

## Multimodal machine learning application to optimize cocoa production through precision agriculture in tropical areas

### Aplicación de aprendizaje automático multimodal para optimizar la producción de cacao mediante agricultura de precisión en zonas tropicales

**Autor:**

Mgs. Mendoza-Cabrera, Denis Darío  
UNIVERSIDAD ESTATAL DE MILAGRO  
Magister en Gerencia en Tecnología de la Información y Comunicación  
Milagro – Ecuador

 [dmendozac2@unemi.edu.ec](mailto:dmendozac2@unemi.edu.ec)  
 <https://orcid.org/0000-0002-4625-7835>

Mgs. Guevara-Serrano, Jorge Dumar  
UNIVERSIDAD ESTATAL DE MILAGRO  
Magister en Tecnologías de la información  
Mención en Transformación Digital e Innovación  
Milagro – Ecuador

 [jguevaras1@unemi.edu.ec](mailto:jguevaras1@unemi.edu.ec)  
 <https://orcid.org/0009-0004-2700-5941>

Mgs. Guevara-Reyes, Rodrigo Josué  
UNIVERSIDAD ESTATAL DE MILAGRO  
Magister en Sistemas de Información  
Mención en Inteligencia de Negocios y Analítica de Datos Masivos  
Milagro – Ecuador

 [ruevarar@unemi.edu.ec](mailto:rguevarar@unemi.edu.ec)  
 <https://orcid.org/0009-0009-7575-6148>

Fechas de recepción: 19-SEP-2025 aceptación: 15-NOV-2025 publicación: 30-DIC-2025

 <https://orcid.org/0000-0002-8695-5005>  
<http://mqrinvestigar.com/>



Vol.9-N° 4, 2025, pp. 01-19

Journal Scientific MQRInvestigar 1

## Resumen

Este artículo presenta la propuesta de PREDICT-CACAO, una plataforma inteligente basada en aprendizaje automático multimodal orientada a optimizar la producción de cacao en zonas tropicales. La investigación aborda las limitaciones estructurales del sistema cacaotero, como la baja capacidad predictiva, el uso ineficiente de recursos y la alta incidencia de enfermedades, mediante la integración de datos climáticos, edáficos, visuales e históricos en modelos de machine learning, visión por computador e inteligencia artificial explicable (XAI).

El marco metodológico PREDICT-CACAO, comprende cinco fases: adquisición de datos multiescalares, modelado predictivo de rendimiento y enfermedades, implementación de XAI, desarrollo de una plataforma digital adaptativa y validación participativa en campo. Los resultados esperados incluyen una mejora significativa en la precisión de las predicciones, una detección temprana de enfermedades con más del 90% de efectividad, y una mayor eficiencia en el uso de insumos agrícolas. Asimismo, se proyecta el fortalecimiento de las capacidades tecnológicas de los productores y una mayor apropiación de la agricultura digital en contextos rurales.

El proyecto demuestra que la integración de inteligencia artificial y agricultura de precisión no solo impulsa la productividad y sostenibilidad del cacao, sino que también promueve la inclusión tecnológica, la resiliencia climática y el desarrollo rural sostenible.

**Palabras clave:** Aprendizaje automático multimodal; agricultura de precisión; inteligencia artificial explicable; cacao; sostenibilidad agrícola; inclusión digital rural.



## Abstract

This article presents the development and impact of PREDICT-CACAO, an intelligent platform based on multimodal machine learning designed to optimize cocoa production in tropical regions. The research addresses structural limitations in cocoa farming—such as low predictive capacity, inefficient resource use, and high disease incidence—by integrating climatic, edaphic, visual, and historical data into models using machine learning, computer vision, and explainable artificial intelligence (XAI).

The study was developed under the PREDICT-CACAO methodological framework, consisting of five phases: multiscale data acquisition, predictive modeling of yield and diseases, XAI implementation, adaptive digital platform development, and participatory field validation. Expected results include a significant improvement in prediction accuracy, early disease detection exceeding 90% effectiveness, and greater efficiency in agricultural input management. Additionally, the system is expected to strengthen farmers' technological capacities and enhance digital agriculture adoption in rural settings.

The project demonstrates that integrating artificial intelligence and precision agriculture not only enhances cocoa productivity and sustainability, but also fosters technological inclusion, climate resilience, and sustainable rural development.

**Keywords:** Multimodal machine learning; precision agriculture; explainable artificial intelligence; cocoa; agricultural sustainability; digital inclusion.



## Introducción

La producción de cacao en regiones tropicales representa no solo un motor económico relevante para países como Colombia, Ecuador, Perú o Costa de Marfil, sino también un elemento clave en la configuración sociocultural de muchas comunidades rurales. Este cultivo, altamente sensible a variaciones climáticas, enfermedades y deficiencias en el manejo agronómico, se encuentra en un punto crítico: la creciente demanda internacional por cacao de calidad contrasta con las limitaciones estructurales de los sistemas de producción tradicionales (Obando et al., 2023).

La revolución digital en la agricultura, marcada por la incorporación de sensores, satélites, imágenes multiespectrales, machine learning (ML) y análisis de grandes volúmenes de datos, ha demostrado su eficacia en cultivos como arroz, maíz, café o frutas tropicales. En estos sistemas, el uso de datos heterogéneos ha permitido diseñar estrategias adaptativas para maximizar la productividad, reducir pérdidas por enfermedades y optimizar el uso de recursos naturales (Talero-Sarmiento et al., 2025).

Sin embargo, el cacao aún no se ha beneficiado plenamente de estos desarrollos tecnológicos. A pesar de algunos avances en la recolección de datos y modelamiento ambiental para el cultivo, la integración efectiva de herramientas de inteligencia artificial en su gestión agronómica sigue siendo incipiente (De la Peña & Granados, 2024). Esta brecha tecnológica limita la capacidad de los agricultores para anticiparse a eventos críticos como brotes de enfermedades, deficiencias nutricionales o estrés hídrico, y perpetúa una dependencia de decisiones empíricas más que científicamente informadas.

En este contexto, surge la necesidad de un enfoque multimodal de aprendizaje automático que combine diversas fuentes de datos como: visuales, climáticas, edáficas e históricas con técnicas modernas de inteligencia artificial, como redes neuronales profundas, aprendizaje auto-supervisado y modelos interpretables. Este enfoque tiene el potencial de transformar la forma en que se toma decisiones, especialmente en zonas tropicales vulnerables a la variabilidad climática y limitaciones estructurales (Logacjov, 2024).

## Fundamentación teórica

El concepto de agricultura de precisión ha evolucionado en las últimas décadas como una estrategia para incrementar la eficiencia del manejo agrícola mediante el uso de tecnologías digitales, sensores, sistemas de información geográfica y análisis de datos. Esta aproximación se basa en la recopilación continua y localizada de información sobre variables críticas del entorno agrícola con el fin de apoyar la toma de decisiones específicas para cada parcela o cultivo (Riaño et al., 2023).



En este contexto, el ML se ha convertido en una herramienta clave para transformar grandes volúmenes de datos en conocimiento práctico. El uso de modelos supervisados y no supervisados permite detectar patrones, predecir rendimientos y anticipar condiciones adversas, sin embargo, uno de los desafíos actuales es la integración de múltiples fuentes de datos dentro de modelos robustos y explicables (Talero-Sarmiento et al., 2025).

Aplicaciones exitosas de ML en agricultura incluyen la detección de enfermedades en hojas mediante redes convolucionales (CNN), el análisis de imágenes satelitales para evaluar salud del cultivo y la predicción de rendimiento a través de algoritmos de ensamble que consideran variables edafoclimáticas (Şener & Ergen, 2024). Estas metodologías ya se han aplicado con éxito en sistemas productivos como el café, los cítricos o la palma aceitera, pero su adopción en este tema ha sido limitada.

Además, estudios recientes han destacado la importancia de incorporar inteligencia artificial explicable (XAI), que permita interpretar las decisiones del modelo y generar confianza en los usuarios finales, especialmente cuando se trata de pequeños productores que no están familiarizados con sistemas complejos (De la Peña & Granados, 2024). Esta necesidad de interpretabilidad es aún más crítica en zonas rurales, donde la apropiación tecnológica está mediada por factores socioculturales, acceso a infraestructura digital y asistencia técnica.

Por otro lado, investigaciones como la de (Obando et al., 2023) demuestran que la gestión hídrica en el cultivo de cacao está fuertemente influenciada por la escasa disponibilidad de datos precisos y la falta de tecnologías de monitoreo, lo que subraya la necesidad de enfoques automatizados y basados en datos para cerrar brechas.

De esta forma, la presente investigación se fundamenta en una intersección crítica entre agricultura de precisión, aprendizaje automático multimodal, teledetección, ontologías agrícolas y participación campesina mediada por inteligencia artificial explicable. Esta integración busca no solo optimizar la producción de cacao, sino también generar herramientas accesibles, adaptables y contextualizadas para zonas tropicales con alto potencial productivo.

## Propuesta de solución

Frente a los múltiples desafíos que se enfrenta desde la baja eficiencia en el uso de insumos hasta la limitada anticipación de enfermedades y rendimientos, se propone el diseño de una plataforma inteligente basada en aprendizaje automático multimodal, además, esta solución busca integrar datos heterogéneos provenientes del entorno agrícola y transformarlos en decisiones agronómicas precisas, contextualizadas e interpretables. La propuesta se fundamenta en tres ejes tecnológicos interrelacionados:

### 1. Modelos de predicción agroproductiva con datos heterogéneos

Se desarrollará un sistema predictivo basado en algoritmos de ML que combine datos estructurados y no estructurados. Estos incluirán variables agroclimáticas, características edáficas, series históricas de rendimiento y datos visuales, además, este enfoque multimodal permitirá superar las limitaciones de modelos univariantes y capturar interacciones no lineales entre factores ambientales y productivos (Talero-Sarmiento et al., 2025).



Se explorarán arquitecturas como XGBoost, redes neuronales profundas y modelos híbridos (MLP + CNN) capaces de manejar entradas heterogéneas. El sistema se validará mediante técnicas de cross-validation y se entrenará sobre datos regionales previamente normalizados y etiquetados por expertos agrónomos y técnicos de campo (Riaño et al., 2023).

## 2. Detección automatizada de enfermedades mediante visión por computador

Un segundo componente estará orientado a la identificación precoz de enfermedades foliares y de fruto, a partir del procesamiento de imágenes. Se implementarán algoritmos de segmentación semántica entrenados con bases de datos curadas de hojas enfermas de cacao, con clasificación por tipo de patología. Esta herramienta permitirá a los agricultores diagnosticar de forma anticipada la presencia de enfermedades como monilia, mazorca negra o escoba de bruja, reduciendo la dependencia de observaciones visuales tardías y subjetivas (Şener & Ergen, 2024).

## 3. Sistema de XAI orientado al agricultor

Reconociendo la necesidad de transparencia en los modelos predictivos, se integrará un componente de XAI. Esto permitirá generar explicaciones comprensibles sobre por qué el sistema recomienda una determinada acción. Estas explicaciones estarán basadas en técnicas como SHAP (SHapley Additive exPlanations) o LIME (Local Interpretable Model-Agnostic Explanations), y serán traducidas a un lenguaje comprensible para usuarios no técnicos (De la Peña & Granados, 2024).

Este enfoque, además de aumentar la confianza del usuario, facilita la apropiación tecnológica por parte de agricultores de distintas generaciones y niveles educativos, como lo evidencian los estudios sobre adopción tecnológica en el sector cacaotero colombiano (Obando et al., 2023).

## 4. Infraestructura digital distribuida y adaptable al contexto tropical

La plataforma se diseñará con una arquitectura modular escalable, capaz de operar tanto en entornos de conectividad limitada como en sistemas centralizados. Esto incluye almacenamiento en la nube para entrenamiento de modelos y versiones locales para uso en campo, además, la interfaz será co-diseñada con agricultores, técnicos y extensionistas a través de una metodología de living lab que recoja necesidades reales del entorno (Pinardi et al., 2023).



## Metodología:

Framework PREDICT-CACAO – Plataforma de Recolección, Evaluación, Diagnóstico, Integración y Computación Territorial del Cacao

La metodología desarrollada para este estudio se denomina PREDICT-CACAO, un enfoque sistemático que combina ciencia de datos, XAI, visión por computador y participación rural activa, orientado al diseño, implementación y validación de una plataforma inteligente para la gestión agronómica del cacao en zonas tropicales. Este marco metodológico está diseñado para ser escalable, adaptable y replicable en diferentes contextos agroecológicos y socioproyectivos.

El modelo se despliega en cinco macrofases operativas:

### Fase 1: Adquisición estructurada de datos multiescalares y multimodales

Esta primera etapa se enfoca en la creación de un ecosistema de datos robusto y representativo del entorno agronómico del cacao. Se recopilan y organizan datos desde cinco dominios clave:

- Datos climáticos históricos y actuales, extraídos de fuentes abiertas y estaciones meteorológicas locales. Se consideran variables como temperatura media, amplitud térmica, evapotranspiración, precipitación acumulada, humedad relativa y radiación solar (Talero-Sarmiento et al., 2025).
- Datos edafológicos, recolectados mediante muestreo georreferenciado, con análisis de textura, pH, contenido de materia orgánica, capacidad de campo y retención de nutrientes, integrando además mapas derivados de SoilGrids250m y catastro agrícola (Riaño et al., 2023).
- Datos de rendimiento histórico, construidos en colaboración con asociaciones de cacaoteros, registros de trazabilidad productiva y bases institucionales como las de FEDECACAO.
- Imágenes visuales RGB y multiespectrales, obtenidas por drones (RedEdge, NIR) y satélites Sentinel-2. Se generarán índices como NDVI, GNDVI y SAVI para estimar vigor vegetativo, estrés hídrico y cobertura (Ramos et al., 2024).
- Observaciones fenológicas y sanitarias, recolectadas mediante formularios móviles y captura directa en campo para validar eventos agronómicos.

### Fase 2: Modelado predictivo multimodal de rendimiento y enfermedades

#### 2.1. Módulo de predicción de rendimiento

- Se implementará un modelo de aprendizaje automático supervisado, estructurado en dos canales.
- Canal tabular, para datos estructurados, procesado mediante algoritmos ensemble y redes densas.
- Canal visual, para imágenes multiespectrales, procesado mediante CNN y atención espacial.



Ambos canales convergen en un modelo de fusión para emitir predicciones ajustadas a cada finca/parcela. Se evaluará mediante cross-validation, R<sup>2</sup>, RMSE y MAE sobre datos independientes.

## 2.2. Módulo de detección automatizada de enfermedades

Se entrenarán modelos de segmentación semántica de imágenes de hojas y frutos para identificar síntomas de enfermedades clave: moniliasis, escoba de bruja y antracnosis. Se compararán arquitecturas como:

- PDRDSegNet (red ligera y eficiente para dispositivos móviles).
- DeepLabV3+ (alta precisión en clasificación multiclas).
- UNet y SegNet (modelos base).

Las imágenes serán anotadas por fitopatólogos, usando herramientas como LabelBox, y se aplicará aumento de datos (data augmentation) para mejorar la generalización, además, se validará con IoU, F1-Score, precisión pixel a pixel y exactitud global.

## Fase 3: Implementación de XAI

Para aumentar la confianza y usabilidad del sistema, se desarrollará un módulo de explicabilidad basado en:

- SHAP, para modelos tabulares, mostrando la contribución de cada variable a la predicción de rendimiento.
- LIME y Grad-CAM, para visualizar qué zonas de la imagen fueron determinantes en la clasificación de enfermedades.

Las salidas interpretativas serán textualizadas y traducidas a representaciones visuales intuitivas, que se desplegarán en la interfaz móvil como mensajes explicativos, íconos codificados por colores y consejos agronómicos ajustados al contexto (De la Peña & Granados, 2024).

## Fase 4: Desarrollo de la plataforma digital adaptativa

El sistema PREDICT-CACAO será desplegado como una plataforma híbrida web-móvil, diseñada para ambientes con conectividad variable. Las características incluyen:

- Frontend en Flutter or React Native.
- Backend con microservicios en Flask y bases de datos.
- Módulo offline que almacena decisiones temporales hasta sincronización.
- La interfaz será desarrollada bajo principios de UX rural, con lenguaje visual amigable, íconos reconocibles y compatibilidad multilingüe.

## Fase 5: Validación piloto con enfoque participativo

Se ejecutará un piloto controlado en dos regiones productoras de cacao con distintas condiciones agroecológicas. La muestra incluirá:





- Se estima diez fincas organizadas en clústeres de afinidad tecnológica.
- Comparación entre lotes con y sin uso del sistema durante al menos 9 meses.

#### Las métricas de validación incluirán:

- Precisión del modelo vs rendimiento real.
- Efectividad de la detección temprana vs diagnóstico manual.
- Tiempo de reacción ante enfermedades.
- Evaluación de percepción tecnológica: confianza, utilidad, comprensibilidad y disposición de uso.

Se aplicará también un enfoque de retroalimentación iterativa, mediante el cual los datos nuevos recopilados alimentarán el modelo para recalibración periódica, asegurando su mejora continua.

En síntesis, PREDICT-CACAO propone una metodología innovadora, integral y contextualizada, que articula el rigor técnico del aprendizaje automático con el conocimiento local y la participación activa de los actores rurales. Este enfoque no solo busca mejorar los resultados productivos, sino también empoderar a los agricultores como usuarios activos de tecnologías avanzadas, generando impactos en eficiencia, sostenibilidad y resiliencia frente al cambio climático.

## Resultados esperados

La implementación del enfoque metodológico PREDICT-CACAO está orientada a generar una transformación integral en la gestión de la producción de cacao en zonas tropicales, a través del aprovechamiento de herramientas de inteligencia artificial y agricultura de precisión. Se espera que el sistema desarrollado aporte resultados significativos en al menos cuatro dimensiones: técnica, agronómica, social y tecnológica.

En primer lugar, desde el punto de vista técnico y científico, se anticipa la obtención de un modelo predictivo robusto, capaz de integrar múltiples fuentes de datos heterogéneos para estimar con alta precisión el rendimiento del cacao a nivel de parcela. A diferencia de los modelos tradicionales univariantes, esta propuesta multimodal permitirá capturar interacciones complejas y no lineales entre las variables del entorno y las respuestas fisiológicas del cultivo. Se espera alcanzar coeficientes de determinación superiores a 0.85 y errores de predicción significativamente bajos, incluso bajo condiciones de alta variabilidad climática o limitaciones edáficas (Talero-Sarmiento et al., 2025).

Simultáneamente, se proyecta la validación de un sistema automatizado de diagnóstico fitosanitario basado en visión por computador, capaz de identificar con alta sensibilidad y especificidad enfermedades como la moniliasis, la escoba de bruja o la antracnosis. Mediante modelos de segmentación semántica, entrenados con imágenes de campo y curados por expertos fitopatólogos, se espera alcanzar precisiones superiores al 90% en la clasificación de patologías y una capacidad operativa adaptada a dispositivos móviles de gama media, lo



cual es crucial en contextos rurales donde el acceso a hardware especializado es limitado (Şener & Ergen, 2024).

Desde la dimensión agronómica, se espera que la implementación del sistema contribuya a una mejora medible en la eficiencia productiva del cacao, tanto en términos de rendimiento por hectárea como de uso racional de insumos. Al contar con predicciones anticipadas y diagnósticos automatizados, los agricultores podrán planificar mejor la aplicación de fertilizantes, ajustar el riego según las condiciones climáticas proyectadas y tomar decisiones más informadas sobre el manejo del cultivo. Esto se traducirá en una reducción significativa del uso ineficiente de recursos, una menor exposición a riesgos fitosanitarios, y un manejo más sostenible y adaptativo del sistema productivo (Obando et al., 2023).

En lo social, uno de los principales impactos esperados es el fortalecimiento de las capacidades de los agricultores frente al uso de tecnologías digitales. La integración de XAI tiene un papel clave en este sentido, ya que permitirá generar confianza en el sistema al ofrecer explicaciones claras, comprensibles y contextualizadas sobre el porqué de cada recomendación.

Se espera, en este marco, un alto grado de aceptación por parte de los agricultores que participen en el piloto de campo, especialmente si el sistema demuestra utilidad práctica y se adapta a sus realidades productivas, lingüísticas y culturales. Además, la interacción con la plataforma servirá como vehículo de capacitación digital, permitiendo a los productores familiarizarse con nuevas herramientas, conceptos y prácticas de agricultura de precisión (De la Peña & Granados, 2024).

En el plano tecnológico, se espera culminar el desarrollo de un prototipo funcional de la plataforma digital, con capacidad para operar en condiciones reales de campo, tanto en su versión web como móvil. Esta herramienta no solo será útil para el cultivo de cacao, sino que podrá ser adaptada a otros cultivos tropicales que enfrentan desafíos similares, como el café, la palma aceitera o los frutales de exportación.

De este modo, se establecerán las bases para un modelo replicable y escalable, con potencial de impacto regional en América Latina y otras zonas tropicales de África o el sudeste asiático (González-Orozco et al., 2024).

Finalmente, como resultado transversal del proyecto, se generará una base de datos agroclimática, edáfica y visual de alta resolución, que podrá ser utilizada tanto para futuras investigaciones científicas como para el diseño de políticas públicas en materia de planificación agrícola, manejo de riesgos climáticos y desarrollo rural. Esta base de datos constituirá un insumo estratégico para fortalecer la soberanía tecnológica en contextos rurales, permitiendo que las decisiones productivas se basen cada vez más en evidencia empírica y no exclusivamente en experiencia acumulada o intuición.

En suma, los resultados esperados del sistema PREDICT-CACAO no se limitan a mejorar indicadores agronómicos, sino que apuntan a una transformación estructural de la forma en que se produce, se gestiona y se toma decisiones en la cacaocultura tropical, con una mirada



de largo plazo que articula sostenibilidad, eficiencia, innovación y justicia digital para los pequeños productores.

## Discusión

El desarrollo e implementación del sistema PREDICT-CACAO debe ser comprendido no solo como una innovación tecnológica, sino como una respuesta estructural a múltiples limitaciones históricas que enfrenta la cacaocultura tropical, especialmente en América Latina. La producción de cacao, aunque de alta relevancia económica y cultural, ha sido tradicionalmente gestionada bajo esquemas empíricos, con limitada incorporación de herramientas científicas para el monitoreo, la predicción o la toma de decisiones agronómicas (Obando et al., 2023). Esta brecha entre el potencial productivo del cultivo y las capacidades reales de gestión tecnológica se ha traducido en bajos rendimientos, alta vulnerabilidad fitosanitaria y dependencia de prácticas reactivas ante las adversidades climáticas o biológicas.

En ese contexto, la propuesta de integrar aprendizaje automático multimodal para optimizar la producción de cacao no representa una simple digitalización de procesos, sino una reestructuración profunda de la relación entre el productor, el ambiente agrícola y la información técnica. A diferencia de otras soluciones fragmentadas que solo abordan componentes aislados, PREDICT-CACAO plantea un enfoque holístico, en el que la información edafoclimática, visual y productiva es fusionada y procesada en tiempo casi real, ofreciendo al agricultor recomendaciones personalizadas, interpretables y adaptadas a su realidad específica (Talero-Sarmiento et al., 2025).

Un aporte crucial de este enfoque es el uso de XAI, que rompe con la opacidad algorítmica característica de muchos modelos de aprendizaje profundo, sin embargo, este componente no solo es deseable desde una perspectiva ética y de transparencia, sino que es fundamental para la apropiación tecnológica en entornos rurales.

La evidencia empírica indica que los pequeños productores tienden a desconfiar de las herramientas digitales cuando sus resultados no pueden ser comprendidos o contextualizados, además, al permitir que el usuario entienda por qué el sistema sugiere una determinada acción, se potencia su autonomía decisional, se reduce la resistencia al cambio tecnológico y se consolida un proceso de aprendizaje colectivo basado en la experiencia local y la ciencia de datos (De la Peña & Granados, 2024).

En el plano técnico, el uso de modelos multimodales representa un avance sustantivo frente a las metodologías convencionales. Mientras los modelos univariantes o los algoritmos basados en regresión simple suelen fallar en contextos de alta variabilidad agroclimática, los modelos multimodales tienen la capacidad de capturar sinergias y compensaciones entre variables, identificar patrones ocultos y ofrecer predicciones más robustas y precisas (Logacjov, 2024). Este tipo de arquitectura permite, por ejemplo, anticipar una disminución de rendimiento en función de una combinación específica de estrés hídrico, condiciones edáficas y señales visuales incipientes de enfermedades, algo prácticamente inalcanzable mediante el análisis humano tradicional.



No obstante, la implementación efectiva de este tipo de sistemas también enfrenta desafíos significativos. Por un lado, existe una marcada asimetría en el acceso a datos de calidad, especialmente en territorios donde no hay infraestructura meteorológica densa ni cartografía edáfica actualizada. Por otro, persiste la dificultad de traducir modelos complejos en aplicaciones que sean funcionales en dispositivos móviles básicos y entornos de baja conectividad. Para superar estas barreras, es clave la adopción de infraestructuras digitales distribuidas, el entrenamiento de modelos ligeros y la generación de entornos híbridos online/offline que aseguren operatividad en campo (Pinardi et al., 2023).

Finalmente, es necesario destacar que el valor del sistema PREDICT-CACAO no radica exclusivamente en su capacidad técnica, sino en su potencial para catalizar un nuevo paradigma de gestión agroecológica basada en datos, donde el conocimiento científico se articula con el saber campesino, y donde la inteligencia artificial no reemplaza la experiencia del productor, sino que la potencia y la sistematiza. En este sentido, la discusión sobre este tipo de tecnologías debe trascender la eficacia algorítmica y avanzar hacia una reflexión ética, participativa y política sobre el futuro del desarrollo rural en los trópicos.

## Trabajos futuros

El presente trabajo, si bien propone una arquitectura metodológica sólida y un conjunto de componentes tecnológicos avanzados para la optimización de la producción de cacao mediante aprendizaje automático multimodal, constituye una etapa inicial dentro de un proceso más amplio y evolutivo de transformación digital del sector cacaotero tropical. Por tanto, se vislumbran múltiples líneas de desarrollo y profundización que podrán abordarse en futuras investigaciones, pilotos ampliados y colaboraciones interdisciplinarias.

Una primera línea de trabajo consiste en la ampliación geográfica y agroecológica del sistema PREDICT-CACAO, con el objetivo de adaptar y validar sus modelos en otros territorios con condiciones contrastantes, sin embargo, si bien este proyecto se centra en zonas tropicales de América Latina, su estructura permite escalar el enfoque hacia regiones cacaoteras de África Occidental o Asia Sudoriental, donde se enfrentan desafíos similares en términos de enfermedades, variabilidad climática, baja productividad y limitada digitalización, además, la expansión internacional del modelo requerirá ajustes lingüísticos, culturales y técnicos, así como la incorporación de bases de datos regionales y alianzas con instituciones locales de investigación agraria (Masolele et al., 2024).

Una segunda línea de trabajo está asociada al mejoramiento continuo de los modelos de aprendizaje automático mediante aprendizaje federado y activo. A medida que se desplieguen versiones funcionales del sistema en campo, será posible incorporar mecanismos de retroalimentación directa que permitan que el modelo aprenda y se actualice a partir de los datos que generan los propios usuarios. Por otro lado, este enfoque de ML continuo no solo aumentará la precisión del sistema con el tiempo, sino que también democratizará la mejora del modelo, ya que cada agricultor contribuirá, de forma anónima y segura, al fortalecimiento colectivo de la herramienta. Asimismo, el uso de aprendizaje federado podría reducir la



necesidad de centralizar datos sensibles, preservando la privacidad de las fincas y facilitando la adopción del sistema en comunidades con restricciones sobre el uso de datos personales. Un tercer eje de desarrollo futuro es la integración de sensores IoT de bajo costo, como sensores de humedad de suelo, estaciones meteorológicas modulares o dispositivos de seguimiento de crecimiento vegetal, además, la combinación de estos sensores con el sistema actual permitiría una mayor granularidad temporal en la toma de decisiones, incorporando datos en tiempo real y generando alertas automatizadas ante condiciones críticas. Esta sinergia entre el modelo de predicción y los datos sensoriales abriría la posibilidad de construir una red inteligente de monitoreo cacaotero, conectando múltiples fincas y generando inteligencia colectiva sobre el comportamiento del cultivo a escala territorial (Riaño et al., 2023).

Otra dimensión relevante para el futuro del proyecto es el enriquecimiento semántico del sistema a través de ontologías agrícolas adaptadas al cultivo de cacao. Incorporar una ontología específica permitiría estructurar mejor el conocimiento agronómico, vincular variables con prácticas recomendadas y facilitar la interoperabilidad con otras plataformas de gestión agrícola. Esta capa semántica no solo favorecería la precisión y transparencia del sistema, sino que también permitiría conectar bases de datos heterogéneas bajo un mismo lenguaje técnico, facilitando el desarrollo de sistemas expertos y agentes conversacionales especializados en cacao.

Desde una perspectiva sociotecnológica, uno de los trabajos futuros más importantes será profundizar en el impacto social, económico y cultural de la introducción de IA en comunidades productoras de cacao, también, será clave estudiar cómo estas tecnologías transforman las prácticas agrarias, las relaciones entre agricultores y técnicos, las dinámicas de género y juventud, así como los procesos de toma de decisiones a nivel familiar y colectivo. Este análisis requerirá metodologías cualitativas complementarias, como etnografía digital, investigación acción participativa y estudios de adopción tecnológica con enfoque de derechos digitales rurales (Pinardi et al., 2023).

Finalmente, en el largo plazo, se vislumbra la posibilidad de integrar el sistema con cadenas de trazabilidad agroalimentaria y plataformas de certificación sostenible, de manera que los datos generados por PREDICT-CACAO no solo sirvan para optimizar la producción, sino también para fortalecer la transparencia comercial, garantizar prácticas agroecológicas y facilitar el acceso a mercados diferenciados.



## Referencias bibliográficas

- Afsar, M. M., Bakhshi, A. D., Hussain, E., & Iqbal, J. (2024). A deep learning-based framework for object recognition in ecological environments with dense focal loss and occlusion. *Neural Computing and Applications*, 36(16), 9591–9604. <https://doi.org/10.1007/s00521-024-09582-5>
- Agbotui, D. K., Ingold, M., Wiegle, M., & Buerkert, A. (2023). Can carbon payments improve profitability of traditional conventional and organic cocoa agroforests? A case study in the Eastern Region of Ghana. *Agroforestry Systems*, 97(5), 813–831. <https://doi.org/10.1007/s10457-023-00828-0>
- Akpoti, K., Dembélé, M., Forkuor, G., Obuobie, E., Mabhaudhi, T., & Cofie, O. (2023). Integrating GIS and remote sensing for land use/land cover mapping and groundwater potential assessment for climate-smart cocoa irrigation in Ghana. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-43286-5>
- Alexandre, C., Tresch, L., Sarron, J., Lavarenne, J., Bringer, G., Chaham, H. R., Bendahou, H., Carmeni, S., Borianne, P., Koffi, J. M., & Faye, E. (2023). Creating shared value(s) from On-Farm Experimentation: ten key lessons learned from the development of the SoYield® digital solution in Africa. *Agronomy for Sustainable Development*, 43(3). <https://doi.org/10.1007/s13593-023-00888-7>
- Alvarado, M. C., Sanchez, P. D. C., & Polongasa, S. G. N. (2023). Emerging rapid and non-destructive techniques for quality and safety evaluation of cacao: recent advances, challenges, and future trends. In *Food Production, Processing and Nutrition* (Vol. 5, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s43014-023-00157-w>
- Arakeri, M., Dhatvik, M. P., Kavan, A. V., Murthy, K. S., Nishitha, N. L., & Lakshmi, N. (2024). Intelligent pesticide recommendation system for cocoa plant using computer vision and deep learning techniques. *Environmental Research Communications*, 6(7). <https://doi.org/10.1088/2515-7620/ad58ae>
- Ashiagbor, G., Asare-Ansah, A. O., Amoah, E. B., Asante, W. A., & Mensah, Y. A. (2023). Assessment of machine learning classifiers in mapping the cocoa-forest mosaic landscape of Ghana. *Scientific African*, 20. <https://doi.org/10.1016/j.sciaf.2023.e01718>
- Atalaya-Marin, N., Goñas, M., Tineo, D., Chuquibala-Checan, B., Arce-Inga, M., Tarrillo, E., Alvarez-Robledo, Y. A., Tafur-Culqui, J., Cabrera-Hoyos, H., & Gómez-Fernández, D. (2025). Integrating remote sensing and in-situ data to determine climate diversity and variability in cocoa systems in the provinces of Jaén and San Ignacio, Cajamarca (NW Perú). *Trees, Forests and People*, 19. <https://doi.org/10.1016/j.tfp.2024.100749>





Attiogb  , A. A. C., Abotsi, K. E., Adjossou, K., Parkoo, E. N., Adjonou, K., & Kokou, K. (2022). Climate vulnerability of coffee-cocoa agrosystems in the sub-humid mountain ecosystems in south-west Togo (West Africa). Environmental Systems Research, 11(1). <https://doi.org/10.1186/s40068-022-00274-4>

Ayikpa, K. J., Ballo, A. B., Mamadou, D., & Gouton, P. (2024). Optimization of Cocoa Pods Maturity Classification Using Stacking and Voting with Ensemble Learning Methods in RGB and LAB Spaces. *Journal of Imaging*, 10(12). <https://doi.org/10.3390/jimaging10120327>

Bosompem, M. (2021). Potential challenges to precision agriculture technologies development in Ghana: scientists' and cocoa extension agents' perspectives. *Precision Agriculture*, 22(5), 1578–1600. <https://doi.org/10.1007/s11119-021-09801-2>

Carrillo, K., Martínez, M., Ramírez, L., Argüello, D., & Chavez, E. (2023). Cadmium (Cd) distribution and soil-plant relationship in cacao farms in Costa Rica. Environmental Monitoring and Assessment, 195(10). <https://doi.org/10.1007/s10661-023-11817-2>

Cerón-Souza, I., Delgadillo-Duran, D., Polo-Murcia, S. M., Sarmiento-Naizque, Z. X., & Reyes-Herrera, P. H. (2023). Prioritizing Colombian plant genetic resources for investment in research using indicators about the geographic origin, vulnerability status, economic benefits, and food security importance. *Biodiversity and Conservation*, 32(7), 2221–2261.  
<https://doi.org/10.1007/s10531-023-02599-7>

Chapman, R., Cock, J., Samson, M., Janetski, N., Janetski, K., Gusyana, D., Dutta, S., & Oberthür, T. (2021). Crop response to El Niño-Southern Oscillation related weather variation to help farmers manage their crops. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-87520-4>

De la Peña, N., & Granados, O. M. (2024). Artificial intelligence solutions to reduce information asymmetry for Colombian cocoa small-scale farmers. *Information Processing in Agriculture*, 11(3), 310–324. <https://doi.org/10.1016/j.inpa.2023.03.001>

Deldari, S., Xue, H., Saeed, A., Smith, D. V., & Salim, F. D. (2022). COCOA: Cross Modality Contrastive Learning for Sensor Data. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 6(3). <https://doi.org/10.1145/3550316>

Donfouet, O., & Ngouhouo, I. (2024). Impact of artificial intelligence on the total productivity of agricultural factors in Africa. Environment, Development and Sustainability. <https://doi.org/10.1007/s10668-024-05528-y>

Dung, C. D., Trueman, S. J., Wallace, H. M., Farrar, M. B., Gama, T., Tahmasbian, I., & Bai, S. H. (2023). Hyperspectral imaging for estimating leaf, flower, and fruit macronutrient concentrations and predicting strawberry yields. *Environmental Science and Pollution Research*, 30(53), 114166–114182. <https://doi.org/10.1007/s11356-023-30344-8>





González-Orozco, C. E., Porcel, M., Byrareddy, V. M., Rahn, E., Cardona, W. A., Velandia, D. A. S., Araujo-Carrillo, G. A., & Kath, J. (2024). Preparing Colombian coffee production for climate change: Integrated spatial modelling to identify potential robusta coffee (*Coffea canephora* P.) growing areas. *Climatic Change*, 177(4). <https://doi.org/10.1007/s10584-024-03717-2>

Harvey, C. A., Pritts, A. A., Zwetsloot, M. J., Jansen, K., Pulleman, M. M., Armbrecht, I., Avelino, J., Barrera, J. F., Bunn, C., Hoyos García, J., Isaza, C., Munoz-Ucros, J., Pérez-Alemán, C. J., Rahn, E., Robiglio, V., Somarriba, E., Valencia, V., & Ni, V. V. (n.d.). Transformation of coffee-growing landscapes across Latin America. A review. <https://doi.org/10.1007/s13593-021-00712-0/Published>

Indore, N. S., Karunakaran, C., & Jayas, D. S. (2022). Synchrotron tomography applications in agriculture and food sciences research: a review. In *Plant Methods* (Vol. 18, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s13007-022-00932-9>

Isong, I. A., John, K., Okon, P. B., Ogban, P. I., & Afu, S. M. (2022). Soil quality estimation using environmental covariates and predictive models: an example from tropical soils of Nigeria. *Ecological Processes*, 11(1). <https://doi.org/10.1186/s13717-022-00411-y>  
 Kalischek, N., Lang, N., Renier, C., Daudt, R. C., Addoah, T., Thompson, W., Blaser-Hart, W. J., Garrett, R., Schindler, K., & Wegner, J. D. (2023). Cocoa plantations are associated with deforestation in Côte d'Ivoire and Ghana. *Nature Food*, 4(5), 384–393. <https://doi.org/10.1038/s43016-023-00751-8>

Kanmegne Tamga, D., Latifi, H., Ullmann, T., Baumhauer, R., Thiel, M., & Bayala, J. (2023). Modelling the spatial distribution of the classification error of remote sensing data in cocoa agroforestry systems. *Agroforestry Systems*, 97(1), 109–119. <https://doi.org/10.1007/s10457-022-00791-2>

Kulesza, E., Thomas, P., Prewitt, S. F., Shalit-Kaneh, A., Wafula, E., Knollenberg, B., Winters, N., Esteban, E., Pasha, A., Provart, N., Praul, C., Landherr, L., dePamphilis, C., Maximova, S. N., & Guiltinan, M. J. (2024). The cacao gene atlas: a transcriptome developmental atlas reveals highly tissue-specific and dynamically-regulated gene networks in *Theobroma cacao* L. *BMC Plant Biology*, 24(1). <https://doi.org/10.1186/s12870-024-05171-9>

Lara-Estrada, L., Rasche, L., & Schneider, U. A. (n.d.). Land in Central America will become less suitable for coffee cultivation under climate change. <https://doi.org/10.1007/s10113-021-01803-0/Published>

Lesiv, M., Schepaschenko, D., Buchhorn, M., See, L., Dürauer, M., Georgieva, I., Jung, M., Hofhansl, F., Schulze, K., Bilous, A., Blyshchyk, V., Mukhortova, L., Brenes, C. L. M., Krivobokov, L., Ntie, S., Tsogt, K., Pietsch, S. A., Tikhonova, E., Kim, M., ... Fritz, S. (2022). Global forest management data for 2015 at a 100 m resolution. *Scientific Data*, 9(1). <https://doi.org/10.1038/s41597-022-01332-3>



Lie Cabrestante, J. M., Difuntorum, A. V., & Gino Wata, M. (2024). Disease Identification and Severity Analysis on Lanzones Leaf Using Android-Based Application. <https://doi.org/10.1145/3647649>

Logacjov, A. (2024). Self-supervised Learning for Accelerometer-based Human Activity Recognition: A Survey. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 8(4). <https://doi.org/10.1145/3699767>

Longchamps, L., Tisseyre, B., Taylor, J., Sagoo, L., Momin, A., Fountas, S., Manfrini, L., Ampatzidis, Y., Schueller, J. K., & Khosla, R. (2022). Yield sensing technologies for perennial and annual horticultural crops: a review. In Precision Agriculture (Vol. 23, Issue 6, pp. 2407–2448). Springer. <https://doi.org/10.1007/s11119-022-09906-2>

Lopes, J. F., da Costa, V. G. T., Barbin, D. F., Cruz-Tirado, L. J. P., Baeten, V., & Barbon Junior, S. (2022). Deep computer vision system for cocoa classification. Multimedia Tools and Applications, 81(28), 41059–41077. <https://doi.org/10.1007/s11042-022-13097-3>  
Marcel Heming, N., Schroth, G., Talora, D. C., & Faria, D. (n.d.). Cabruca agroforestry systems reduce vulnerability of cacao plantations to climate change in southern Bahia. <https://doi.org/10.1007/s13593-022-00780-w/Published>

Masolele, R. N., Marcos, D., De Sy, V., Abu, I. O., Verbesselt, J., Reiche, J., & Herold, M. (2024). Mapping the diversity of land uses following deforestation across Africa. Scientific Reports, 14(1). <https://doi.org/10.1038/s41598-024-52138-9>

Obando, Ó. E. T., Escobar, J. A. D., & Torres, E. A. (2023). Socio-economic aspects associated with water management in cocoa cultivation in Valle del Cauca, Colombia. Acta Agronomica, 72(4), 368–376. <https://doi.org/10.15446/acag.v72n4.113615>

Omeka, M. E., Igwe, O., Onwuka, O. S., Nwodo, O. M., Ugar, S. I., Undiandeye, P. A., & Anyanwu, I. E. (2024). Efficacy of GIS-based AHP and data-driven intelligent machine learning algorithms for irrigation water quality prediction in an agricultural-mine district within the Lower Benue Trough, Nigeria. Environmental Science and Pollution Research, 31(41), 54204–54233. <https://doi.org/10.1007/s11356-023-25291-3>

Orji, R., Deters, R., Kumi, S., & Lomotey, R. K. (2024). Automatic detection and diagnosis of cocoa diseases using mobile tech and deep learning. International Journal of Sustainable Agricultural Management and Informatics, 10(1). <https://doi.org/10.1504/ijسامی.2024.10059217>

Pinardi, S., Salis, M., Sartor, G., & Meo, R. (2023). EU–Africa: Digital and Social Questions in a Multicultural Agroecological Transition for the Cocoa Production in Africa. Social Sciences, 12(7). <https://doi.org/10.3390/socsci12070398>

Ramos, M. J., Beilhe, L. B., Alvarado, J., Rapidel, B., & Allinne, C. (2024). Disentangling shade effects for cacao pest and disease regulation in the Peruvian Amazonia. Agronomy for Sustainable Development, 44(1). <https://doi.org/10.1007/s13593-024-00948-6>



Riaño, M. A., Rodriguez, A. O. R., Velandia, J. B., García, P. A. G., & Marín, C. E. M. (2023). Design and application of an ontology to identify crop areas and improve land use. *Acta Geophysica*, 71(3), 1409–1426. <https://doi.org/10.1007/s11600-022-00808-5>

Schmidt, J. E., DuVal, A., Isaac, M. E., & Hohmann, P. (2022). At the roots of chocolate: understanding and optimizing the cacao root-associated microbiome for ecosystem services. A review. In *Agronomy for Sustainable Development* (Vol. 42, Issue 2). Springer-Verlag Italia s.r.l. <https://doi.org/10.1007/s13593-021-00748-2>

Şener, A., & Ergen, B. (2024). Advanced CNN Approach for Segmentation of Diseased Areas in Plant Images. *Journal of Crop Health*, 76(6), 1569–1583. <https://doi.org/10.1007/s10343-024-01054-z>

Seyum, E. G., Bille, N. H., Abtew, W. G., Munyengwa, N., Bell, J. M., & Cros, D. (2022). Genomic selection in tropical perennial crops and plantation trees: a review. In *Molecular Breeding* (Vol. 42, Issue 10). Springer Science and Business Media B.V. <https://doi.org/10.1007/s11032-022-01326-4>

Singh, K., Fuentes, I., Al-Shammari, D., Fidelis, C., Butubu, J., Yinil, D., Sharififar, A., Minasny, B., Guest, D. I., & Field, D. J. (2023). E-Agriculture Planning Tool for Supporting Smallholder Cocoa Intensification Using Remotely Sensed Data. *Remote Sensing*, 15(14). <https://doi.org/10.3390/rs15143492>

Somarriba, E., Saj, S., Orozco-Aguilar, L., Somarriba, A., & Rapidel, B. (2024). Shade canopy density variables in cocoa and coffee agroforestry systems. *Agroforestry Systems*, 98(3), 585–601. <https://doi.org/10.1007/s10457-023-00931-2>

Talero-Sarmiento, L., Roa-Prada, S., Caicedo-Chacon, L., & Gavanzo-Cardenas, O. (2025). A Data-Driven Approach to Improve Cocoa Crop Establishment in Colombia: Insights and Agricultural Practice Recommendations from an Ensemble Machine Learning Model. *AgriEngineering*, 7(1). <https://doi.org/10.3390/agriengineering7010006>

Vásquez-Ordóñez, A. A., Torres-López, W., & Monmany-Garzia, A. C. (2024). A multi-scale approach to study palm-weevils in a tropical agroecosystem. *Landscape Ecology*, 39(3). <https://doi.org/10.1007/s10980-024-01796-1>

Vera, D. B., Oviedo, B., Casanova, W. C., & Zambrano-Vega, C. (2024). Deep Learning-Based Computational Model for Disease Identification in Cocoa Pods (*Theobroma cacao* L.). <http://arxiv.org/abs/2401.01247>

Villalobos-Culqui, C., García-Rivas-plata, C., & Tuesta-Hidalgo, O. A. (2025). Artificial vision model based on convolutional neural networks for black pod identification in cacao plantations. *Revista Cientifica de Sistemas e Informatica*, 5(1). <https://doi.org/10.51252/rksi.v5i1.678>

Zhang-Zheng, H., Deng, X., Aguirre-Gutiérrez, J., Stocker, B. D., Thomson, E., Ding, R., Adu-Bredou, S., Duah-Gyamfi, A., Gvozdevaite, A., Moore, S., Oliveras Menor, I., Prentice,



I. C., & Malhi, Y. (2024). Why models underestimate West African tropical forest primary productivity. *Nature Communications*, 15(1). <https://doi.org/10.1038/s41467-024-53949-0>

**Conflictos de intereses:**

Los autores declaran que no existe conflicto de interés posible.

**Financiamiento:**

No existió asistencia financiera de partes externas al presente artículo.

**Agradecimiento:**

N/A

**Nota:**

El artículo no es producto de una publicación anterior.

