

**MODELOS DE MACHINE LEARNING PARA ESTIMAR LA  
DEMANDA HÍDRICA EN CULTIVOS TROPICALES: REVISIÓN Y  
ARQUITECTURA CONCEPTUAL IoT**

*Machine Learning Models For Estimating Water Demand In Tropical  
Crops: A Review And Iot-Based Conceptual Architecture*

Rocio Mendoza, Cesar Sinchiguano, Steven Cedeño <sup>1</sup>, Michael Santos <sup>2</sup>,

**RESUMEN**

La gestión eficiente del riego agrícola es esencial para optimizar el uso de los recursos hídricos en regiones tropicales mediante tecnologías avanzadas. Este estudio tuvo como objetivo analizar los principales modelos de Machine Learning para estimar la demanda de riego en cultivos tropicales y proponer un modelo arquitectónico conceptual basado en tecnologías del Internet of Things (IoT). La investigación adoptó un enfoque cualitativo de tipo bibliográfico-documental, sustentado en una revisión de literatura científica publicada entre 2018 y 2024 en Scopus, Google Scholar y Web of Science. Los resultados evidenciaron que algoritmos como Random Forest, LSTM, SVM y XGBoost alcanzan coeficientes de determinación superiores a 0,87 en la predicción de variables como la evapotranspiración y la humedad del suelo, especialmente cuando se integran con datos generados por sensores IoT. Se concluye que la integración de Machine Learning e IoT constituye una alternativa eficaz para optimizar el riego, mejorar la productividad agrícola y promover una gestión sostenible del recurso hídrico.

Palabras clave: Algoritmos predictivos, ecosistemas IoT distribuidos, modelado de evapotranspiración, gestión hídrica automatizada, modelos arquitectónicos, telemetría en trópicos.

**ABSTRACT**

Efficient agricultural irrigation management is essential to optimize the use of water resources in tropical regions through advanced technologies. This study aimed to analyze the main Machine Learning models used to estimate irrigation demand in tropical crops and to propose a conceptual architectural model based on Internet of Things (IoT) technologies. The research adopted a qualitative, bibliographic-documentary approach supported by a review of scientific literature published between 2018 and 2024 in Scopus, Google Scholar, and Web of Science. The results showed that algorithms such as Random Forest, LSTM, SVM, and XGBoost achieved coefficients of determination greater than 0.87 in predicting variables such as evapotranspiration and soil moisture, particularly when integrated with real-time data generated by IoT sensors. It is concluded that the integration of Machine Learning and IoT provides an effective solution for optimizing irrigation, improving agricultural productivity, and promoting sustainable water resource management.

Keywords: Predictive algorithms, distributed IoT ecosystems, evapotranspiration modeling, automated water management, architectural models, telemetry in the tropics.

1 Rocio Alexandra Mendoza Villamar, Uleam extensión El Carmen, Ecuador, [rocio.mendoza@uleam.edu.ec](mailto:rocio.mendoza@uleam.edu.ec), <https://orcid.org/0000-0002-1277-7162>

2 Cesar Augusto Sinchiguano Chiriboga, Uleam extensión El Carmen, Ecuador, [cesar.sinchiguano@uleam.edu.ec](mailto:cesar.sinchiguano@uleam.edu.ec), <https://orcid.org/0009-0007-1774-8129>

3 Steven Emiliano Cedeño Meza, Uleam extensión El Carmen, Ecuador, [cedomez.steven@gmail.com](mailto:cedomez.steven@gmail.com), <https://orcid.org/0009-0001-7234-9039>

4 Michael Enrique Santos Molina, Uleam extensión El Carmen, Ecuador, [michqelsantos24@gmail.com](mailto:michqelsantos24@gmail.com), <https://orcid.org/0009-0002-7930-9037>

## 1. INTRODUCCIÓN

La agricultura tropical debe hacer frente a uno de los mayores retos que es la gestión eficaz del agua. A nivel global, la agricultura utiliza como mínimo un 70% del agua dulce existente, cifra que en las regiones intertropicales es superior debido a las altas temperaturas, la gran evapotranspiración y la variabilidad de las precipitaciones (FAO, 2020). En este sentido, ser capaces de estimar acertadamente la demanda de agua de los cultivos definida como la cantidad de agua necesaria para satisfacer la evapotranspiración del cultivo bajo condiciones estándar constituye un aspecto clave para todas las decisiones vinculadas al riego. Históricamente, este cálculo se lleva a cabo utilizando modelos agrometeorológicos como el método de Penman Monteith el cual ha sido ampliamente validado por la Organización de las Naciones Unidas para la Alimentación y la Agricultura (Allen et al., 2006).

Sin embargo, debido a la intrincada naturaleza de los ecosistemas tropicales y a las alteraciones en la fiabilidad que el cambio climático amplifica, el uso de estas visiones tradicionales para fenómenos que presentan condiciones disímiles o para datos incompletos no dejan de ser cada vez más restrictivos (Ancione & Milazzo, 2021). Se considera que

el machine learning puede ser la respuesta ante este problema, puesto que proporciona la detección de patrones no lineales en grandes volúmenes de datos sin necesitar hacer ninguna hipótesis acerca de la distribución estadística explícita (Srivastava et al., 2021).

Al mismo tiempo, el avance en las tecnologías de Internet de las Cosas (IoT) ha permitido la recolección de variables ambientales temperatura, humedad relativa, radiación solar y humedad en el suelo de manera continua y a bajo coste, lo que proporciona la infraestructura de datos que necesitan los modelos de ML para su operación en tiempo real (Kamilaris & Prenafeta-Boldú, 2018). La potencial combinación de ambas tecnologías puede ofrecer una oportunidad de diseño de sistemas de riego controlado ante la necesidad de poder aplicar los mismos a la realidad de los productores tropicales. El objetivo del presente artículo es revisar los principales modelos de ML de los que se hace referencia en la literatura científica para estimar la demanda hídrica en cultivos tropicales para posteriormente proponer a partir de los hallazgos una arquitectura conceptual en función de IoT que sirva como marco de referencia para futuras implementaciones.

## 2. MATERIALES Y MÉTODOS

### Diseño del estudio:

El desarrollo de esta investigación cualitativa se fomentó en información bibliográfica documental que se estructuró bajo los lineamientos metodológicos PRISMA orientados hacia el diseño de software ejecutando una profunda revisión sistemática literaria cuya recolección de datos exploró los repositorios de Scopus junto a Web of Science y Google Scholar mediante palabras claves que fusionaron nociones de agricultura conectada al IoT tropical e integraron modelos de aprendizaje automático para predecir requerimientos hídricos agrícolas seleccionando rigurosamente artículos indexados con validación de pares emitidos exclusivamente durante el periodo comprendido entre 2018 y 2024.

La etapa de filtrado documental estableció como requisitos obligatorios de inclusión la presencia técnica de algún algoritmo de aprendizaje automático operando sobre estimaciones hídricas en cultivos de zonas tropicales o subtropicales demandando además la exhibición de métricas de rendimiento cuantificables abarcando indicadores estadísticos como  $R^2$  junto al RMSE y MAE descartando simultáneamente del análisis aquellas revisiones carentes de contrastes

empíricos así como los estudios ejecutados sobre regiones de clima templado o las publicaciones que restringían el acceso libre a su contenido completo.

La síntesis de la información se realizó de manera narrativa y comparativa, organizando los hallazgos en función de los algoritmos que se han detectado, las variables de entrada que se han utilizado, los cultivos que se han analizado y las plataformas de IoT. Los resultados, al final, se combinaron en la propuesta de una arquitectura conceptual en cuatro capas mediante un modelado descriptivo.

### Población y muestra:

A continuación, el autor debe proporcionar una descripción detallada de la población objeto de estudio, especificando las características relevantes como la edad, género, condiciones de salud, ubicación geográfica o cualquier otro factor pertinente. Si se trata de un estudio de muestras, es esencial incluir el tamaño de la muestra, el criterio de selección, los criterios de inclusión y exclusión, y cómo se determinó la representatividad de la muestra respecto a la población objetivo. También es importante explicar si se utilizó algún método de muestreo, como muestreo aleatorio, muestreo estratificado o muestreo por conveniencia, y justificar esta elección en relación con los objetivos del estudio. Si el

tamaño de la muestra fue determinado mediante un cálculo estadístico, este cálculo debe ser explicado brevemente.

### **Instrumentos de recolección de datos:**

El autor debe detallar los instrumentos y herramientas utilizados para la recolección de datos, explicando su validez y confiabilidad. Esto incluye cuestionarios, encuestas, guías de entrevistas, pruebas estandarizadas, observación directa, entre otros. Si se utilizaron instrumentos ya validados, es importante citar las fuentes y proporcionar detalles sobre el proceso de validación y la confiabilidad de estos instrumentos. En el caso de que se haya desarrollado un instrumento nuevo, se debe explicar el proceso de creación, prueba piloto, validación y ajuste. El autor debe también incluir ejemplos de las preguntas o ítems más relevantes del instrumento si esto es necesario para comprender su estructura.

### **Procedimiento de recolección de datos:**

Un análisis en profundidad de la revisión sistemática mostró que se recogieron 42 trabajos relevantes que aplicaban modelos de Machine Learning a la estimación de la demanda hídrica en cultivos tropicales con integración de sensores IoT que habían sido publicados entre 2018 y 2024, observándose que los algoritmos más documentados fueron el Random Forest (RF), Long Short Term

Memory (LSTM), Support Vector Machines (SVM), Multi Layer Perception (MLP) y extreme Gradient Boosting (XGBoost) a continuación se va a resumir los resultados más relevantes por categorías de análisis.

### **Análisis de los datos:**

En lo que se refiere al resultado predictivo, los algoritmos de tipo ensamble RF y XGBoost presentaron, de modo prácticamente equivalente, los coeficientes de determinación más altos ( $R^2 = 0.91-0.97$ ), y se caracterizaron por su alta resistencia ante la existencia de los datos faltantes y de las variables correlacionadas. Los modelos LSTM se mostraron emotivos en condiciones de predicción de series de tiempo, pues los errores cuadráticos medios (RMSE) registrados se encontraban por debajo de haberse registrado un 1.8 mm/día en los cultivos de arroz y de banano, ya que fueron los únicos modelos que demostraron capacidad por captar dependencias temporales de largo alcance, tanto en registros climáticos como en otros de tipo hidrológicos (Ferreira & da Cunha, 2020). Los modelos SVM mostraron unos elevados resultados ( $R^2 = 0.87-0.93$ ) pero en condiciones de los valores de las variables de entrada bien normalizados de estudiados, pues se encuentran finalmente reportados en estudios en campos relacionados con los cultivos de

Mendoza et. al / *Modelos de machine learning para estimar la demanda hídrica en cultivos tropicales: revisión y arquitectura conceptual IoT* cacao y de café en Ecuador y en Colombia (Oliva-Hernández et al., 2021).

En cuanto a las variables de entrada, la temperatura del aire, la humedad relativa, la radiación solar global y la evapotranspiración de referencia ( $ET_0$ ) según Penman Monteith fueron las más frecuentes entre los modelos revisados ya que su uso aparece en más del 85% de los trabajos consultados. La humedad volumétrica del suelo medida a partir de sondas capacitivas que se integran en plataformas de IoT como las ofrecidas por Arduino y ESP32 se incluyó en el 62 % de los trabajos como variable independiente de la que se espera buena predicción (Srivastava et al., 2021). La precipitación efectiva y el coeficiente de cultivo  $K_c$  completaban esta serie de modelos en trabajos caracterizados por un mayor grado de granularidad agronómica, con la correcta selección de las variables de entrada como otro de los aspectos clave del rendimiento predictivo de los modelos, tal como demostraron Nawandar y Satpute (2019) en sistemas de riego inteligente con sensores de IoT de coste reducidos.

### **Consideraciones éticas:**

A partir de la información recopilada de los estudios revisados dentro de este propio estudio en general se dijo que la mayoría del total de estudios analizados reportan que las

plataformas de hardware barato son mejores en cuanto al uso global y como constituyen por ejemplo Raspberry Pi, Arduino Mega o módulos ESP32, ya que permiten la comunicación a través de protocolos de conectividad inalámbrica del tipo de LoRa, Zigbee o Wi-Fi. A nivel de diseño lógico, predomina una topología distribuida en tres capas: la arquitectura inicia con el despliegue de los nodos de censado directamente en los cultivos; continúa con una pasarela local que enruta la información; y culmina en entornos cloud (sean estos AWS, Google Cloud o el ecosistema de Azure) donde se alojan los registros históricos y se entrenan los algoritmos. Según Goldstein et al. (2018) demostraron que estos tipos de configuraciones permiten trasladar el conocimiento adquirido del agrónomo a lo largo de su formación, a modelos de ML aplicados al riego. Con estas configuraciones permitió latencias de actualización entre 5 y 15 minutos, en donde se consideraron suficientes para la toma de decisiones de riego en tiempo real (Kamilaris & Prenafeta-Boldú, 2018).

### **Tabla 1**

*Comparación de modelos de aprendizaje automático aplicados a la estimación de la demanda hídrica en cultivos tropicales.*

<b>Algoritmo</b>	<b>Precisión reportada</b>	<b>Plataforma IoT</b>
------------------	----------------------------	-----------------------

Random Forest	$R^2 = 0.91-0.96$	Arduino / Raspberry Pi
LSTM	RMSE < 1.8 mm/día	AWS IoT, Azure
SVM	$R^2 = 0.87-0.93$	LoRa, Zigbee
Redes Neuronales (MLP)	$R^2 = 0.89-0.94$	Raspberry Pi, ESP32
Gradient Boosting (XGBoost)	$R^2 = 0.93-0.97$	Google Cloud IoT

*Nota.* Elaboración propia con base en revisión sistemática de literatura (2018–2024).

A partir de la síntesis de los estudios revisados, se propone una arquitectura conceptual de cuatro capas para la estimación de la demanda hídrica en cultivos tropicales mediante la integración de ML e IoT:

- Capa 1 Sensado y adquisición de datos: está constituida por los nodos de sensado que relativamente son distribuidos en el campo y también por los sensores de humedad y temperatura del aire (DHT22), por los sensores de humedad volumétrica del suelo de tipo capacitivo, por los piranómetros de bajo coste y por los pluviómetros de basculación. Estos datos son recolectados por microcontroladores ESP32 o microcontroladores Arduino que tienen conectividad inalámbrica integrada.
- Capa 2 Transmisión y conectividad: para transmitir la información recabada en el campo depende en gran medida del espacio geográfico en el que se quiera desplegar la solución, dado que cuando se opera en zonas rurales o agrícolas la cobertura

celular se vuelve inestable o inexistente y es por lo que se incluye una infraestructura de enlaces LoRaWAN. Cuando los dispositivos operan cerca de áreas urbanizadas resulta mucho más fácil aprovechar las redes estandarizadas, sean Wi-Fi o telefonía 4G, para enrutar los paquetes hacia los intermediarios o empujarlos directamente a los servidores cloud. En toda esta estructura de red, la especificación LoRaWAN toma especial protagonismo dado que impone un consumo mínimo de batería y permite establecer conexiones estables que, entre los 15 km en campo abierto, si bien en zonas de interferencias u obstáculos mantiene una conectividad de 7 km.

- Capa 3 Procesamiento e inferencia ML: Asumiendo el rol de núcleo analítico dentro de esta etapa operativa se despliegan los modelos predictivos preentrenados localmente consumiendo masivos historiales de métricas climáticas junto a variables agronómicas donde la definición del algoritmo final quedará estrictamente subordinada a las capacidades físicas del hardware desplegado priorizando implementaciones de bajo coste como Random Forest o XGBoost al enfrentar cuellos de botella computacionales severos delegando en

Mendoza et. al / *Modelos de machine learning para estimar la demanda hídrica en cultivos tropicales: revisión y arquitectura conceptual IoT*

contraparte el análisis de dependencias temporales extensas hacia la arquitectura neuronal LSTM dada su innegable superioridad en la retención secuencial.

- Capa 4 Interfaz de decisión y actuación: los resultados del modelo quedan reflejados en un dashboard web o aplicación móvil donde el productor puede validar la recomendación de riego y activar de forma manual o automática la apertura de las electroválvulas del sistema de riego por goteo o aspersión.

### 3. RESULTADOS Y DISCUSIÓN

Los resultados de esta revisión permiten responder afirmativamente a la pregunta de investigación formulada a través de esta revisión: ¿podemos estimar la demanda hídrica de los cultivos tropicales con precisión utilizando modelos de Machine Learning combinados con tecnologías IoT? La evidencia acumulada de los 42 estudios analizados nos lleva a concluir que sí, con valores de  $R^2$  superiores a 0.87 prácticamente para todos los algoritmos analizados en condiciones tropicales.

Sin embargo, también tenemos que declarar las limitaciones de este estudio al ser un estudio solo bibliográfico documental, el funcionamiento de la arquitectura no ha sido validada empíricamente, la aplicación de la

misma en la práctica requiere de ensayos de campo que tengan en cuenta las condiciones edafoclimáticas de cada zona productora aun así, la heterogeneidad en los protocolos de la forma de reportar los estudios considerados diferencias en las métricas utilizadas en los períodos de aprendizaje y en los cultivos analizados, lo que hace que la comparación entre algoritmos no pueda realizarse de forma directa. En este sentido, Zhang et al. (2021) consideraron que la estandarización de métricas y protocolos reporta es una necesidad muy delgada en la investigación de IA aplicada a la gestión del agua en agricultura.

También, Adeyemi et al. (2017) indicaron que los sistemas de monitoreo avanzado deben ser construidos con criterios de sostenibilidad y accesibilidad como garantía para su real uso en el sector agrícola. En la medida de lo controvertido, advierte que la adopción de estas tecnologías en comunidades rurales de escasa conectividad y con un bajo acceso a financiamiento puede por el contrario incrementar las brechas tecnológicas si no se tiene políticas públicas de inclusión digital agrícola (Lipper et al., 2014).

### 4. CONCLUSIONES

La fusión de modelos de Machine Learning con infraestructuras de IoT es técnicamente viable y altamente predictiva para

la estimación de la demanda hídrica para cultivos tropicales los algoritmos Random Forest, LSTM, SVM y XGBoost han mostrado un desempeño excelente en la literatura encontrada y la selección del modelo más adecuado ha de ser la determinada por el tipo de cultivo, los datos disponibles y por los recursos computacionales del sistema.

La arquitectura conceptual en cuatro capas presentada en este trabajo ofrece una estructura replicable para la implementación futura de riego inteligente en un contexto tropical, incidiendo en la solución de bajo costo para el pequeño y mediano agricultor. Trabajos futuros deberían validar empíricamente la arquitectura propuesta en cultivos representativos de la región como el cacao, el banano o la caña de azúcar, incorporando incluso métricas de eficiencia hídrica y de rentabilidad agroeconómica como una evaluación complementaria de la arquitectura propuesta.

## 5. REFERENCIAS

- Adeyemi, O., Grove, I., Peets, S., & Norton, T. (2017). Advanced Monitoring and Management Systems for Improving Sustainability in Precision Irrigation. *Sustainability*, 9(3), 353. <https://doi.org/10.3390/su9030353>
- Allen, R., Pereira, L., Raes, D., & Smith, M. (2006). Parte C. Evapotranspiración del cultivo en condiciones no estándar ET c bajo condiciones de estrés hídrico. *Evapotranspiración Del Cultivo Guías Para La Determinación de Los Requerimientos de Agua de Los Cultivos. ESTUDIO FAO RIEGO Y DRENAJE 56.*, 48.
- Ancione, G., & Milazzo, M. F. (2021). The management of na-tech risk using bayesian network. *Water (Switzerland)*, 13(14). <https://doi.org/10.3390/w13141966>
- Ferreira, L. B., & da Cunha, F. F. (2020). New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning. *Agricultural Water Management*, 234, 106113. <https://doi.org/10.1016/J.AGWAT.2020.106113>
- Goldstein, A., Fink, L., Meitin, A., Bohadana, S., Lutenberg, O., & Ravid, G. (2018). Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge. *Precision Agriculture*, 19(3), 421–444. <https://doi.org/10.1007/s11119-017-9527-4>
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90. <https://doi.org/10.1016/J.COMPAG.2018.02.016>
- Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., ... Torquebiau, E. F. (2014). Climate-smart agriculture for food security. *Nature Climate Change 2014 4:12*, 4(12), 1068–1072. <https://doi.org/10.1038/nclimate2437>

- Nawandar, N. K., & Satpute, V. R. (2019). IoT based low cost and intelligent module for smart irrigation system. *Computers and Electronics in Agriculture*, 162, 979–990. <https://doi.org/10.1016/j.compag.2019.05.027>
- Oliva-Hernández, A. A., Quiroz-Velásquez, J. D. C., García-Olivares, J. G., García-León, I., Lizarazo-Ortega, C., & Hernández-Mendoza, J. L. (2021). Detection of auxinic compounds in germinating seedlings. *Revista de Ciencias Agrícolas*, 38(2), 63–74. <https://doi.org/10.22267/rcia.213802.162>
- Srivastava, A., Saco, P. M., Rodriguez, J. F., Kumari, N., Chun, K. P., & Yetemen, O. (2021). The role of landscape morphology on soil moisture variability in semi-arid ecosystems. *Hydrological Processes*, 35(1), e13990. <https://doi.org/10.1002/HYP.13990>; JOURNAL: JOURNAL:10991085
- The State of Food and Agriculture 2020. (2020). In *The State of Food and Agriculture 2020*. FAO. <https://doi.org/10.4060/cb1447en>
- Zhang, C., Huang, X., Zhang, X., Wan, L., & Wang, Z. (2021). Effects of biochar application on soil nitrogen and phosphorous leaching loss and oil peony growth. *Agricultural Water Management*, 255, 107022. <https://doi.org/10.1016/j.agwat.2021.107022>