

Predicción de la evapotranspiración de referencia mediante NASA-POWER: contraste con estaciones meteorológicas en el sudeste de Córdoba

Lucas Alberto Gusmerotti^{1,2}, Natalia Noemí Gattinoni¹, Carlos Marcelo Di Bella^{3,4} y Jorge Luis Mercau⁵

¹ Instituto de Clima y Agua, INTA. De Los Reseros y N. Repetto S/N (Hurlingham, Buenos Aires). gusmerotti.lucas@inta.gob.ar, gattinoni.natalia@inta.gob.ar

² Cátedra de Manejo y Conservación de Suelos, FCyF, UNLP.

³ Dpto. de Métodos Cuantitativos y Sistemas de Información (FAUBA). carlos.m.dibella@gmail.com

⁴ IFEVA (FAUBA, CONICET)

⁵ AER San Luis, INTA. mercau.jorge@inta.gob.ar

Resumen. La cuantificación de la evapotranspiración de referencia (ET_o) resulta de utilidad para diferentes aplicaciones. Sin embargo, su estimación está limitada por la carencia de instrumentos y redes de observación terrestres. Actualmente, se dispone de productos satelitales grillados de reanálisis y modelos de aprendizaje automático que surgen como alternativa para estimar variables meteorológicas espacialmente explícitas a escala local y regional. El objetivo de este trabajo fue evaluar la predicción de la ET_o mediante el producto NASA-POWER en el sudeste de Córdoba. A partir de la temperatura, humedad, velocidad de viento y radiación solar registradas en estaciones meteorológicas y estimadas por NASA-POWER, se estimó la ET_o mediante el modelo Penman-Monteith. Además, se empleó el modelo de aprendizaje automático *Extreme Gradient Boosting* (XGBoost) para corregir las estimaciones de NASA-POWER necesarias para el cálculo de ET_o. Los resultados mostraron que NASA-POWER estimó la ET_o con errores inferiores a 1 mm.día⁻¹ cuando se comparó con la observada en las estaciones. XGBoost mejoró sensiblemente la precisión. A partir del producto NASA-POWER y XGBoost se puede reconstruir la falta de registros meteorológicos en el sudeste de Córdoba y a partir de ello estimar la ET_o en forma precisa.

Palabras clave: datos de reanálisis, Penman-Monteith, aprendizaje automático

Received May 2024; Accepted June 2024; Published July 2024

Prediction of reference evapotranspiration using NASA-POWER: testing with weather stations in southeastern Córdoba

Lucas Alberto Gusmerotti^{1,2}, Natalia Noemí Gattinoni¹, Carlos Marcelo Di Bella^{3,4} y Jorge Luis Mercau⁵

¹ Instituto de Clima y Agua, INTA. De Los Reseros y N. Repetto S/N (Hurlingham, Buenos Aires). gusmerotti.lucas@inta.gob.ar, gattinoni.natalia@inta.gob.ar

² Cátedra de Manejo y Conservación de Suelos, FCyF, UNLP.

³ Dpto. de Métodos Cuantitativos y Sistemas de Información (FAUBA). carlos.m.dibella@gmail.com

⁴ IFEVA (FAUBA, CONICET)

⁵ AER San Luis, INTA. mercau.jorge@inta.gob.ar

Abstract. Computation of reference evapotranspiration (ET_o) is useful for different applications. However, its estimation is limited by the lack of terrestrial observation instruments and networks. Currently, gridded satellite reanalysis products and machine learning models are available as an alternative to estimate spatially explicit meteorological variables at local and regional scales. The aim of this work was to evaluate and predict ET_o using the NASA-POWER product in southeastern Córdoba. Through the Penman-Monteith model, ET_o was estimated from temperature, humidity, wind speed and solar radiation recorded at stations, those estimated by NASA-POWER and those predicted by the Extreme Gradient Boosting machine learning model (XGBoost). NASA-POWER estimated ET_o with errors less than 1 mm.day⁻¹ when compared to that observed at the stations. XGBoost obtained a better accuracy and lower estimation errors. From the NASA-POWER product and XGBoost it is possible to reconstruct the lack of meteorological records in southeastern Córdoba to estimate ET_o accurately.

Keywords: reanalysis data, Penman-Monteith, machine learning

1 Introducción

Después de la precipitación, la evapotranspiración (ET) es el componente más importante del ciclo terrestre del agua. A través de la evaporación desde la superficie y la transpiración de los vegetales, la ET devuelve a la atmósfera más del 60% de las lluvias [1], lo que modula la disponibilidad de agua en el suelo. La ET es también un importante flujo energético, ya que utiliza más de la mitad de la energía solar total absorbida por las superficies terrestres [2]. Una estimación precisa de la ET es necesaria no sólo debido a la creciente competencia por las limitadas reservas hídricas, sino que, dado que constituye una de las variables más afectadas por el cambio climático, también es esencial para proyectar los posibles cambios en el ciclo hidrológico mundial.

La gestión del agua en sistemas agrícolas y forrajeros requiere la estimación de las necesidades hídricas de los cultivos en condiciones estándar (ETc). La ETc se estima a partir de la evapotranspiración de una cobertura de referencia (ETo) y un coeficiente que varía con el tipo y el estado fenológico de un determinado cultivo (Kc) [3]. La ETo es, entonces, un componente fundamental en el cálculo del balance hídrico para el manejo de la producción vegetal, la realización de estudios ambientales, la programación del riego y la planificación integral de los recursos hídricos.

La ETo se define como la tasa de evaporación y transpiración de un cultivo de referencia hipotético, una superficie extensa de hierba verde de altura uniforme y en activo crecimiento, con un suministro adecuado de agua y que sombrea completamente al suelo [3]. Sobre la base de mediciones directas de la ETo con lisímetros en diferentes localidades, se han desarrollado diferentes modelos que permiten estimarla a partir de sus forzantes meteorológicas (*e.g.* temperatura, velocidad de viento, humedad relativa y/o radiación solar). La Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO) considera la ecuación Penman-Monteith (PM)-FAO56 [3] como un estimador robusto de la ETo a nivel mundial [4-6], cuyo error en la estimación es aceptable en diferentes regiones y puede utilizarse globalmente sin necesidad de calibración de parámetros adicionales [7]. Se han empleado diferentes métodos para estimar la ETo con información provista por estaciones meteorológicas en la Provincia de Córdoba, tales como Hargreaves-Samani (1985), Priestley-Taylor (1972), Turc (1963), Penman (1948) y Penman-Monteith (1963) [8-11]. La mayoría de los autores concluyen que el método más recomendado para estimar la ETo corresponde a la metodología de PM, incluso en situaciones de alta nubosidad [8]. Sin embargo, el modelo PM requiere forzantes que no siempre se encuentran disponibles (*e.g.* la radiación solar), lo cual limita su implementación a través de estaciones meteorológicas que carecen de los instrumentos de medición apropiados. La utilización de datos grillados de reanálisis surge como alternativa para superar la insuficiencia de observaciones y para realizar estimaciones espacialmente explícitas de la dinámica de la ETo, tanto a nivel local como regional [12].

Los productos de reanálisis estiman variables atmosféricas y de la superficie terrestre a partir de sistemas numéricos de asimilación y distintas fuentes de información (satélites, imágenes de radar y redes de estaciones meteorológicas). Existen diferentes productos que proporcionan datos de reanálisis diarios [13-16]. Sin embargo, debido a la accesibilidad de la información y a su facilidad de uso, uno de los más ampliamente utilizados es el *National Aeronautics and Space Administration*

Prediction of Worldwide Energy Resource (NASA-POWER) [17], que proporciona, además de otras variables, datos diarios de temperatura del aire, humedad relativa, velocidad y dirección del viento, precipitaciones y radiación solar. Todos los conjuntos de datos están disponibles a escala local, regional y global, con una amplia cobertura temporal que se extiende desde 1981 hasta la actualidad. No obstante, la utilización local de productos de reanálisis requiere evaluaciones previas mediante el empleo de datos de observaciones para cuantificar posibles errores y sesgos en la estimación [18]. Investigaciones realizadas en los últimos años en diferentes regiones del mundo mostraron una significativa correlación entre variables estimadas por NASA-POWER y registradas en estaciones meteorológicas, principalmente para la temperatura del aire y la radiación solar [12,19]. Sin embargo, la precisión y la bondad de ajuste de las estimaciones de ETo derivadas del producto NASA-POWER varían a lo largo de diferentes regiones mediterráneas [7,20,21].

En los últimos años, los modelos de aprendizaje automático (*machine learning*) se han utilizado en el campo de la hidrología, la agricultura y las ciencias ambientales para la estimación de la temperatura, la radiación solar y la ETo [27]. Estos algoritmos tienen la capacidad de captar relaciones complejas entre variables dependientes e independientes, lo que los convierte en potentes herramientas para la modelización de variables agrometeorológicas. Entre los modelos evaluados pueden destacarse las redes neuronales artificiales [22], las máquinas de vectores de soporte (SVM, en inglés) [23], *Random Forests* [24] y *Gradient Boosting Machines* [25]. En este sentido, una versión mejorada de los modelos *Gradient Boosting* denominada *Extreme Gradient Boosting* (XGBoost) ha sido recientemente aplicada en diferentes campos porque ha demostrado superioridad frente a otros algoritmos en términos de precisión y eficiencia (*i.e.* menor costo computacional y una mejor capacidad para superar los problemas de sobreajuste) [26,27]. Fan et al. [26] evaluaron diferentes modelos de aprendizaje automático en diferentes climas de China, y concluyeron que algoritmos relativamente sencillos como XGBoost son altamente recomendables para predecir la ETo diaria, incluso con baja disponibilidad de datos. En el mismo sentido, en un estudio realizado en el sur de India se demostró que los algoritmos XGBoost fueron los que presentaron los menores errores en la predicción diaria de la ETo [28]. No obstante, a pesar de haber sido evaluados en diferentes regiones, la estabilidad y eficacia de los modelos XGBoost dependen de determinados parámetros que deben ser validados previo a su implementación, entre los cuales pueden mencionarse el número de árboles, la máxima profundidad de árboles y la tasa de aprendizaje [27].

Dada la importancia de contar con estimaciones precisas de la ETo, los objetivos de este trabajo fueron (1) estimar la ETo diaria en el sudeste de Córdoba a través del método PM mediante el uso de las forzantes climáticas provistas por NASA-POWER, (2) estimar la ETo a partir de una mejora en la predicción de esas variables mediante un modelo XGBoost e información de estaciones meteorológicas disponibles en la región y (3) evaluar la precisión con que ambas aproximaciones predicen la ETo calculada con datos de dichas estaciones regionales. La obtención de un producto que permita estimar la ETo, a partir de datos de reanálisis, permite afrontar las problemáticas asociadas a la distribución heterogénea de estaciones y a la falta de registros en el área de estudio. Además, el empleo de productos grillados de reanálisis, disponibles en tiempo real resulta de gran utilidad para diferentes

aplicaciones operativas que requieren información sobre la ETo, tanto en ámbitos agropecuario como ambientales.

2 Materiales y métodos

2.1 Área de estudio

El sudeste de Córdoba (Figura 1) es una de las zonas más productivas de la provincia y se encuentra ocupado mayoritariamente por cultivos agrícolas y pasturas implantadas [29]. Posee un clima templado húmedo con una temperatura media anual de 17°C. El régimen pluviométrico oscila alrededor de 900 mm anuales, concentradas en el período primavera-estival. La zona fue seleccionada debido a que la producción agrícola-ganadera se realiza principalmente en condiciones de secano, donde la disponibilidad de agua en el suelo es crucial para asegurar un buen desarrollo y, por lo tanto, un alto rendimiento de los cultivos. Además, en los últimos años, la zona ha estado sujeta a la ocurrencia de eventos extremos de sequías e inundaciones que han generado pérdidas de gran magnitud en la producción agropecuaria.

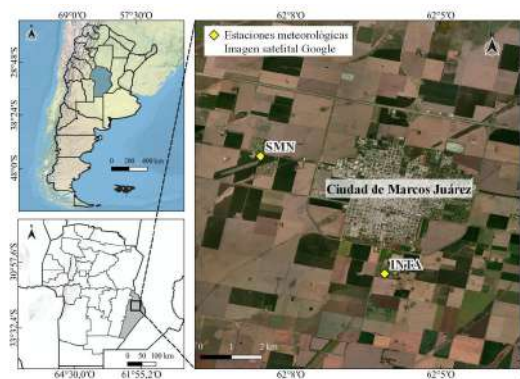


Figura 1. Área de estudio y ubicación de las estaciones meteorológicas pertenecientes al Instituto Nacional de Tecnología Agropecuaria (INTA) y al Servicio Meteorológico Nacional (SMN). Fuente: elaboración propia.

2.2 Fuentes de información

El período analizado comprendió desde el 01/01/1984 hasta el 31/12/2022. Se recopiló información diaria de dos estaciones meteorológicas convencionales pertenecientes al Instituto Nacional de Tecnología Agropecuaria (INTA) y al Servicio Meteorológico Nacional (SMN), separadas a una distancia de aproximadamente 6 km y aledañas a la Ciudad de Marcos Juárez (Figura 1). Ambas están incluidas en el Sistema de Información y Gestión Agrometeorológica del INTA (<http://siga.inta.gob.ar/#/>) y fueron seleccionadas porque brindan información homologada sobre variables meteorológicas y por el amplio período de registro (desde 1967 hasta la actualidad). Los datos diarios recopilados fueron: temperatura del aire máxima y mínima (Tmax y Tmin en °C), humedad relativa media (HR en %),

velocidad del viento a 2 metros de altura (U_2 en $\text{km}\cdot\text{h}^{-1}$) y radiación solar (R_s en $\text{MJ}\cdot\text{m}^{-2}\cdot\text{día}^{-1}$). En las dos estaciones meteorológicas, la R_s se estima a partir del método de Angström-Prescott [30], el cual utiliza como variable de entrada la heliofanía relativa (en %). En el caso del registro de viento, en la estación SMN se estima la velocidad a 2 metros de altura como el 80% de la velocidad de viento medida a 10 metros [3,31].

Los datos de reanálisis se descargaron del geoportal climático de NASA-POWER, el cual es de acceso libre y gratuito y permite extraer información diaria grillada a escala mundial, con un tamaño de celda de $0,5^\circ \times 0,5^\circ$. Las variables meteorológicas necesarias para el cálculo de la ETo se obtuvieron de la comunidad “*agrocimatology*” del producto NASA-POWER, a partir del píxel correspondiente a la ubicación de cada estación.

2.3 *Extreme Gradient Boosting (XGBoost)*

El algoritmo XGBoost, desarrollado por Chen & Guestrin [32], construye de forma iterativa una colección de árboles de decisión, donde cada árbol se crea a partir de la información de árboles creados previamente. XGBoost puede determinar de forma independiente los tipos de funciones de pérdida utilizados para la evaluación del modelo. Además, para reducir el riesgo de sobreajuste, se añade al modelo un término de regularización adicional. La puntuación media de cada árbol se utiliza como valor predictivo para la regresión.

El modelo XGBoost fue utilizado para predecir la información necesaria para la estimación de la ETo a partir de la información provista por NASA-POWER. De esta manera, los valores de T_{max} , T_{min} , R_s , HR y U_2 registrados en las estaciones fueron utilizados como variables objetivo (*target*) para ser predichas a partir del producto NASA-POWER (T_{max} , T_{min} , R_s , HR y U_2 como *inputs*). Los hiperparámetros del modelo se definieron mediante búsqueda en cuadrícula (*Grid Search Cross-Validation*). Para ello, el conjunto de datos disponible para entrenar el modelo se dividió en dos subconjuntos: el primero, denominado subconjunto de entrenamiento (80% elegido aleatoriamente), se utilizó para entrenar los modelos, y el segundo (20%) se utilizó como subconjunto de validación para el ajuste de los hiperparámetros. De esta manera, se eligieron los valores de los hiperparámetros que minimizaban el error en el subconjunto de validación. Por último, los modelos se probaron con datos del conjunto de testeo. En la determinación de los hiperparámetros, se ajustó el número de árboles (*n_estimators*), la profundidad máxima del árbol (*max_depth*) y la tasa de aprendizaje (*learning rate*). Los valores de número de árboles evaluados fueron 100, 200, 300, 400 y 500; la profundidad máxima de árbol osciló entre 3 y 6; y la tasa de aprendizaje varió entre 0,05 y 0,55. El modelo XGBoost fue implementado a través de la librería *xgboost* del software estadístico R©.

2.4 *Evapotranspiración de referencia*

La ETo a escala diaria se calculó mediante el método FAO-56 PM (Ecuación 1) a partir de las variables registradas en las estaciones, aquellas estimadas por NASA-POWER y las predichas por el modelo XGBoost. FAO-56 PM es considerado

como el mejor método para la estimación de ETo y es uno de los más precisos comparado con mediciones lisimétricas [33].

$$ET_o (mm) = \frac{0,408 \Delta (R_n - G) + \gamma \frac{900}{T + 273} U_2 (e_s - e_a)}{\Delta + \gamma (1 + 0,34 U_2)} \quad (1)$$

Donde Δ es la pendiente de la curva de presión de saturación de vapor ($\text{kPa} \cdot ^\circ\text{C}^{-1}$), R_n es la radiación neta en la superficie del cultivo de referencia ($\text{MJ} \cdot \text{m}^{-2} \cdot \text{día}^{-1}$), G es el flujo del calor del suelo ($\text{MJ} \cdot \text{m}^{-2} \cdot \text{día}^{-1}$), γ es la constante psicrométrica ($\text{kPa} \cdot ^\circ\text{C}^{-1}$), T es la temperatura media del aire ($^\circ\text{C}$), U_2 es la velocidad media del viento a 2 metros de la superficie ($\text{m} \cdot \text{s}^{-1}$), e_s constituye la presión de vapor de saturación (kPa), e_a es la presión real de vapor (kPa) y $(e_s - e_a)$ es el déficit de presión de vapor (kPa). Para intervalos de tiempo diarios, los valores de G son relativamente pequeños y, por lo tanto, el término es despreciado [3]. R_n se calcula con la Ecuación 2.

$$R_n = R_{ns} - R_{nl} \quad (2)$$

Donde R_{ns} es la radiación neta de onda corta ($\text{MJ} \cdot \text{m}^{-2}$), dada por la Ecuación 3, y R_{nl} es la radiación neta de onda larga ($\text{MJ} \cdot \text{m}^{-2}$), calculada por la Ecuación 4.

$$R_{ns} = (1 - \alpha) \cdot R_s \quad (3)$$

$$R_{nl} = \sigma \left[\frac{T_{max}^4 + T_{min}^4}{2} \right] (0,34 - 0,14 \sqrt{ea} \left(1,35 \frac{R_s}{R_{so}} - 0,35 \right)) \quad (4)$$

Donde R_s es la radiación solar global, α es el albedo de un cultivo herbáceo hipotético de referencia, T_{max} y T_{min} son la temperatura máxima y mínima del aire (K) respectivamente, y R_{so} es la radiación solar con cielo despejado ($\text{MJ} \cdot \text{m}^{-2}$), dada por la Ecuación 5.

$$R_{so} = (0,75 + 2z \cdot 10^{-5}) \cdot R_a \quad (5)$$

Donde z es la altitud de la estación (metros sobre el nivel del mar) y R_a es la radiación extraterrestre ($\text{MJ} \cdot \text{m}^{-2}$). Los valores de z en cada estación fueron de 114 m para la del SMN y de 110 m para la del INTA. Los parámetros R_a , e_a y Δ fueron calculados a través de las ecuaciones recomendadas por la FAO [3], las cuales no fueron descriptas en este trabajo. La información de viento a 2 metros de altura, expresada en $\text{km} \cdot \text{h}^{-1}$ en las estaciones meteorológicas, se convirtió a $\text{m} \cdot \text{s}^{-1}$ para su implementación en el modelo PM.

2.5 Evaluación de la precisión

Para cada estación meteorológica se contó con tres estimaciones de ETo: una a partir de los registros meteorológicos de las propias estaciones, otra a partir de las estimaciones de NASA-POWER y, finalmente, aquellas estimadas por XGBoost. Se consideró como variable observada a la ETo obtenida a partir de los datos de las estaciones meteorológicas. Los días que presentaron datos faltantes de temperatura,

humedad, velocidad de viento y/o radiación solar fueron descartados. Los estadísticos analizados fueron: el Coeficiente de Determinación (R^2), el Error Medio (EM) y su normalización (EMN), la Raíz del Error Cuadrático Medio (RMSE), la RMSE Relativa (RRMSE) y el coeficiente de Eficiencia de Nash-Sutcliffe (ENS) (Tabla 1). Un $R^2 > 0,75$ indica que la mayor parte de la varianza de los valores observados es explicada por el modelo de estimación con ajuste significativo, mientras que valores de R^2 de 0,25 y 0,50 corresponden a un ajuste débil y moderado, respectivamente [34]. El EM provee información sobre del sesgo de la estimación y permite cuantificar la sobreestimación o subestimación, mientras que el EMN relaciona el EM con la media de las observaciones. La RMSE mide las diferencias globales entre los valores observados y estimados, por lo que debe ser lo más baja posible. El RRMSE relaciona la RMSE con el promedio de datos observados y es útil para comparar errores en diferentes bases de datos. El ENS cuantifica la precisión en la estimación de la ETo por parte de NASA-POWER en relación con la media de la ETo observada. Valores de ENS cercanos a 1 indican una estrecha correlación entre la ETo medida y estimada.

Tabla 1. Estadísticos analizados para evaluar la precisión de NASA-POWER y XGBoost. E: dato observado en la estación meteorológica, \bar{E} : promedio de datos observados, S: dato estimado, \bar{S} : promedio de datos estimados, n: número de observaciones.

Estadístico	Ecuación	Valor deseable
Coeficiente de Determinación	$R2 = \left\{ \frac{\sum_{i=1}^n (E_i \bar{E}) (S_i \bar{S})}{\left[\sum_{i=1}^n (E_i \bar{E}) \right]^{0.5} \left[\sum_{i=1}^n (S_i \bar{S}) \right]^{0.5}} \right\}^2$	1
Error Medio	$EM = \frac{\sum_{i=1}^n (S_i - E_i)}{n}$	0
Error Medio Normalizado	$EMN = \frac{EM}{\bar{E}} \times 100$	< 15%
Raíz del Error Cuadrático Medio	$RMSE = \sqrt{\frac{\sum_{i=1}^n (S_i - E_i)^2}{n}}$	0
Raíz del Error Cuadrático Medio Relativa	$RRMSE = \frac{RMSE}{\bar{E}} \times 100$	< 15%
Coeficiente de Eficiencia de Nash-Sutcliffe	$ENS = 1 - \frac{\sum (S - E)^2}{\sum (E - \bar{E})^2}$	1

3 Resultados y discusión

Las estimaciones de NASA-POWER fueron mejores para temperatura máxima, temperatura mínima y humedad relativa, con un sesgo positivo para radiación y baja precisión para la velocidad de viento (Figura 2). En la estimación de la Tmax, el EM de NASA-POWER osciló entre 0,85 y 1,00 °C, con un EMN < 5%. La RMSE fue de aproximadamente 2,50 °C y la RRMSE se ubicó alrededor de 10%. La Tmin en la estación INTA mostró cierta tendencia a la subestimación (EM= -0,50 °C), con un RMSE=2,11 °C y una RRMSE de 18%. En la estación SMN, el EM en la estimación

de T_{min} fue de $0,87\text{ }^{\circ}\text{C}$, con una $RMSE=2,31\text{ }^{\circ}\text{C}$ y una $RRMSE=22\%$. Sin embargo, la eficiencia en la estimación de ambas estaciones fue muy elevada ($ENS>0,97$). Rodrigues & Braga [12] y Negm et al. [35] reportaron valores similares a los hallados en este trabajo en la evaluación de la estimación de la T_{max} y T_{min} por parte de NASA-POWER en Portugal e Italia, respectivamente. La humedad relativa también fue bien estimada ($RRMSE\approx 14\%$) con una tendencia a la subestimación de los valores superiores a 65% (Figura 2). NASA-POWER sobreestimó los valores de radiación en ambas estaciones, con un EM de $1,71\text{ MJ.m}^{-2}$ y de $1,94\text{ MJ.m}^{-2}$ para INTA y SMN, respectivamente. No obstante, el EMN fue inferior al 15% , con una $RMSE\approx 3,40\text{ MJ.m}^{-2}$ y una $RRMSE=22\%$. Además, las regresiones en ambas estaciones tuvieron pendientes muy cercanas a 1 con un buen ajuste ($R^2>0,87$ y $ENS>0,95$). En ambas estaciones meteorológicas no se midió radiación, sino que se estimó a partir de la heliofanía relativa (método de Angström-Prescott). Si bien esta metodología es una de las que menores errores produce en la estimación de la radiación en la región, para Marcos Juárez se reportó un valor de $RMSE$ de $3,33\text{ MJ.m}^{-2}\text{.día}^{-1}$, similar al que mostró NASA-POWER en este trabajo [36]. Sin embargo, hay antecedentes de que el índice de claridad de la atmósfera en distintas localidades y épocas del año es un determinante de la precisión de estimación de la radiación de NASA-POWER [37,38]. Por último, la velocidad de viento a 2 metros fue la variable que presentó los mayores errores en la estimación (Figura 2). El EMN fue mayor a 15% , con una tendencia a la subestimación por parte de NASA-POWER en la estación SMN ($-20,85\%$). La $RMSE$ osciló entre $1,03$ y $1,34\text{ m.s}^{-1}$ y la $RRMSE$ entre 46 y 59% para INTA y SMN, respectivamente. La velocidad de viento en NASA-POWER se estima en base al promedio de mediciones horarias con parámetros de rugosidad de la superficie elaborados a escala global [39]. Por su parte, en la estación SMN corresponde al promedio de los 24 registros horarios de los anemómetros, mientras que en la del INTA se consideran sólo 3 observaciones (9, 15 y 21 h UTC-3). Además, en la estación SMN los anemómetros se encuentran instalados a mayor altura y se realiza una corrección grosera para estimar la velocidad del viento a 2 metros (80% de la velocidad registrada a 10 metros).

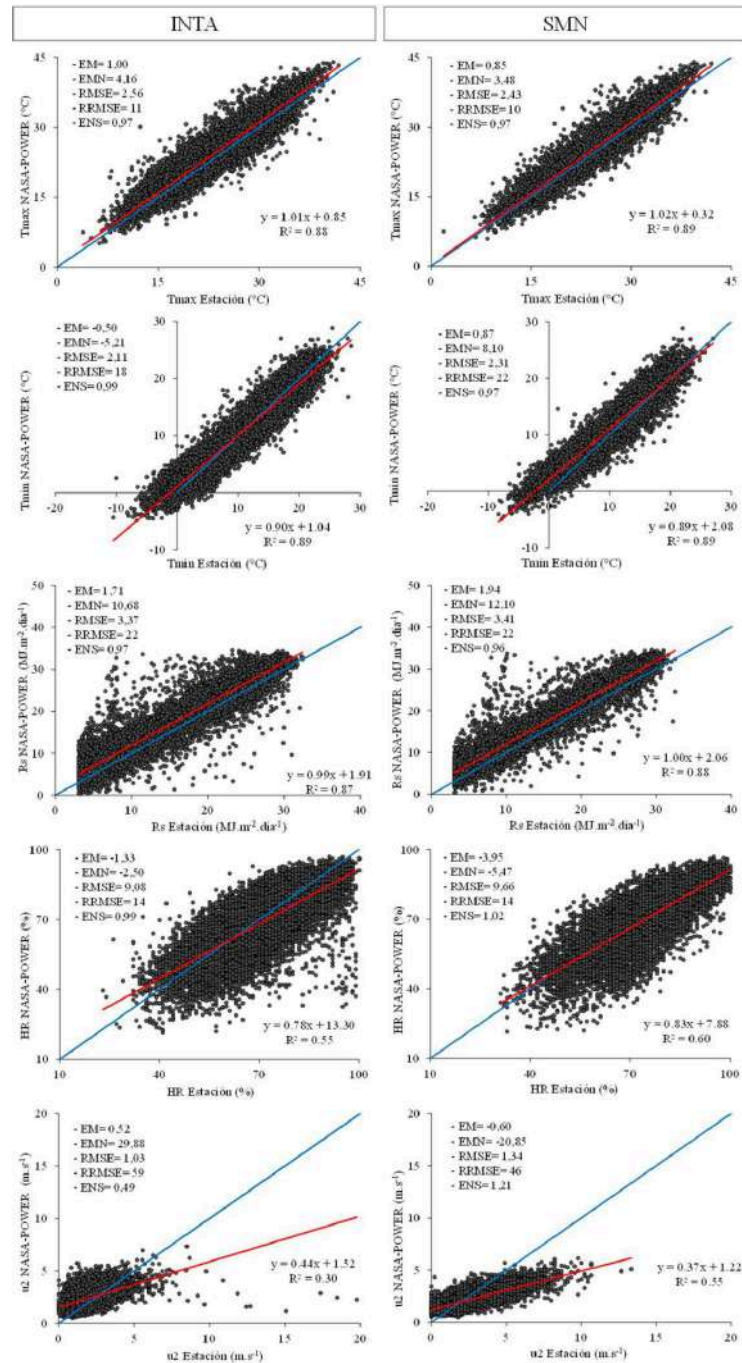


Figura 2. Diagramas de dispersión entre las variables registradas en las estaciones y aquellas estimadas por NASA-POWER. La línea sólida roja corresponde a la tendencia lineal, mientras que la línea sólida azul representa el ajuste 1:1.

La estimación de la ETo PM a partir de las variables climáticas provistas por NASA-POWER, respecto al calculado con datos de ambas estaciones, mostró una precisión moderada con una leve sobreestimación. La ETo media calculada para la estación meteorológica de INTA fue de 3,17 mm.día⁻¹, con un mínimo de 0,42 mm.día⁻¹ y un máximo de 10,06 mm.día⁻¹ (n=13126). En la estación SMN, la ETo osciló entre 0,24 mm.día⁻¹ y 11,59 mm.día⁻¹, con un valor promedio de 3,37 mm.día⁻¹ (n=9129). A partir de los datos NASA-POWER se obtuvo un promedio de 3,69 mm.día⁻¹, con un mínimo de 0,50 mm.día⁻¹ y un máximo de 12,05 mm.día⁻¹. Para ambas estaciones, aunque la sobreestimación de la ETo fue baja, con EMN de 15 y 10%, la precisión fue moderada, con una RRMSE de 28 y 25% para INTA y SMN, respectivamente (Figura 3). Valores similares fueron reportados por Paredes et al. [6] y Pelosi et al. [37], quienes evaluaron la estimación de la ETo mediante el producto NASA-POWER a través de diferentes modelos en la península itálica y en Portugal, respectivamente.

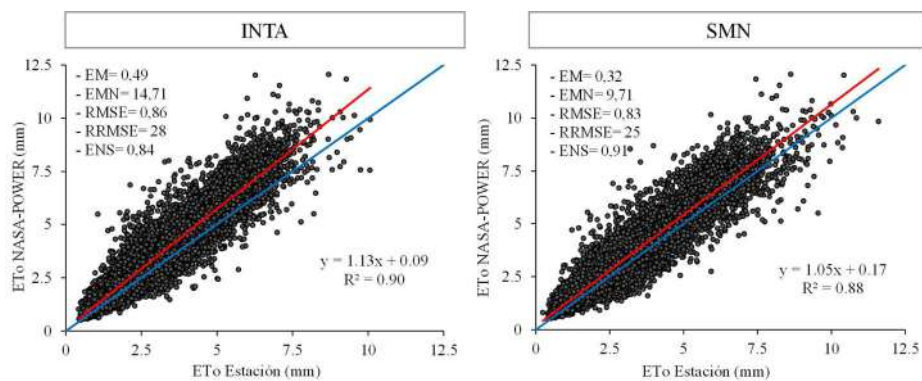


Figura 3. Diagramas de dispersión entre la evapotranspiración de referencia estimada a partir de los datos observados (ETo Estación) y aquella estimada mediante los datos NASA-POWER (ETo NASA-POWER). La línea sólida roja corresponde a la tendencia lineal, mientras que la línea sólida azul representa el ajuste 1:1.

La estimación de la ETo diaria mostró diferencias estacionales en la sobreestimación y grado de precisión. En los trimestres primaveral (septiembre, octubre y noviembre) y otoñal (marzo, abril y mayo), el cálculo con NASA-POWER, tanto comparado con la estación del INTA como la del SMN, alcanzó mayor precisión y baja sobreestimación (Figuras 4 y 5). En el trimestre estival (diciembre, enero y febrero) se observaron las mayores sobreestimaciones, que en varios meses superaron el 19%, tanto para los días con baja como con alta ETo (Figuras 4 y 5). En cambio, en el trimestre invernal (junio, julio y agosto) se observó la menor precisión y una tendencia a subestimar los valores más elevados para esos meses (Figuras 4 y 5).

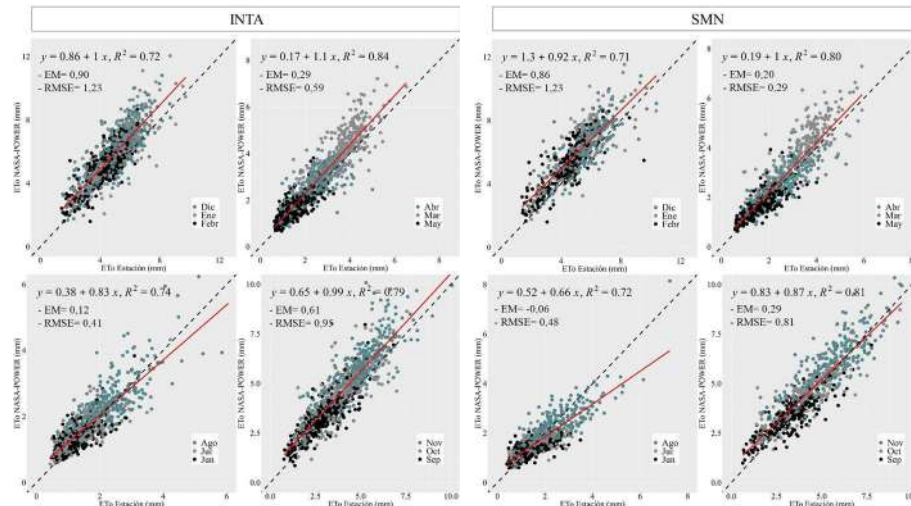


Figura 4. Diagramas de dispersión trimestrales entre la evapotranspiración de referencia estimada a partir de los datos observados en las estaciones INTA y SMN (ETo Estación) y aquella estimada mediante los datos NASA-POWER (ETo NASA-POWER). La línea sólida roja corresponde a la tendencia lineal, mientras que la línea punteada negra representa el ajuste 1:1.

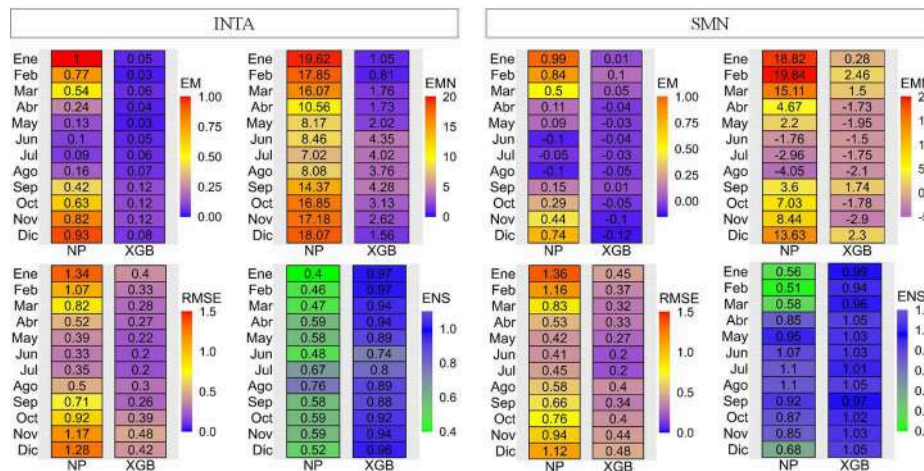


Figura 5. Variación mensual de los estadísticos de la comparación entre la ETo observada en las estaciones y la estimada con las variables climáticas de NASA-POWER (NP) y su ajuste mediante XGBoost (XGB).

Para evaluar el impacto de las forzantes con más error en NASA-POWER, velocidad del viento y radiación, se realizaron estimaciones de la ETo a través del reemplazo de los datos de reanálisis por aquellos observados en cada estación meteorológica. Cuando se utilizó la velocidad del viento a 2 metros registrada en la estación INTA, el error en la estimación de la ETo se redujo de 0,49 a 0,29 mm.día⁻¹,

mientras que la RMSE, de 0,86 a 0,64 mm.día⁻¹. Por el contrario, en la estación SMN, cuya medición a 10 m se multiplica por 0,8, el EM para ETo aumentó de 0,32 a 0,45 mm.día⁻¹ y el RMSE se elevó levemente, de 0,83 a 0,85 mm.día⁻¹. Estacionalmente, en la primavera, el verano y el otoño el EM se redujo entre 18 y 24% cuando se utilizaron datos reales de U2 en la estación INTA, mientras que la RMSE se redujo entre 19 y 22%. En el invierno la magnitud de la reducción del EM y la RMSE fue de alrededor de 11%. Contrariamente, en la estación SMN, el EM aumentó entre un 13 y 29% y la RMSE fue aproximadamente un 10% mayor en el período primavero-estival. Estas inconsistencias podrían estar asociadas a la diferente altura en la colocación de los anemómetros en las estaciones.

Cuando se reemplazaron los datos de radiación NASA-POWER por aquellos estimados a partir de heliofania en las estaciones, el error en la estimación de la ETo se redujo de 0,49 a 0,33 mm.día⁻¹, mientras que la RMSE se redujo de 0,86 a 0,67 mm.día⁻¹, para la estación del INTA. También, para la del SMN, el EM se redujo de 0,32 a 0,14 mm.día⁻¹ y la RMSE, de 0,83 a 0,65 mm.día⁻¹. El EM y la RMSE de la estimación de la ETo se redujeron entre 22 y 32% en primavera y verano y entre un 7 y 12 % en otoño e invierno para la estación INTA, y entre 18 y 33 % y entre 6 y 18 %, respectivamente, para la del SMN. Las diferencias en ETo calculadas solo con datos de NASA-POWER están entonces asociados a sus mayores valores de radiación, especialmente durante los meses más cálidos, en los que, además, el término de radiación de la ecuación PM ejerce un mayor peso en la estimación de la ETo [33,35]. En este aspecto, Bois et al. [33] realizaron un análisis de sensibilidad del uso de radiación obtenida por sensores remotos en la fórmula PM en Francia y reportaron mayores errores durante el verano, donde la radiación solar tiene mayor impacto en el modelo. El uso de modelos que se rigen principalmente por las transferencias radiativas (*e.g.* modelos de Hargreaves y de Priestley–Taylor) para estimar ETo conduce a sesgos más elevados cuando se reemplazan valores de radiación medidos por aquellos estimados por sensores remotos [35]. La corrección de las variables climáticas antes de su utilización en la estimación de la ETo mejora sensiblemente la precisión [12].

En el ajuste de los hiperparámetros del modelo XGBoost se utilizaron los datos de las dos estaciones meteorológicas, para obtener así un modelo de predicción de cada variable con un número mayor de observaciones ($n=22263$). El número de árboles, la tasa de aprendizaje y la máxima profundidad de cada árbol fueron diferentes para cada una de las variables objetivo (Tabla 3). El comportamiento del modelo XGBoost en el conjunto de testeo mostró un ajuste significativo en la predicción de la temperatura, la humedad relativa y la radiación solar, mientras que para la velocidad de viento el ajuste fue moderado (Tabla 4). Aunque en este trabajo no se ha evaluado la predicción en base a cada variable por separado, el ajuste de la combinación de Tmax, Tmin, Rs, HR y U2 como *inputs* indicaría un buen comportamiento en la predicción de la variable objetivo. Al respecto, en el Norte de China se halló que la Tmax, la Tmin, la Rs, la HR y la U2, empleadas conjuntamente como predictoras de la ETo a partir de diferentes modelos de aprendizaje automático, arrojaron los mejores resultados en términos de precisión que las variables tomadas por separado [27].

Tabla 3. Hiperparámetros ajustados del modelo XGBoost para la predicción de la temperatura máxima (Tmax), temperatura mínima (Tmin), radiación solar (Rs), humedad relativa (HR) y velocidad de viento (U2) a partir de las variables NASA-POWER. ETA: tasa de aprendizaje.

Target (estación)	Inputs (NASA-POWER)	Máx.Prof	ETA	Nro. de árboles
Tmax	Tmax, Tmin, Rs, HR, U2	11	0,05	500
Tmin	Tmax, Tmin, Rs, HR, U2	11	0,05	200
Rs	Tmax, Tmin, Rs, HR, U2	11	0,05	200
HR	Tmax, Tmin, Rs, HR, U2	10	0,05	200
U2	Tmax, Tmin, Rs, HR, U2	3	0,15	100

Tabla 4. Estadísticos de ajuste del modelo XGBoost durante la fase de testeo para cada una de las variables de entrada.

	EM	EMN	RMSE	RRMSE	R ²
Temperatura máxima	0,05	0,24	1,58	0,15	0,94
Temperatura mínima	0,03	0,27	1,98	0,18	0,91
Radiación solar	0,01	0,04	1,26	0,11	0,97
Humedad relativa	0,05	0,08	4,37	0,22	0,88
Velocidad de viento	0,01	-0,16	0,53	0,28	0,71

El modelo XGBoost demostró un buen comportamiento en la predicción de las variables de estimación de la ETo. En la Tmax, Tmin y Rs, los errores se redujeron entre un 70 y un 95% (EM, EMN, RMSE y RRMSE), mientras que el ENS osciló entre 0,99 y 1, lo que sugiere estrecha correlación entre las variables predichas y registradas (Figura 6). Por otro lado, en cuanto a la velocidad de viento y la humedad relativa, si bien hubo una mejora en la predicción mediante XGBoost, su magnitud osciló entre un 10 y un 15%, además de observarse una tendencia a la subestimación en los valores predichos en la estación SMN (EM<0). En cuanto a la ETo, el promedio estimado por XGBoost en la estación INTA fue de 3,24 mm.día⁻¹, con un mínimo de 0,44 mm.día⁻¹ y un máximo de 9,88 mm.día⁻¹. En la estación SMN, los valores de ETo estimados oscilaron entre 0,44 y 10,35 mm.día⁻¹, con un promedio de 3,30 mm.día⁻¹. Esto implica una diferencia de 0,07 mm.día⁻¹ entre el valor estimado por XGBoost y el valor medio registrado en ambas estaciones. Los errores se redujeron entre 50-80% cuando se utilizaron las variables predichas por XGBoost (Figuras 3 y 7). La eficiencia en la predicción fue elevada (ENS ≈1), aunque se observó una leve tendencia a la subestimación de los valores de ETo en la estación SMN (EM y EMN <0). Esta tendencia también se observó en la comparación trimestral de la estación SMN, especialmente en el trimestre invernal y primaveral. No obstante, en ambas estaciones, se alcanzó una buena predicción en todos los trimestres cuando se utilizaron las variables predichas por XGBoost (Figuras 8 y 9). El modelo XGBoost permitió reducir los errores durante los meses primavero-estivales, en los cuales la RMSE no superó los 0,50 mm.día⁻¹ y el EMN fue inferior a 5%, con un EM < 0,12 mm-día⁻¹ (Figura 5). Valores de RMSE, ENS y EM fueron hallados en la evaluación de diferentes modelos de aprendizaje automático en China [26], en España [40] y en el Este de Brasil [41].

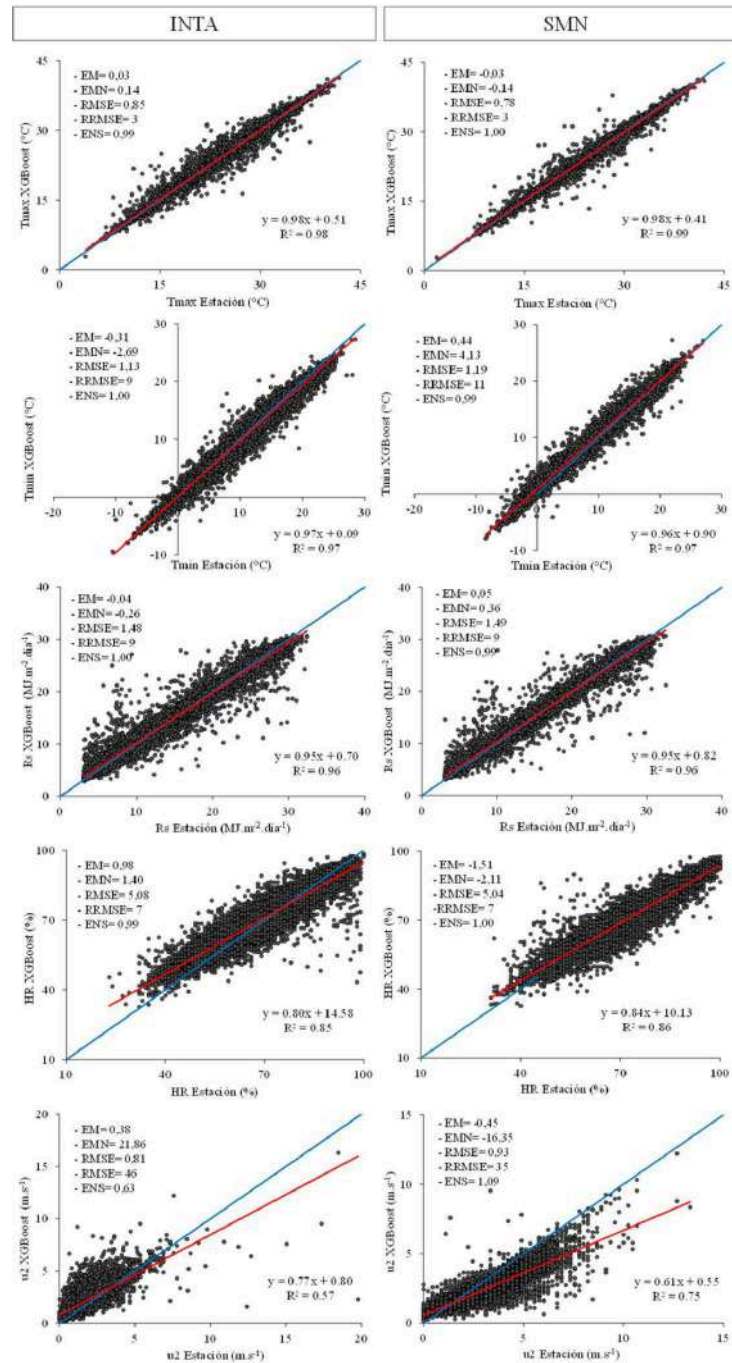


Figura 6. Diagramas de dispersión entre las variables registradas en las estaciones y aquellas predichas por XGBoost. La línea sólida roja corresponde a la tendencia lineal, mientras que la línea sólida azul representa el ajuste 1:1.

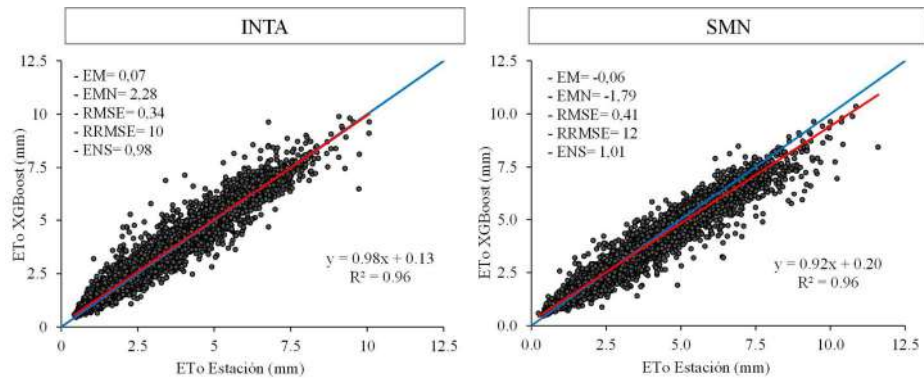


Figura 7. Diagramas de dispersión entre la evapotranspiración de referencia estimada a partir de los datos observados (ETo Estación) y aquellos predichos por XGBoost (ETo XGBoost). La línea sólida roja corresponde a la tendencia lineal, mientras que la línea sólida azul representa el ajuste 1:1.

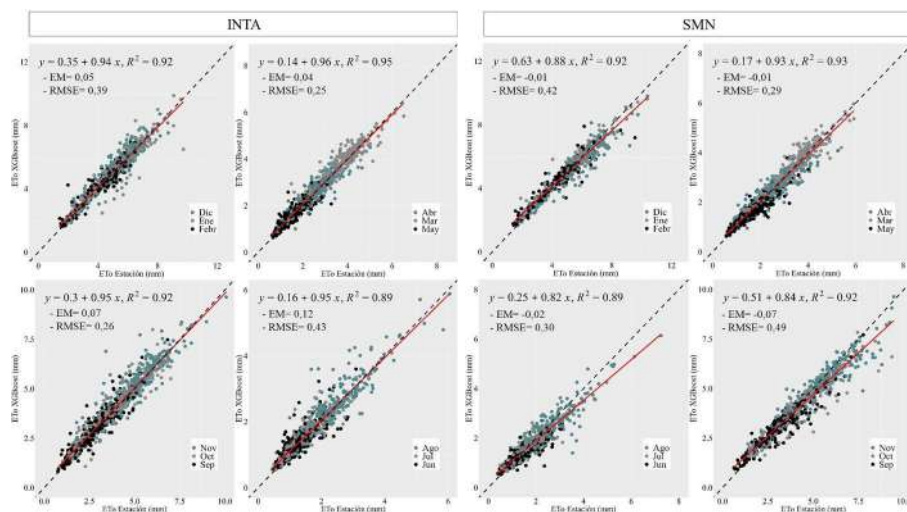


Figura 8. Diagramas de dispersión trimestrales entre la evapotranspiración de referencia estimada a partir de los datos observados en la estación INTA (ETo Estación) y aquella predicha mediante XGBoost (ETo XGBoost). La línea sólida roja corresponde a la tendencia lineal, mientras que la línea punteada negra representa el ajuste 1:1.

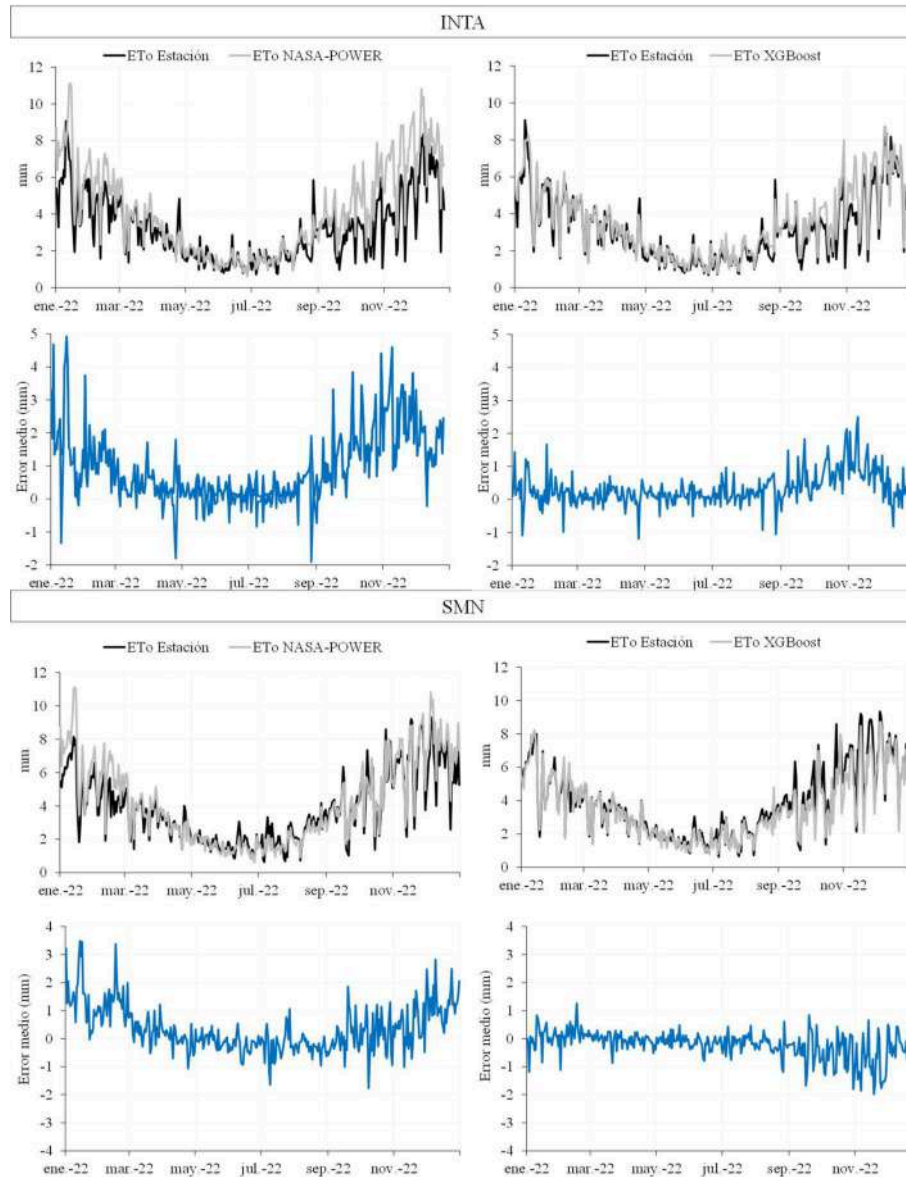


Figura 9. Evolución temporal y error medio de la ETo registrada en las estaciones, la estimada con los datos NASA-POWER y aquella estimada con los datos predichos por XGBoost.

Finalmente, dado que la disponibilidad de los datos de radiación NASA-POWER presenta una demora de aproximadamente 10 días, se buscó predecir la ETo a partir de la T_{max} , T_{min} , HR y U_2 corregidas mediante XGBoost. El ajuste del modelo arrojó 200 árboles, una máxima profundidad de 6 y una tasa de aprendizaje de 0,05 como hiperparámetros. En el conjunto de testeo, se obtuvo un $EM = -0,01 \text{ mm.día}^{-1}$, un $EMN = -0,27 \%$, una $RMSE = 0,72 \text{ mm.día}^{-1}$ y un $R^2 = 0,85$. El modelo ajustado tuvo

un buen comportamiento en la predicción de la ETo sin la inclusión de la radiación solar (Figura 10). El EM fue de $0,06 \text{ mm.día}^{-1}$ en la estación INTA, con un EMN= $1,89 \%$ y una ENS= $0,98$. En la estación SMN se observó una débil tendencia a la subestimación ($EM = -0,08 \text{ mm.día}^{-1}$ y $EMN = -2,43\%$). En ambas la RMSE fue de $0,54 \text{ mm.día}^{-1}$ y la RRMSE= 16% . La predicción de la ETo a partir de los datos NASA-POWER corregidos por XGBoost sin incluir la radiación mejoró la estimación con respecto a la información NASA-POWER sin corregir. Sin embargo, los errores fueron levemente superiores a aquellos obtenidos cuando se utilizaron todas las variables, incluida la radiación corregida por XGBoost. Los resultados obtenidos indican que XGBoost permitió captar las diferencias en la estimación de las variables NASA-POWER en comparación con aquellas registradas en las estaciones y lograr un modelo de estimación con un buen ajuste, incluso si no existieran datos de radiación disponibles. Esto lo convierte en una herramienta de gran utilidad para estimar en forma precisa las variables necesarias para la estimación de la ETo en el sudeste de Córdoba.

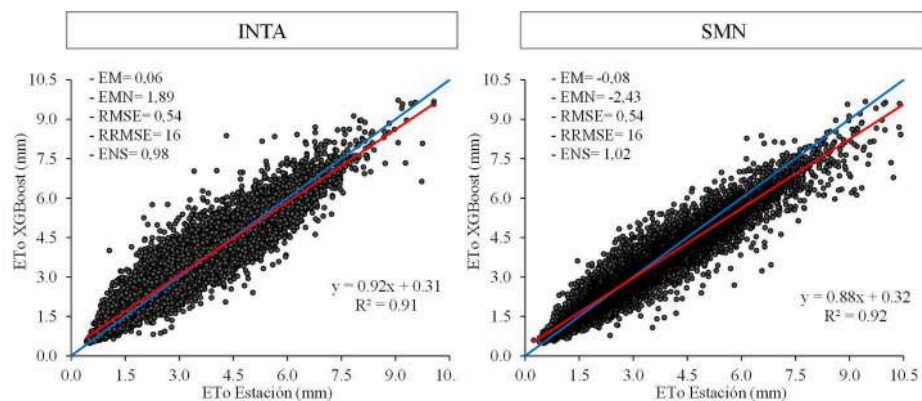


Figura 10. Diagramas de dispersión entre la evapotranspiración de referencia estimada a partir de los datos observados (ETo Estación) y aquellos predichos por XGBoost sin incluir la radiación solar (ETo XGBoost). La línea sólida roja corresponde a la tendencia lineal, mientras que la línea sólida azul representa el ajuste 1:1.

4 Conclusiones

Se evaluó la información provista por NASA-POWER como alternativa para estimar la ETo mediante el modelo PM en el sudeste de Córdoba. Los resultados obtenidos muestran que, a partir del uso del producto NASA-POWER, se puede realizar una estimación de la ETo con bajo error, lo cual resulta aceptable para determinados análisis agrícolas y/o ambientales, sobre todo en áreas de secano. En la comparación de la ETo estimada con la observada en dos estaciones meteorológicas convencionales se alcanzaron errores (EM y RMSE) inferiores a 1 mm.día^{-1} y una ENS $> 0,75$. En la evaluación intermensual, los mejores ajustes se observaron durante los meses primaverales y otoñales, mientras que los meses invernales y estivales presentaron mayores errores y una menor correlación entre la ETo estimada y observada. Las diferencias en los valores de radiación diaria, especialmente los meses

más cálidos, son la principal causa de desvíos en las estimaciones de ETo. En cambio, la velocidad del viento genera un error no consistente en ambas estaciones meteorológicas.

La estimación de ETo tuvo una mejora importante cuando se empleó el modelo de aprendizaje automático XGBoost para predecir las variables de entrada del método PM. Se corrigió el sesgo y aumentó la precisión respecto de los datos originales de NASA-POWER. El EM fue inferior a $0,10 \text{ mm.día}^{-1}$ y la RMSE no superó los $0,41 \text{ mm.día}^{-1}$, con una elevada precisión ($\text{ENS} \approx 1$). La mejora en la estimación se produjo principalmente durante los meses primavera-estivales. Los resultados obtenidos indican que XGBoost permitió captar las diferencias en la estimación de las variables NASA-POWER en comparación con aquellas registradas en las estaciones y lograr un modelo de estimación de la ETo en el sudeste de Córdoba, incluso cuando no se disponga de datos de radiación. Esto resulta de utilidad para diferentes aplicaciones hidrológicas, agronómicas o ambientales en áreas con escasa información meteorológica, sobre todo histórica, o con redes de medición dispersas e irregulares que requieran estimar la ETo en tiempo real.

5 Referencias

1. Guo D., Westra, S., Maier, H. R. Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones. *Hydrology and Earth System Sciences* 21(4), 2107–2126 (2017).
2. Mu, Q., Zhao, M. Running, S.W. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment* 115, 1781–1800 (2011).
3. Allen, R.G., Pereira, L.S., Raes, D., Smith, M. Crop evapotranspiration —guidelines for computing crop water requirements. *FAO Irrigation and drainage paper* 56 (1998).
4. Mohawesh, O.E., Talazi, S.A. Comparison of Hargreaves and FAO56 equations for estimating monthly evapotranspiration for semi-arid and arid environments. *Archives of Agronomy and Soil Science* 58, 321–334 (2012).
5. Song, X., Lu, F., Xiao, W., Zhu, K., Zhou, Y., Xie, Z. Performance of 12 reference evapotranspiration estimation methods compared with the Penman-Monteith method and the potential influences in northeast China. *Meteorological Applications* 26, 83–96 (2019).
6. Paredes, P., Fontes, J.C., Azevedo, E.B., Pereira, L.S. Daily reference crop evapotranspiration in the humid environments of Azores islands using reduced data sets: Accuracy of FAO-PM temperature and Hargreaves-Samani methods. *Theoretical and Applied Climatology* 134, 595–611 (2018).
7. Rodrigues, G. C., Braga, R. Estimation of Daily Reference Evapotranspiration from NASA POWER Reanalysis Products in a Hot Summer Mediterranean Climate. *Agronomy* 11(10), 2077 (2021).
8. López, M. L. Análisis de los métodos de estimación de evaporación y evapotranspiración a las condiciones locales de la ciudad de Córdoba. Tesis de Maestría en Ciencias Hídricas. Santa Rosa, La Pampa, (2011).
9. De la Casa, A., Ovando, G.G. Climate change and its impact on agricultural potential in the central region of Argentina between 1941 and 2010. *Agricultural and Forest Meteorology* 195-195, 1-11 (2014).
10. Carmona, F., Holzman, M., Rivas, R., Degano, M. F., Kruse, E., Bayala, M. Evaluación de dos modelos para la estimación de la evapotranspiración de referencia con

- datos CERES. Asociación Española de Teledetección, *Revista de Teledetección* 51, 87-98 (2018).
11. Degano, M. F., Rivas, R. E., Carmona, F., Niclós, R., Sánchez, J.M. Evaluation of the MOD16A2 evapotranspiration product in an agricultural area of Argentina, the Pampas region. *The Egyptian Journal of Remote Sensing and Space Sciences* 24, 319-328 (2021).
 12. Rodrigues, G. C., Braga, R. Evaluation of NASA POWER Reanalysis Products to Estimate Daily Weather Variables in a Hot Summer Mediterranean Climate. *Agronomy* 11(6), 1207 (2021).
 13. Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., Vitart, F. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.* 137, 553-597 (2011).
 14. Kobayashi, S., Yukinari, O.T.A., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Miyaoka, K., Takahashi, K. The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Ser. II Ser. II* 93, 5–48, (2015).
 15. Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.K., Hnilo, J.J., Fiorino, M., Potter, G.L. NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Am. Meteorol. Soc.* 83, 1631–1644 (2002).
 16. Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., Woollen, J. MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate* 24 (14), 3624–3648 (2011).
 17. Chandler, W.S., Hoell, J. M., Westberg, D., Zhang, T., Stackhouse, P. W., Jr. NASA prediction of worldwide energy resource high resolution meteorology data for sustainable building design. In *Proceedings of the Solar 2013 Conference of American Solar Energy Society*, Baltimore, MD, USA, 16–20 April 2013.
 18. Sheffield, J., Ziegler, A.D., Wood, E.F., Chen, Y. Correction of the high-latitude rain day anomaly in the NCEP–NCAR reanalysis for land surface hydrological modeling. *Journal of Climate* 17, 3814–3828 (2004).
 19. Sayago, S., Ovando, G., Almorox, J., Bocco, M. Daily solar radiation from NASA-POWER product: assessing its accuracy considering atmospheric transparency. *International Journal of Remote Sensing* 41(3), 897-910 (2019).
 20. Srivastava, P.K., Singh, P., Mall, R.K., Pradhan, R.K., Bray, M., Gupta, A. Performance assessment of evapotranspiration estimated from different data sources over agricultural landscape in Northern India. *Theoretical and Applied Climatology* 140, 145–156 (2020).
 21. Monteiro, A.L., Sentelhas, P.C., Pedra, G.U. Assessment of NASA/POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation. *International Journal of Climatology* 38, 1571–1581 (2018).
 22. Meshram, S.G., Singh, V.P., Kisi, O., Karimi, V., Meshram, C. Application of artificial neural networks, support vector machine and multiple model-ANN to sediment yield prediction. *Water Resour. Manage.* 34, 4561–4575 (2020).

23. Nourani, V., Elkiran, G., Abdullahi, J.. Multi-station artificial intelligence based ensemble modeling of reference evapotranspiration using pan evaporation measurements. *J. Hydrol. (Amst)*. 577, 123958 (2019). <https://doi.org/10.1016/j.jhydrol.2019.123958>.
24. Wang, S., Lian, J., Peng, Y., Hu, B., Chen, H. Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi .China. *Agric. Water Manag.* 221, 220–230 (2019). <https://doi.org/10.1016/j.agwat.2019.03.027>.
25. Saggi, M. K., Jain, S. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. *Computers and Electronics in Agriculture* 156, 387–398, (2019).
26. Fan, J., Yue, W., Wu, L., Zhang, F., Cai, H., Wang, X., Lu, X., Xiang, Y. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. *Agricultural and Forest Meteorology* 263, 225–241 (2018). <https://doi.org/10.1016/j.agrformet.2018.08.019>.
27. Han, Y. Wu, J., Zhai, B., Pan, Y., Huang, G., Wu, L., Zeng, W. Coupling a Bat Algorithm with XGBoost to Estimate Reference Evapotranspiration in the Arid and Semiarid Regions of China. *Advances in Meteorology* 2019, Article ID 9575782. <https://doi.org/10.1155/2019/9575782>.
28. Agrawal, Y., Kumar, M., Ananthkrishnan, S., Kumarapuram, G. Evapotranspiration Modeling Using Different Tree Based Ensembled Machine Learning Algorithm. *Water Resources Management* 36, 1025–1042 (2022). <https://doi.org/10.1007/s11269-022-03067-7>.
29. García, C. L., I. Teich, A. Kindgard, M. J. Bosio, A. Ravelo, M. Piumetto, H. Morales, Fuentes, M. L. Mapas de cobertura del suelo de la provincia de Córdoba 2017/2018. Informe técnico. Comisión Nacional de Investigaciones Científicas y Técnicas (CONICET) y Ministerio de Finanzas de la provincia de Córdoba. Córdoba, Córdoba, Argentina. 2018.
30. Rietveld, M. R. A new method for estimating the regression coefficients in the formula relating solar radiation to sunshine. *Agricultural Meteorology* 19, 243-525 (1977).
31. Fernández Long, M. E., Spescha, L., Barnatán, I., Murphy, G. Modelo de Balance Hidrológico Operativo para el Agro (BHOA). *Revista Agronomía & Ambiente FAUBA* 32 (1-2), 31-47 (2012).
32. Chen, T., Guestrin, C., XGBoost: a scalable tree boosting system. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785-794 (2016). <https://doi.org/10.1145/2939672.2939785>.
33. Bois, B., Pieri, P., Van Leeuwen, C., Wald, L., Huard, F., Gaudillere, J.P., Saur, E. Using remotely sensed solar radiation data for reference evapotranspiration estimation at a daily time step. *Agricultural and Forest Meteorology* 148, 619-630 (2008).
34. Henseler, J., Ringle, C., Sinkovics, R. The use of partial least squares path modeling in international marketing. In *New Challenges to International Marketing*; Emerald Group Publishing Limited: Bingley, UK, 20, 277–320 (2009).
35. Negm, A., Jabro, J., & Provenzano, G. Assessing the suitability of American National Aeronautics and Space Administration (NASA) agro-climatology archive to predict daily meteorological variables and reference evapotranspiration in Sicily, Italy. *Agricultural and Forest Meteorology*, 244-245, 111-121 (2017). <https://doi.org/10.1016/j.agrformet.2017.05.022>.
36. De La Casa, A., Ovando, G., Rodríguez, A. Estimación de la radiación solar global en la provincia de Córdoba, Argentina, y su empleo en un modelo de rendimiento potencial de papa. *Revista de Investigaciones Agropecuarias (INTA)*, 32 (2), 45-62 (2003).

37. Pelosi, A., Terribile, F., D'Urso, G., Chirico, Giovanni Battista. Comparison of ERA5-Land and UERRA MESCAN-SURFEX Reanalysis Data with Spatially Interpolated Weather Observations for the Regional Assessment of Reference Evapotranspiration. *Water* 12, 1669 (2020).
38. Sayago, S., Ovando, G., Bocco, M. Estimación de radiación solar diaria a partir de imágenes NASA-POWER con distintos índices de claridad. *Avances en Energías Renovables y Medio Ambiente* 21, 49-58 (2017).
39. Chandler, W. S., Whitlock, C.H., Stackhouse, P.W. Determining Wind Resources as a Function of Surface Roughness and Height from NASA Global Assimilation Analysis. *Proceedings of the International Solar Energy Society 2005 Solar World Congress, August 6-12, Orlando, Florida.*
40. Bellido-Jiménez, J. A., Estévez, J., García-Marín, A. P. New machine learning approaches to improve reference evapotranspiration estimates using intra-daily temperature-based variables in a semiarid region of Spain. *Agricultural Water Management* 245, 106558 (2021). <https://doi.org/10.1016/j.agwat.2020.106558>.
41. Ferreira, L. B., da Cunha, F. F. 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 (2020). <https://doi.org/10.1016/j.agwat.2020.106113>.