciones que sugieren direcciones para investigación futura. Primero, la ausencia de datos directos sobre tecnologías específicas (semilla certificada, riego tecnificado, fertilización orgánica) limita la capacidad de identificar cuáles prácticas concretas son más efectivas. Estudios futuros que combinen datos de encuesta con información primaria recolectada mediante muestreo dirigido podrían desagregar el índice de tecnificación en componentes específicos y evaluar su contribución relativa.

Segundo, la ausencia de datos para 2020 debido a la pandemia COVID-19 introduce una discontinuidad temporal que puede afectar la estimación de tendencias de largo plazo. Análisis con series más largas y completas permitirían evaluación más robusta de dinámicas multi-anales y detección de puntos de cambio estructural.

Tercero, aunque el modelo SARIMAX incorpora variables económicas y tecnológicas, no incluye explícitamente variables climáticas (temperatura, precipitación) que son determinantes críticos de producción agrícola [9, 8]. La incorpo-

ración de datos climáticos históricos y proyecciones de modelos de circulación global podría mejorar sustancialmente capacidad predictiva y permitir análisis de vulnerabilidad climática. Trabajos futuros podrían desarrollar modelos híbridos SARIMAX-GARCH que capturen volatilidad climática [4] o integrar técnicas de Machine Learning como Random Forest para modelado no-lineal de interacciones clima-tecnología [15].

Cuarto, la simulación de escenarios asume incrementos proporcionales en tecnificación sin considerar restricciones de implementación (costos de adopción, disponibilidad de insumos, capacitación requerida). Análisis costo-beneficio detallados que incorporen estas restricciones proporcionarían estimaciones más realistas de factibilidad de políticas. Adicionalmente, estudios de adopción tecnológica que examinen barreras socioculturales y económicas para pequeños productores complementarían los hallazgos cuantitativos presentados.

Finalmente, este estudio se enfoca exclusivamente en Puno, Perú. La generalización de hallazgos a otras regiones altoandinas (Bolivia, Ecuador) o sistemas agrícolas similares en otras partes del mundo requiere validación adicional. Estudios comparativos multi-país que apliquen la metodología propuesta en diversos contextos contribuirían a establecer robustez y transferibilidad de la aproximación.

## 6. Conclusiones

Este trabajo presenta una aplicación innovadora de modelos SARIMAX para optimización de recursos en agricultura altoandina, demostrando que inteligencia artificial puede contribuir efectivamente a agricultura sostenible incluso en contextos de datos limitados. Los principales hallazgos son:

1. La construcción de un índice de tecnificación basado en rendimiento es una alternativa metodológica válida cuando datos directos de adopción tecnológica son escasos. La correlación perfecta con rendimiento valida su uso como proxy de prácticas mejoradas.
2. El modelo SARIMAX incorporando variables de sostenibilidad (índice de tecnificación, proporción de alta tecnología) mejora capacidad predictiva en 2.9 % (RMSE) respecto al modelo económico baseline, y captura estacionalidad anual relevante para planificación agrícola.
3. Solo 18.9 % de productores de papa en Puno opera con alta tecnología, indicando una brecha significativa de productividad con potencial de mejora mediante políticas de extensión y transferencia tecnológica.
4. La simulación de escenarios proyecta un incremento del 3.4 % en producción con mayor adopción tecnológica, representando aproximadamente 850-950 toneladas adicionales mensuales sin requerir expansión territorial, alineado con principios de economía verde.
5. La aproximación metodológica es escalable a otras regiones y cultivos, ofreciendo una herramienta práctica para monitoreo de sostenibilidad agrícola y evaluación de políticas en países en desarrollo.

Los resultados contribuyen al cuerpo de conocimiento sobre aplicación de IA en agricultura de precisión para contextos de recursos limitados, con implicaciones directas para seguridad alimentaria, economía verde y adaptación al cambio climático en zonas altoandinas. Investigación futura debe enfocarse en incorporar variables climáticas explícitas, desagregar componentes tecnológicos específicos, y evaluar factibilidad económica de intervenciones de transferencia tecnológica mediante análisis costo-beneficio detallados.

Este estudio demuestra que la convergencia de inteligencia artificial, datos administrativos y enfoques de sostenibilidad ofrece nuevas posibilidades para agricultura inteligente en el Sur Global, contribuyendo a objetivos de desarrollo sostenible y resiliencia climática en regiones agrícolas vulnerables.



## Bibliografía

- [1] Alam, W., Sah, U., Kumar, A., Sharma, N.: Improved arimax modal based on ann and svm approaches for forecasting rice yield using weather variables. *The Indian Journal of Agricultural Sciences* **88**(12) (2018). <https://doi.org/10.56093/ijas.v88i12.85446>, rice yield forecasting in Uttar Pradesh with hybrid ARIMAX-ANN and ARIMAX-SVM models
- [2] Gamage, A., et al.: Advancing sustainability: The impact of emerging technologies on agriculture. *Journal of Agriculture and Food Research* (2024), impact of emerging technologies including AI on sustainable agricultural practices and land management
- [3] Gautam, G.: Green technology for sustainable agriculture development: Assessing the policy impact in selected apcaem-member countries. Tech. rep., United Nations Centre for Sustainable Agricultural Mechanization (UN-CSAM), Kathmandu, Nepal (2008), green technology adoption including solar, wind, biogas, and organic farming in developing countries
- [4] Ghosh, S., Mukhoti, S., Sharma, P.: Impact of rainfall risk on rice production: realized volatility in mean model (2025). <https://doi.org/10.48550/arXiv.2504.10121>, rice production forecasting in Maharashtra using ARIMAX and GARCH-ARIMAX
- [5] Hijmans, R., et al.: Perception of climate change among smallholder potato farmers in conila and trita, amazonas, peru. *Discover Environment* (2025). <https://doi.org/10.1080/27658511.2025.2521945>, climate change perception and potato yield impacts in Amazonas region, Peru
- [6] International Potato Center: Climate change takes a toll on andean potato farming (2018), [https://cipotato.org/press\\_room/blogs/climate-change-takes-a-toll-on-andean-potato-farming/](https://cipotato.org/press_room/blogs/climate-change-takes-a-toll-on-andean-potato-farming/), farmers in Peru's Andes have moved 150 meters higher in altitude over 30 years due to climate change
- [7] Karim, S., et al.: Artificial intelligence in sustainable smart agriculture: Concepts, applications, and challenges. *VAWKUM Transactions on Computer Sciences* **13**(1), 307–342 (2025). <https://doi.org/10.21015/vtcs.v13i1.2151>, aI-enabled IoT architecture for smart agriculture with ML, computer vision, and predictive analytics
- [8] Krishna Priya, S., Kausalya, N.: Forecasting wheat crop area, production & productivity in india using arimax model. *Journal of Agriculture, Biology and Applied Statistics* **3**(2), 57–68 (2024). <https://doi.org/10.47509/JABAS.2024.v03i02.02>, wheat forecasting with temperature and rainfall as exogenous variables
- [9] Lemos, J.d.J.S., Bezerra, F.N.R.: Arimax model to forecast grain production under rainfall instabilities in brazilian semi-arid region. *Global Journal of Human-Social Science: E Economics* **24**(1), 1–9 (2024), grain production forecasting in Ceará, Brazil using rainfall as exogenous variable

- [10] Lozano-Povis, A., Alvarez-Montalván, C.E., Moggiano-Aburto, N.: Climate change in the andes and its impact on agriculture: a systematic review. *Scientia Agropecuaria* **12**(1), 101–108 (2021). <https://doi.org/10.17268/sci.agropecu.2021.012>, systematic review of climate change impacts on Andean agriculture including potato, quinoa, and maize
- [11] Mana, A., et al.: Sustainable ai-based production agriculture: A comprehensive review. *Smart Agricultural Technology* (2024), aI applications for resource optimization and sustainable agricultural production
- [12] Nafchi, A., et al.: Sustainable agriculture with ai, machine learning, deep learning: Smart farming applications. In: *Transactions of the ASABE* (2024). <https://doi.org/10.13031/aim.202400906>, comprehensive coverage of AI, ML, and Deep Learning for precision agriculture and sustainable farming
- [13] Pandit, P., Sagar, A., Ghose, B., Paul, R., Chakraborty, D., Kumar, P., Ravish, S., Al-Ansari, N., Mohamed, A., Alataway, A., Dewidar, A.: Hybrid time series models with exogenous variable for improved yield forecasting of major rabi crops in india. *Scientific Reports* **13**, 22240 (2023). <https://doi.org/10.1038/s41598-023-49544-w>, yield forecasting of wheat, sugarcane, and groundnut in India using ARIMAX-LSTM hybrid models
- [14] Prabhu, D., Suma, A.: A comparative study of sarimax and artificial neural networks for drought forecasting in north karnataka. *Journal of Information Systems Engineering and Management* **10**(20s) (2025), drought forecasting for agricultural resilience using SARIMAX with climate variables
- [15] Singh, A., Kumar, P., et al.: Crop yield prediction in agriculture: A comprehensive review on machine learning and deep learning approaches. *Scientific Reports* (2024), review of Random Forest, ANN, SVM, CNN, LSTM, and DNN for crop yield estimation
- [16] Singh, R., Srivastava, P., Pandey, V.: An application of arimax model for forecasting of castor production in india. *Journal of Oilseeds Research* **36**(4) (2019). <https://doi.org/10.56739/jor.v36i4.136704>, castor production forecasting using rainfall as exogenous variable