

ALTERNATIVA PROBABILÍSTICA BAYESIANA PARA UNA ESTIMACIÓN MEJOR DEL COEFICIENTE DE CULTIVO BASADO EN ÍNDICES DE VEGETACIÓN

A BAYESIAN PROBABILISTIC ALTERNATIVE FOR A BETTER ESTIMATION OF THE CROP COEFFICIENT BASED ON VEGETATION INDEXES

Víctor Manuel **Salas-Aguilar**^{*}, María Elena **Torres-Olave**, Hugo Luis **Rojas-Villalobos**,
Luis Carlos **Alatorre-Cejudo**, Luis Carlos **Bravo-Peña**

Geoinformática, Universidad Autónoma de Ciudad Juárez. DMC Cuauhtémoc. Carretera Cuauhtémoc-Anáhuac Km 3.5, Ejido Cuauhtémoc. 31600 Ciudad Cuauhtémoc, Chihuahua. (victor.salas@uacj.mx).

RESUMEN

La estimación robusta y operativa de los requerimientos hídricos en los cultivos, es una premisa que tienen que cumplir los especialistas en el manejo del agua. El objetivo de este estudio fue implementar el enfoque bayesiano (EB) para estimar el coeficiente de cultivo (K_c) a partir del índice de vegetación de diferencia normalizada (NDVI), y comparar con los enfoques frecuentistas, ecuaciones generales y específicas para un cultivo. El modelo bayesiano se desarrolló con base en distribuciones de probabilidad de parámetros (a y b) *a priori*, recopilados de 18 investigaciones, que estimaron la relación lineal (K_c -NDVI). En una primera etapa se compararon los enfoques en siete tamaños de muestra (TM) (5, 10, 30, 60, 90, 120 y 150), en ellas se realizaron 1000 repeticiones sin remplazo. Para validar el ajuste de los enfoques en cada TM se utilizaron 156 datos experimentales de K_c -NDVI. En la segunda etapa se evaluó el EB respecto a ecuaciones ajustadas a cultivos específicos, se añadió el primer y segundo cuartil de los datos para hacer la validación. Los resultados mostraron en primera instancia que el EB superó en todos los TM a los métodos frecuentistas y ecuaciones generales. El EB demostró obtener incertidumbres bajas con solo cinco TM, a diferencia de los otros métodos que necesitaron más de 30 registros para obtener resultados similares. La comparación entre EB y ecuaciones específicas corroboró que cinco datos aleatorios recopilados en el primer cuartil de los registros fueron suficientes para obtener incertidumbres bajas. La metodología propuesta es operativa y los datos de K_c se pueden estimar desde las primeras fases fenológicas con certidumbre alta y con pocos datos del NDVI.

ABSTRACT

The robust and operative estimation of water requirements in crops is a premise that water management specialists must fulfill. The aim of this study was to implement the Bayesian approach (EB) to estimate the crop coefficient (K_c) from the normalized difference vegetation index (NDVI), and to compare them to the frequentist approaches, general and specific equations for a crop. The Bayesian model was developed based on probability distributions of parameters (a and b) *a priori*, gathered from 18 studies, which estimated the linear relation (K_c -NDVI). In the first stage, the approaches were compared in seven sample sizes (TM) (5, 10, 30, 60, 90, 120 and 150), in which 1000 repetitions were carried out without replacements. In order to validate the adjustments of each TM, a base of 156 experimental data from K_c -NDVI were used. In the second stage, the EB was evaluated regarding equations adjusted to specific crops, and the first and second quartiles of the data were added to perform validation. Results first showed that the EB surpassed the frequentist methods and general equations in all TM. EB obtained low uncertainties with only five TM, unlike the other methods, which needed more than 30 records to obtain similar results. The comparison between EB and specific equations corroborated that five random data gathered in the first quartile were enough to obtain low uncertainties. This proposed methodology is operative and the K_c data can be estimated from the first phenological stages with a high certainty and few NDVI data.

Key words: probabilistic modeling, normalized difference vegetation index NDVI, water optimization in crops, crop coefficient, remote sensors.

^{*} Autor para correspondencia ♦ Author for correspondence.
Recibido: agosto, 2020. Aprobado: febrero, 2021.
Publicado como ARTÍCULO en *Agrociencia* **55**: 117-131. 2021.
DOI: <https://doi.org/10.47163/agrociencia.v55i2.2390>

Palabras clave: modelación probabilística, índice de vegetación de diferencia normalizada NDVI, optimización de agua en cultivos, coeficiente de cultivo, sensores remotos.

INTRODUCCIÓN

La agricultura es la actividad humana que usa más agua en el mundo, y la demanda actual de agua para uso agrícola en el mundo es insostenible. En México este sector usa el 77% del agua disponible. El caso es crítico en zonas agrícolas abastecidas por fuentes subterráneas, en donde la extracción de las aguas supera la recarga, lo que ha propiciado la sobreexplotación de 105 acuíferos, en su mayoría destinados para la agricultura (CONAGUA, 2018). La optimización del uso de agua de riego es imperiosa, por ello es necesario conocer con detalle la demanda hídrica de los cultivos (Hunsaker *et al.* 2005). Para estimar esta demanda, se requiere cuantificar la evapotranspiración real (ETa). Una manera de calcular la ETa es a partir de la metodología FAO-56, cuyo planteamiento general supone que la ETa está en función de la evapotranspiración de un cultivo de referencia (ETr) y un coeficiente de cultivo (Kc) (Calera *et al.* 2017).

La ETr se calcula con información de estaciones agrometeorológicas en condiciones de campo y Kc se calibra por medio de la razón entre ETa y ETr (Allen *et al.*, 2011). Sin embargo, la estimación de Kc en campo resulta costosa porque es necesaria una instrumentación especializada y puntual (torres de covarianza de torbellinos o lisímetros de pesada) (Pôças *et al.*, 2020), por lo cual es necesario recurrir a métodos operativos para mapear esta variable en áreas grandes (Palacios *et al.*, 2011).

La estimación de Kc se popularizó con el uso de sensores remotos, en particular por su relación con los índices de vegetación (IV) (Calera *et al.*, 2017). La modelación de Kc en función de IV es posible debido a las correlaciones altas entre estas variables y las correspondientes al desarrollo fenológico del cultivo y la reflectancia espectral (Rozenstein *et al.*, 2018).

Este enfoque simplifica la metodología de estimar Kc, porque no requiere del procesamiento de las bandas termal de un satélite y produce resultados similares a otras metodologías como el sistema de balance de energía multiescala en la superficie para el monitoreo operacional de la evapotranspiración superficial (en inglés, MSSEBS) (Colin *et al.*, 2006) y el método

INTRODUCTION

Agriculture is the human activity that uses the most water in the world, and the current demand for water for agricultural use is not sustainable. In Mexico, this sector uses 77% of the available water. The case is critical in agricultural areas supplied by underground water sources, where water extraction surpasses recharges, which has led to the overuse of 105 aquifers, destined mostly for agriculture (CONAGUA, 2018).

The optimization of the use of irrigation water is a pressing issue, why it is crucial to know the water demand of crops in detail (Hunsaker *et al.* 2005). In order to estimate this demand, it is necessary to quantify real evapotranspiration (ETa). One way to do this is using methodology FAO-56, the general approach of which supposes that ETa is based on the evapotranspiration of a reference crop (ETr) and a crop coefficient (Kc) (Calera *et al.* 2017).

ETr is estimated using information from agri-weather stations in field conditions and Kc is calibrated using the ratio between ETa and ETr (Allen *et al.*, 2011). However, estimating the Kc on the field is costly, since it requires a specialized and particular set of instruments (covariance flux towers or weighing lysimeters) (Pôças *et al.*, 2020). Therefore, resorting to operative methods becomes necessary to map this variable in large areas (Palacios *et al.*, 2011).

Kc estimation became popular with the use of remote sensors, particularly due to its relation with vegetation indices (IV) (Calera *et al.*, 2017). Modelling Kc based on IV is possible due to the high correlations between these variables, and those of crop phenological development, and spectral reflectance (Rozenstein *et al.*, 2018).

This approach simplifies the method to estimate Kc, since it does not require satellite thermal bands to be processed, and it produces similar results to other methodologies, such as the multiscale surface energy balancing system (MSSEBS) on the surface for the operational monitoring of surface evapotranspiration (Colin *et al.*, 2006), and the satellite-based energy balance method for Mapping Evapotranspiration with Internalized Calibration (METRIC™) (Allen *et al.*, 2011).

The normalized difference vegetation index (NDVI) and its variants are used to estimate the

de balance de energía con base en satélites para medir la evapotranspiración con calibración internalizada (en inglés, METRIC™) (Allen *et al.*, 2011).

El índice de vegetación de diferencia normalizada (en inglés, NDVI) y sus variantes se utilizan para estimar el valor de Kc en forma espacialmente explícita, y con este procedimiento es posible estimar la ETa del cultivo de manera operativa (Lei y Yang, 2014). Por esta razón, Montgomery *et al.* (2015) y Toureiro *et al.* (2017) propusieron modelos de Kc-IV para cultivos distintos, o el uso de esta metodología para estimar el consumo hídrico de los cultivos en otras regiones del mundo (Kamble *et al.*, 2013).

Los modelos se basan en una ecuación lineal empírica para cada cultivo; sin embargo, la limitante de usar ecuaciones generadas en condiciones geográficas y fisiográficas diferentes es que se producen variaciones grandes en sus parámetros (Calera *et al.* 2017). El uso de ecuaciones empíricas generadas en otras regiones puede causar una sobrestimación de los requerimientos hídricos de los cultivos y rompe el objetivo deseable de estas aplicaciones que es la optimización del agua aplicada para riego (Odi-Lara *et al.*, 2016).

El enfoque bayesiano (EB) tomó auge en las dos décadas recientes para solventar el problema de usar ecuaciones generales (EG) o ajustes por mínimos cuadrados ordinarios (MCO) en la estimación de variables ambientales (Weiskittel *et al.*, 2015). Su ventaja radica en que incorpora conocimiento *a priori* acerca de los datos, el cual las EG y ajustes de MCO ignoran por completo (Zhang *et al.*, 2013). Zapata-Cuartas *et al.* (2012) utilizaron EB con registros *a priori* de ajustes de parámetros para estimar biomasa en diferentes sitios del mundo, y mostraron que la incertidumbre en la estimación de biomasa era 300% menor, en comparación con el uso de EG y MCO. Wang *et al.* (2019) mostraron la eficiencia del EB para estimar biomasa a través de Lidar e indicaron que la incorporación de información generada *a priori*, ayudó a disminuir el número de muestras por recopilar.

Respecto al uso del EB en actividades agrícolas, Shirley *et al.* (2020) utilizaron esta metodología para estimar la producción de un cultivo a través de variables agroclimáticas y observaron que el enfoque es factible aun cuando existan datos limitados de variables independientes.

El objetivo de este estudio fue implementar el EB para estimar el Kc a partir del NDVI, y comparar

value of Kc in a spatially explicit way. By using this procedure, it is possible to operatively estimate the ETa of the crop (Lei and Yang, 2014). Due to this, Montgomery *et al.* (2015) and Toureiro *et al.* (2017) proposed Kc-IV models for different crops, or the use of this methodology to estimate the water consumption of crops in other areas of the world (Kamble *et al.*, 2013).

The models are based on an empirical linear equation for each crop, although the limitation of using equations based on different weather and physiographic conditions is that large variations are produced in their parameters (Calera *et al.* 2017). The use of empirical equations generated in other regions can cause an overestimation of the water requirements of crops, thus disregarding the desirable aim of these applications, which is optimizing irrigation water (Odi-Lara *et al.*, 2016).

The Bayesian approach (EB) reached new heights in the past two decades to solve the problem of using general equations (EG) or adjustments by ordinary least squares (MCO) to estimate environmental variables (Weiskittel *et al.*, 2015). Its advantage lies in the incorporation of prior knowledge on the data, which EG and MCO adjustments ignore altogether (Zhang *et al.*, 2013). Zapata-Cuartas *et al.* (2012) used EB with previous records of parameter adjustments to estimate biomass in different areas of the world, and showed that the uncertainty in the estimation of biomass was 300% lower, in comparison with the use of EG and MCO. Wang *et al.* (2019) showed the efficiency of EB to estimate the biomass via Lidar; they indicated that incorporating information produced *a priori* helped reducing the number of samples to be gathered.

Regarding the use of EB in agricultural activities, Shirley *et al.* (2020) used this method to estimate the production of a crop using agri-weather variables and they noticed that the approach is feasible, even when data from independent variables are limited.

The objective of this study was to implement EB to estimate the Ks from the NDVI, and to compare them with the frequentist approaches (EG and crop-specific). Our hypothesis was that the use of specific equations for a crop and a particular area produces high uncertainty when applied at places other than those in which they were originally produced. In addition, the information of the linear adjustments between Kc-IV provided by other research will

con los enfoques frecuentistas (EG y específicas para un cultivo). La hipótesis fue que el uso de ecuaciones específicas para un cultivo y para una zona en particular produce incertidumbre alta al aplicarse en sitios distintos a aquellos donde se generaron. Además, la información de los ajustes lineales entre Kc-IV aportados por otras investigaciones servirá como información *a priori* para ajustar el modelo bayesiano y permitirá reducir las incertidumbres para la estimación de Kc en cultivos diversos.

MATERIALES Y MÉTODOS

Marco bayesiano en la aplicación de Kc

La relación entre el Kc y el NDVI se puede expresar con la función siguiente:

$$Kc = f(NDVI|\theta) \quad (1)$$

donde θ son parámetros que se pueden obtener por medio de los métodos de mínimos cuadrados (MMC) o por la estimación de máxima verosimilitud (EMV) dentro de los enfoques frecuentistas. El marco bayesiano se expresa respecto a un vector de datos $y = (y_1, y_2, y_3 \dots y_n)$ y un vector de parámetros $\theta = (\theta_1, \theta_2, \theta_3 \dots \theta_n)$ de la manera siguiente:

$$p(y, \theta) = p(y | \theta) p(\theta) = p(\theta | y) p(y) \quad (2)$$

donde p representa una función de densidad de probabilidad, los valores de θ pueden obtenerse por medio de mínimos cuadrados u otro medio de ajuste de parámetros. En el contexto de este estudio θ tiene una distribución de probabilidad que puede actualizarse con algunos datos observados (IV):

$$\pi(\theta | Kc) = \frac{\pi(\theta) f(IV | \theta)}{\int \pi(\theta) f(IV | \theta) d\theta} \quad (3)$$

La distribución de probabilidad *a priori* de los parámetros para una ecuación está dada por $\pi(\theta)$; la distribución *a posteriori* que corresponde a las estimaciones nuevas del Kc se representa por $\pi(\theta)|Kc$. El uso del método bayesiano se propone para incorporar información *a priori* de los parámetros reportados en la literatura.

Recopilación de base de datos

La elección de las distribuciones *a priori* de los parámetros es crítica para el desarrollo del método bayesiano. El valor de los

serve as *a priori* information to adjust the Bayesian model and will help reducing uncertainties for the estimation of Kc in different crops.

MATERIALS AND METHODS

Bayesian framework in the application of Kc

The relation between Kc and NDVI can be expressed as follows:

$$Kc = f(NDVI|\theta) \quad (1)$$

where θ are parameters that can be obtained using the least squared methods (LSM) or the maximum likelihood estimate (MLE) within the frequentist approaches. The Bayesian framework is expressed in regard to a vector of data $y = (y_1, y_2, y_3 \dots y_n)$ and a vector of parameters $\theta = (\theta_1, \theta_2, \theta_3 \dots \theta_n)$ as follows:

$$p(y, \theta) = p(y | \theta) p(\theta) = p(\theta | y) p(y) \quad (2)$$

where p represents a function of density of probability, and the values of θ can be obtained using least squared tests or other parameter-adjustment methods. In the context of this study θ has a probability distribution that can be updated with some data observed (IV):

$$\pi(\theta | Kc) = \frac{\pi(\theta) f(IV | \theta)}{\int \pi(\theta) f(IV | \theta) d\theta} \quad (3)$$

The *a priori* distribution of the probability of the parameters for an equation is given by $\pi(\theta)$; the *a posteriori* distribution corresponding to the new estimations of the Kc is represented as $\pi(\theta)|Kc$. The use of the Bayesian method is proposed to incorporate *a priori* information of the parameters reported in the literature.

Database gathering

Choosing the *a priori* distributions of the parameters is critical to develop the Bayesian method. The values of parameters *a* and *b* were gathered from 18 scientific publications related to this topic (Table 1). The values of Kc vs. NDVI by Hunsaker *et al.* (2005) were extracted manually from that research to adjust the values in a linear regression, and to obtain the parameters in Table 1. The remaining values of these parameters were obtained from the adjustments performed by the authors.

For the gathered data sets, a linear model was adjusted to estimate Kc based on the NDVI. The Anderson-Darling test was

parámetros a y b se obtuvieron en 18 publicaciones científicas relacionadas con el tema (Cuadro 1). Los valores de Kc vs. NDVI de Hunsaker *et al.* (2005) se extrajeron en forma manual de esa investigación para ajustar los valores en una regresión lineal y así obtener los parámetros del Cuadro 1. Los valores restantes de estos parámetros se obtuvieron de los ajustes realizados por los autores.

A cada conjunto de datos recopilados se ajustó un modelo lineal para estimar Kc en función del NDVI. La prueba Anderson-Darling se utilizó para comprobar que el valor de ambos parámetros a y b estuvieran distribuidos normalmente. La media y desviación estándar de los parámetros del Cuadro 1 se usaron para generar una distribución de probabilidad normal y usar el modelo bayesiano.

Marco bayesiano e iteración con el método Montecarlo

El conocimiento informativo de los parámetros permite definir el dominio que deben tener, y la distribución de los parámetros puede generar miles de soluciones antes de converger en su distribución posterior. Para lograr esto último, es necesario utilizar métodos numéricos que involucren una gran cantidad de simulaciones para obtener datos confiables.

used to verify that the values of both parameters a and b were normally distributed. The mean and standard deviation of the parameters in Table 1 were used to produce a normal probability distribution in order to use the Bayesian model.

Bayesian framework and iteration with the Montecarlo method

The informing knowledge of the parameters helps defining the domain they should have, and the distribution of the parameters can generate thousands of solutions before converging in their later distribution. In order to achieve this, numeric methods involving a large amount of simulations must be used to obtain reliable data.

The procedure was programmed in the Python 2.7 language (Python Software Foundation, Amsterdam, The Netherlands), which is supported on the module PYMC3 (Salvatier *et al.*, 2016), containing specialized libraries on the implementation of Bayesian models. This module incorporates NUTS (Non-U-turn Sampler) method to update the parameters. The NUTS method is an extension of the Markov chain Montecarlo algorithm. The concept is simple, the more samples are iterated, the values of a *posteriori* parameters will converge in more precise values. The

Cuadro 1. Parámetros a y b recopilados para aplicar el enfoque bayesiano.

Table 1. Parameters a and b gathered in order to use the Bayesian approach.

Cultivo	a	b	Referencia
Trigo	-0.12	1.49	Hunsaker <i>et al.</i> (2005)
Algodón	-0.27	1.33	Rozenstein <i>et al.</i> (2018)
Algodón	-0.09	1.37	[†] Montgomery <i>et al.</i> (2015)
General [†]	0.17	1.46	Kamble <i>et al.</i> (2013)
General [†]	0.16	1.56	[†] Calera <i>et al.</i> (2003)
General [†]	0.20	1.25	D'Urso y Calera (2006)
Maíz	-0.05	1.09	Neale <i>et al.</i> (1989)
Maíz	-0.03	1.18	Neale <i>et al.</i> (1989)
Maíz	0.03	1.05	Bausch (1993)
Maíz	-0.01	1.08	Bausch (1993)
Maíz	-0.06	1.13	Bausch (1993)
Maíz	-0.11	1.19	Bausch (1993)
Maíz	0.30	0.92	Toureiro <i>et al.</i> (2017)
Maíz	-0.02	1.37	González-Piqueras <i>et al.</i> (2004)
Trigo	-0.26	1.49	Choudhury <i>et al.</i> (1994)
Trigo	-0.14	1.64	Duchemin <i>et al.</i> (2006)
Trigo	-0.18	1.74	Lei y Yang (2014)
Víñedo	-0.10	1.44	Campos <i>et al.</i> (2010)

[†]Ecuación general obtenida de parcelas con cultivos diferentes. ♦

[†]General equation obtained from plots with different crops.

El procedimiento se programó en el lenguaje Python 2.7 (Python Software Foundation, Amsterdam, The Netherlands), el cual se apoya del módulo PYMC3 (Salvatier *et al.*, 2016) que posee librerías especializadas en la implementación de modelos bayesianos. Dicho módulo incorpora el método NUTS (Non-U-turn Sampler o muestreador sin retornos), para actualizar los parámetros. El método NUTS es una extensión del algoritmo Montecarlo de cadenas de Markov. El concepto es simple, a medida que se iteren más muestras, los valores de los parámetros *a posteriori* convergerán en valores más precisos. El programa se ajustó para que se realizaran 25000 iteraciones del muestreador NUTS, para asegurar una convergencia máxima de los parámetros estimados (el código de programación puede solicitarse a los autores de esta investigación).

La Figura 1 presenta el esquema de ajuste de parámetros con el esquema bayesiano; en la primera parte se define la media (μ) y desviación estándar (σ) de los parámetros *a* y *b* que se tomaron del Cuadro 1. El modelo lineal generalizado (MLG) empleado para ajustar parámetros, supone una distribución normal bivariada en los dos parámetros. Otro supuesto es que los errores están distribuidos normalmente e $\sim N(0, \sigma^2)$. El MLG, además de la distribución de los parámetros *a priori*, usa los datos de NDVI para obtener una actualización óptima de los parámetros *a posteriori* y al mismo tiempo una estimación del coeficiente de cultivo.

Datos reales de Kc versus NDVI

La incertidumbre del modelo se evaluó con 156 datos pareados de valores de Kc y NDVI, provenientes de evaluaciones en

program was adjusted for 25 000 iterations of the NUTS sampler to be carried out, in order to ensure the maximum convergence of the estimated parameters (the adjustment programming code is available upon request to authors).

Figure 1 presents the parameter adjustment scheme with the Bayesian scheme; the first part defines the mean (μ) and standard deviation (σ) of parameters *a* and *b* taken from Table 1. The generalized linear model (GLM) used to adjust parameters assumed a bivariate normal distribution in both parameters. Another assumption is that the errors are normally distributed e $\sim N(0, \sigma^2)$. The GLM, not only uses the distribution of the *a priori* parameters, but also the NDVI data to obtain an optimum update of the *a posteriori* parameters, and in the same run, it renders estimation of the crop coefficient.

Real data Kc versus NDVI

The uncertainty of the model was evaluated with 156 paired data values of Kc and NDVI, from evaluations on maize (Choudhury *et al.*, 1994; Toureiro *et al.*, 2017), wheat (Duchemin *et al.*, 2006), and maize, soybean and grass (Kamble *et al.*, 2013; Rozenstein *et al.*, 2018). The data were obtained from samples of dispersal diagrams of Kc and NDVI relations from the aforementioned studies to validate our proposal in this study.

Kc values were obtained from covariance flux towers, soil water balancing methods, and weighing lysimeters; therefore they were considered real field data. For each record of Kc data obtained, we recorded the NDVI value from the processing of digital multispectral images or samples with a radiometer.

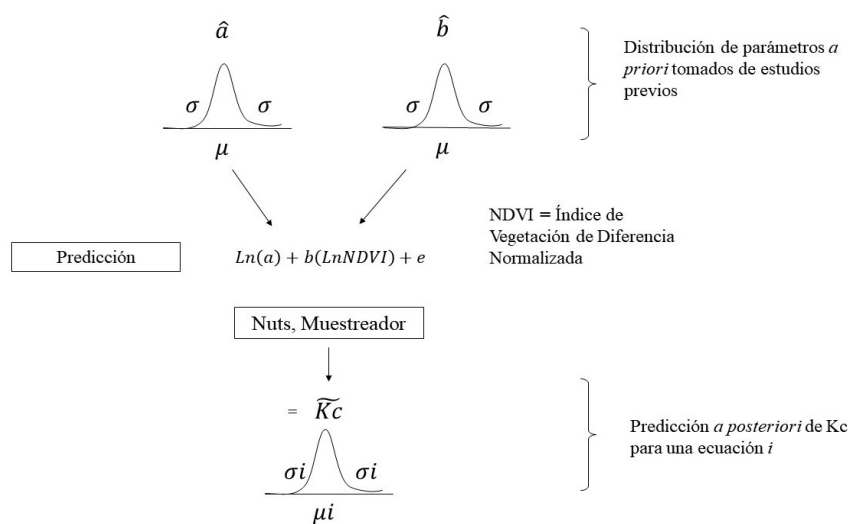


Figura 1. Esquema de la aplicación del enfoque bayesiano para predecir la distribución *a posteriori* y estimar el coeficiente de cultivo.

Figure 1. Application scheme of the Bayesian approach to predict the *a posteriori* distribution, estimating the crop coefficient.

cultivos de maíz (Choudhury *et al.*, 1994; Toureiro *et al.*, 2017), trigo (Duchemin *et al.*, 2006) y maíz, soya y pasto (Kamble *et al.*, 2013; Rozenstein *et al.*, 2018). Los datos se obtuvieron de muestras de los diagramas de dispersión de la relación Kc y NDVI de dichas investigaciones para hacer la validación de la propuesta de este estudio.

Los valores de Kc se obtuvieron de torres de covarianza de torbellinos, métodos de balance de agua del suelo y por lisímetros; por lo tanto, se consideran como datos reales de campo. Para cada dato de Kc obtenido, los autores registraron el valor de NDVI procedente del procesamiento de imágenes digitales multiespectrales o muestras con radiómetro.

El resumen de los 156 registros de Kc y NDVI se presenta en el Cuadro 2. Los datos se obtuvieron en condiciones diferentes de ubicación geográfica, manejo, cultivo y ambientales, y se incluye la cita correspondiente para indagar más sobre cada uno de los estudios.

Evaluación de los modelos

Comparación entre enfoque bayesiano, mínimos cuadrados ordinarios y ecuaciones generales

Para la comparación se realizó una evaluación del EB, mínimos cuadrados ordinarios (MCO) y las ecuaciones generales propuestas por D'urso y Calera (2006), $Kc = 0.20 + 1.25 * NDVI$, y por Kamble *et al.* (2013), $Kc = 0.17 + 1.46 * NDVI$, en los 156 datos de Kc y NDVI ya descritos. Un manejador de base datos tipo SQL (SQLite v. 3.0) se utilizó para definir los tamaños de muestra (TM) de 5, 10, 30, 60, 90, 120 y 150 registros. Cada TM se probó aleatoriamente, sin remplazo, 1000 veces y se ajustaron a los enfoques propuestos.

Los distintos TM permitieron calcular el Kc. En el caso del EB las simulaciones facilitaron la obtención de los parámetros a

Those paired 156 Kc and NDVI records are summarized in Table 2. Data were obtained under different geographic locations, management, crop and environmental conditions. The corresponding reference is included for further information about those studies.

Evaluation of the models

Comparison between the Bayesian approach, ordinary least squares and general equations

For the comparison, the EB was evaluated along with the ordinary least squares (MCO) and the general equations proposed by D'urso and Calera (2006), $Kc = 0.20 + 1.25 * NDVI$, and by Kamble *et al.* (2013), $Kc = 0.17 + 1.46 * NDVI$, in the 156 Kc and NDVI paired data described paragraphs above. An SQL (SQLite, v. 3.0) database manager was used to define the sample sizes (TM) of 5, 10, 30, 60, 90, 120 and 150 records. Each TM was tested at random, without replacements, 1000 times, and it was adjusted to the suggested approaches.

The different TMs helped calculating the Kc. In the case of EB, the simulations helped obtaining those *a posteriori* parameters that helped to measure the efficiency of estimation, and the respective parameters for MCO were also adjusted. In all cases, the mean squared error (ECM) and the mean absolute error (EMA) were used to estimate error in the different sample sizes. The coefficient of variation (Cv) was also obtained to observe the homogeneity of the error in each TM.

Comparison between the Bayesian approach and specific equations for wheat and maize

Another comparison between specific formulas was carried out for a crop and EB. The Kc *versus* NDVI data by Duchemin

Cuadro 2. Datos de Kc y NDVI usados para realizar las evaluaciones del enfoque bayesiano y formulas generales.

Table 2. Kc and NDVI data used to carry out the evaluations of the Bayesian approach and general formulae.

Cultivo	NDVI		Kc		n	Referencia
	Media	SEE	Media	SEE		
Algodón	0.820	0.070	0.826	0.096	5	Rozenstein <i>et al.</i> (2018)
Maíz	0.509	0.257	0.527	0.389	62	Choudhury <i>et al.</i> (1994)
Maíz, soya, pasto	0.662	0.225	0.800	0.318	23	Kamble <i>et al.</i> (2013)
Maíz	0.659	0.175	0.906	0.182	21	Toureiro <i>et al.</i> (2017)
Trigo	0.516	0.229	0.724	0.299	45	Duchemin <i>et al.</i> (2006)

SEE: Desviación estándar; n: número de registros. ♦ SEE: Standard deviation; n: number of records.

posteriori que sirvieron para medir la eficiencia de la estimación, y para MCO también se ajustaron sus parámetros respectivos. En todos los casos se usó el error cuadrático medio (ECM) y el error medio absoluto (EMA) para estimar el error en los tamaños de muestra diferentes. También se obtuvo el coeficiente de variación (Cv) para observar la homogeneidad del error en cada TM.

Comparación entre enfoque bayesiano y ecuaciones específicas para los cultivos de trigo y maíz

Otra comparación entre fórmulas específicas se realizó para un cultivo y el EB. Los datos Kc *versus* NDVI, de Duchemin (2006) para trigo y Kamble *et al.* (2013) para maíz, se usaron como datos reales. Para estimar Kc en cada caso, se evaluaron las fórmulas de Lei y Yang (2014) para trigo y de Toureiro *et al.* (2017) para maíz. Para el trigo se definieron los TM: 5, 10, 20, 40 y el primer y segundo cuartil de los datos del NDVI. Los TM para el maíz fueron: 5, 10, 20 y el primer y segundo cuartil de los datos de NDVI. En los cuartiles solo se eligió un TM de cinco registros aleatorios sin remplazo para hacer la evaluación. La eficiencia de la estimación se realizó con el error medio absoluto (EMA) y el error cuadrático medio (ECM).

RESULTADOS Y DISCUSIÓN

Ajuste de los parámetros del enfoque bayesiano y ajustes por mínimos cuadrados ordinarios

Los parámetros a y b para calcular Kc (Cuadro 1) mostraron una correlación negativa ($r = -0.52$, $p \leq 0.05$), y los resultados de la prueba de normalidad revelaron que ambos parámetros tienden a un comportamiento normal ($p \leq 0.05$). Los valores de la media y desviación estándar introducidos al modelo bayesiano fueron -0.064 y 0.149 para el parámetro a , mientras que para el parámetro b los valores fueron 1.321 y 0.218 .

En la estimación de los parámetros *a posteriori* realizados con las 1000 simulaciones para cada tamaño de muestra, se observa que con un TM mayor los parámetros son más compactos en sus valores (Figura 2). En ambos parámetros el EB mostró homogeneidad mayor, mientras que los valores de MCO presentaron datos extremos conforme el TM descendía. Según Grzenda (2015) estos resultados se deben a que la integración del método Montecarlo permitió realizar las simulaciones con tamaños de muestra diferentes, lo cual eliminó ejecuciones de ajuste e intervenciones imprácticas del usuario.

(2006) for wheat, and Kamble *et al.* (2013) for maize, were used as real data. To estimate Kc in each case, the formulas by Lei and Yang (2014) for wheat, and Toureiro *et al.* (2017) for maize, were evaluated. For wheat, 5, 10, 20, 40 TM were defined, along with the first and second quartile of the NDVI data. TM for maize were 5, 10, 20 records, and the first and second quartile of the NDVI data. In quartiles, only the TM with five random records without replacements was chosen to perform the evaluation. The efficiency of the estimation was evaluated with the mean absolute error (EMA) and the mean squared error (ECM).

RESULTS AND DISCUSSION

Adjustment of the parameters in the Bayesian approach and adjustments by ordinary least squares

Parameters a and b to calculate Kc (Table 1) displayed a negative correlation ($r = -0.52$, $p \leq 0.05$); and the results of the normality test revealed that both parameters tend towards a normal distribution ($p \leq 0.05$). The values of the mean and standard deviation used as inputs in the Bayesian model were -0.064 and 0.149 for parameter a ; whereas for parameter b , the values were 1.321 and 0.218 .

The estimation of *a posteriori* parameters from the 1000 simulations for each sample size shows that with a greater TM, parameter values are more compact (Figure 2). In both parameters, EB displayed greater homogeneity, while the MCO values displayed extreme data as TM decreased. According to Grzenda (2015), these results are due to the integration of the Montecarlo method allowing for simulations with different sample sizes, which eliminated adjustment executions, and impractical interventions by the user.

Kruschke (2014) described that the EB recovers an interval of inferential solutions through previous knowledge and local data; rather than an accurate estimation and empirical adjustments, as in the adjustment regression by least squares.

Validation of methods proposed for Kc estimation

Bayesian approach, general equations and the ordinary least squares method

The 1000 simulations showed that reducing the sample size leads to an increased error, as shown by the EMA and the ECM (Figure 3). The EB and MCO

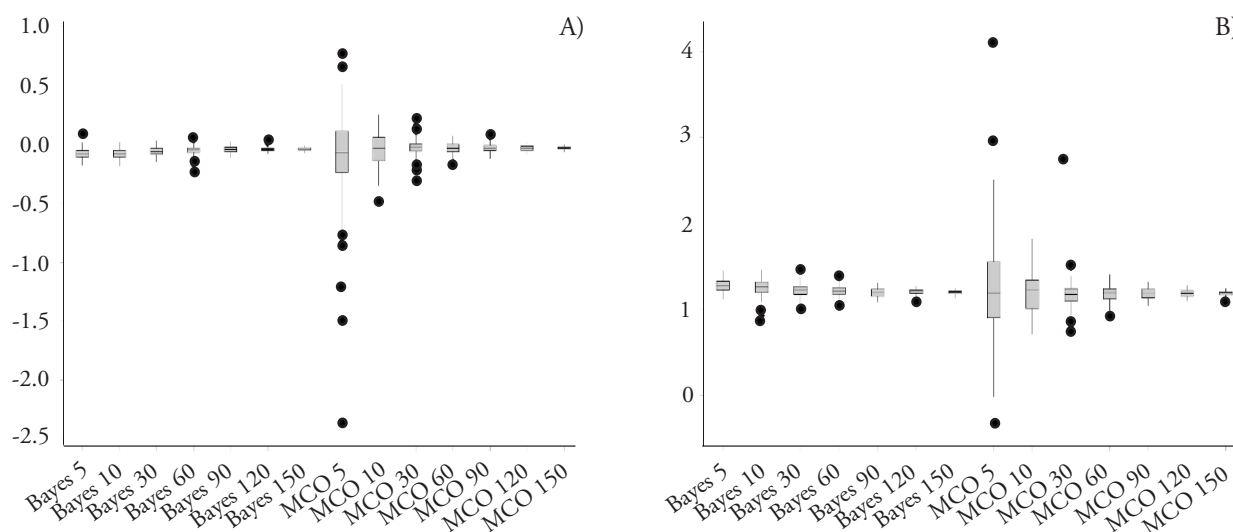


Figura 2. Resultados del ajuste de los parámetros para las 1000 repeticiones en tamaño de muestra diferente. El gráfico de la izquierda son los valores para el parámetro A) y el de la derecha para el parámetro B). Las cajas muestran el rango intercuartílico de los datos para cada TM.

Figure 2. Results of the parameters adjustment for 1000 repetitions in different sample sizes. The graph on the left shows the values for the parameter A) and the graph on the right, for the parameter B). Boxes show the interquartile range of the data for each TM.

Kruschke (2014) describió que el EB recupera un intervalo de soluciones inferenciales a través de conocimiento previo y datos locales; en lugar de una estimación puntual y ajustes empíricos, como sí ocurre en la regresión de ajuste por mínimos cuadrados.

Validación de métodos propuestos para estimar Kc

Enfoque bayesiano, ecuaciones generales y método de mínimos cuadrados

Las 1000 simulaciones mostraron que al disminuir el tamaño de muestra, el error tiende a aumentar, como demuestran el EMA y el ECM (Figura 3). El EB y MCO presentaron un comportamiento homogéneo del error hasta el TM 10. Sin embargo, cuando el TM = 5 el Cv del MCO fue 36.07% mientras que para el EB fue 6.92%.

El MCO mostró un valor máximo de 0.74 de EMA y 0.82 de ECM en el TM=5. Este error fue el más alto de las 1000 simulaciones en los cuatro métodos, lo cual indica que este procedimiento es sensible al error en tamaños de muestra pequeños. Estos resultados son similares a lo documentado por Zhang *et al.* (2013), quienes observaron que el EB resultó

presented a homogenous behavior of the error up to TM 10. However, when TM = 5 the Cv of MCO was 36.07%, whereas it was 6.92% for the EB.

MCO displayed a maximum value of 0.74 for EMA, and 0.82 for ECM when TM=5. This error was the highest out of the 1000 simulations in the four methods, which indicates that this procedure is sensitive to the error in small sample sizes. These results are similar to reports by Zhang *et al.* (2013), who observed that the EB resulted in a lower uncertainty when evaluating the biomass of forest species with different TM. The adaptability of the EB is confirmed by a reduction of uncertainty due to the use of *a priori* probability distributions, since EB helps achieving better results in comparison with the adjustment of frequentist approaches, or general equations because it is integrated on few local data (Sileshi, 2014).

The models by D'urso and Calera (2006) and Kamble *et al.* (2013) displayed high errors in the mean error, with an average of 0.23 and 0.34 of EMA for all sample sizes, and they displayed a heterogenous behavior of the error and high Cv, starting at TM=30. However, the robustness of the EB for each sample size can be observed through the EMA statistics for each method (Table 3). This represents the highest

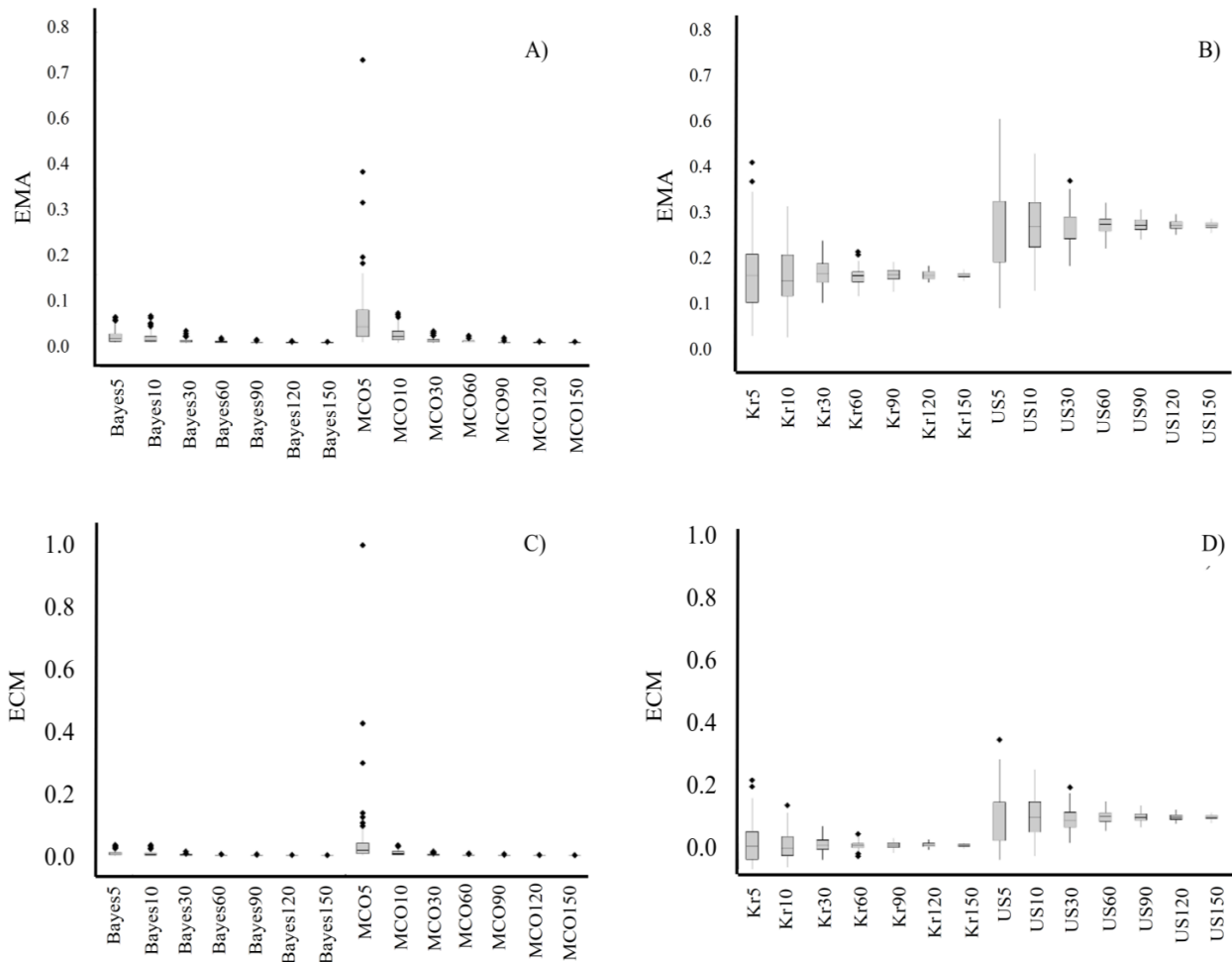


Figura 3. A) Error medio absoluto (EMA) para enfoque bayesiano y mínimos cuadrados ordinarios (MCO) y B) EMA para modelos de Kamble (Kr) y D'urso (Us). C) Error cuadrático medio (ECM) para enfoque bayesiano y mínimos cuadrados ordinarios; D) ECM para modelos de Kramble y D'urso.

Figure 3. A) Mean absolute error (EMA) for the Bayesian approach, and ordinary least squares (MCO); B) EMA for the Kamble (Kr) and D'urso (Us) models; C) Mean squared error (ECM) for the Bayesian approach and ordinary least squares; D) ECM for the Kramble and D'urso models.

con incertidumbre menor al evaluar la biomasa de especies forestales con TM diferente. La adaptabilidad del EB se refleja en un decremento de la incertidumbre, debida al uso de distribuciones de probabilidad *a priori*, porque al integrarse con pocos datos locales permite lograr resultados mejores en comparación con el ajuste de enfoques frecuentistas o ecuaciones generales (Sileshi, 2014).

Los modelos de D'urso y Calera (2006) y Kamble *et al.* (2013) mostraron errores altos en la media del error, con un promedio de 0.23 y 0.34 de EMA para todos los tamaños de muestra, y presentaron un comportamiento heterogéneo del error y Cv altos a partir del TM=30. Pero la robustez del EB para cada

homogeneity at the estimation of mean Kc. In addition, SEE, Cv, minimum (Min), and maximum (Max) of the repetitions were included, since they are important when comparing the variation of error in all four methods (Table 3).

The models by Kamble and D'urso, and Calera require 60 records of Kc *versus* NDVI to obtain the same error variation than the EB in TM = 5, whereas MCO needs 10 pieces of data. Regarding the maximum error that the methods can produce, the EB only increases one EMA of 0.05 from an TM = 150 to a TM = 5. This is significant because it may lead to more accurate estimations of Kc value with few values obtained through field research.

tamaño de muestra se puede observar a través de los estadísticos del EMA para cada método (Cuadro 3). Esto representa una homogeneidad mayor de la media estimada de Kc. Además, se incluyeron la SEE, el Cv, los valores mínimos (Min) y máximos (Max) de las repeticiones, porque resultan importantes al comparar la variación del error en los cuatro métodos (Cuadro 3).

Los modelos de Kamble y D'urso, y Calera necesitan 60 registros de Kc *versus* NDVI para obtener la misma variación del error que el EB en TM = 5, mientras que el MCO necesita 10 datos. Respecto al el error máximo que pueden producir los métodos, el EB solo aumenta un EMA de 0.05 desde un TM = 150 a un TM = 5. Esto es significativo porque puede llevar a estimaciones más precisas del valor de Kc con pocos valores obtenidos en campo.

Nuestro estudio demostró que no es viable usar una ecuación general para estimar Kc, contrario a las recomendaciones de Palacios *et al.* (2011) y Kamble *et al.* (2013). La elección de un modelo empírico es la fuente mayor de incertidumbre (Sileshi, 2014). Por lo tanto, las investigaciones de predicción deberían evitar enfocarse solo en la medida del ajuste del modelo y en observar más la habilidad predictiva del mismo (Zapata-Cuarta *et al.*, 2012).

Our study showed that using a general equation to estimate Ks is not feasible, as opposed to the recommendations by Palacios *et al.* (2011) and Kamble *et al.* (2013). The choice of an empirical model is the greatest source of uncertainty (Sileshi, 2014). Therefore, prediction research should avoid focusing on the measure of models adjustment, and rather observing their predictive ability (Zapata-Cuarta *et al.*, 2012).

Calera *et al.* (2017) pointed out the weaknesses of the Kc-NDVI relation, when adjusted with the frequentist approach. They mentioned that the uncertainties in this relation may be due to the different sensor measurements of NDVI, and differences on image processing. Also, the physiographic and hydrological differences of the sites in which the experiments were implemented (Pôças *et al.*, 2020).

The proposal of this investigation follows the recommendations by Weiskittel *et al.* (2015), which consist in maximizing the existing information to improve the predictions of the crop coefficient. By doing this, the modelling can be used under any management condition or phenological stage, adapting to the differences of the sensor used to generate the indices.

Cuadro 3. Estadísticos del Kc para los cuatro enfoques propuestos en tamaño de muestra diferente.
Table 3. Kc statistics for the four approaches proposed in different samples sizes.

Método	Estimador	5	10	30	60	90	120	150
Enfoque bayesiano	Media	0.172	0.170	0.165	0.162	0.162	0.161	0.161
	SEE	0.012	0.011	0.004	0.002	0.001	0.001	0
	Cv	6.92	6.28	2.47	1.08	0.76	0.42	0.26
	Min	0.161	0.161	0.161	0.161	0.161	0.161	0.161
	Max	0.213	0.216	0.185	0.17	0.167	0.164	0.162
Mínimos cuadrados	Media	0.214	0.177	0.165	0.163	0.162	0.161	0.161
	SEE	0.077	0.013	0.005	0.003	0.001	0	0
	Cv	36.07	7.59	3.07	1.68	0.8	0.29	0.22
	Min	0.161	0.161	0.161	0.161	0.161	0.161	0.161
	Max	0.743	0.221	0.184	0.175	0.17	0.163	0.162
Kramble <i>et al.</i> (2013)	Media	0.236	0.232	0.237	0.231	0.234	0.233	0.234
	SEE	0.08	0.063	0.03	0.017	0.015	0.01	0.006
	Cv	33.88	27.22	12.85	7.33	6.22	4.36	2.45
	Min	0.092	0.088	0.169	0.184	0.194	0.216	0.22
	Max	0.496	0.393	0.314	0.288	0.264	0.255	0.246
D'Urso y Calera (2006)	Media	0.346	0.351	0.343	0.35	0.35	0.35	0.35
	SEE	0.091	0.066	0.037	0.023	0.016	0.01	0.008
	Cv	26.34	18.85	10.73	6.5	4.49	2.79	2.18
	Min	0.157	0.197	0.255	0.296	0.316	0.328	0.331
	Max	0.597	0.517	0.453	0.401	0.387	0.375	0.365

Calera *et al.* (2017) señalaron las debilidades de la relación Kc-NDVI ajustadas con el enfoque frecuentista; mencionaron que las incertidumbres asociadas con esta relación pueden deberse a las mediciones de NDVI con sensor y procesamiento de la imagen diferentes; además de las diferencias fisiográficas e hidrológicas de los sitios donde se llevaron a cabo los experimentos de los cultivos (Pôças *et al.*, 2020).

La propuesta de esta investigación sigue las recomendaciones de Weiskittel *et al.* (2015), que consiste en maximizar la información existente para mejorar las predicciones del coeficiente del cultivo y que la modelación se pueda utilizar en cualquier condición de manejo y etapa fenológica, además de adecuarse a las diferencias del sensor usado para generar los índices.

Validación en datos de maíz y trigo

Los estadísticos asociados con el método bayesiano y la función de Toreiro *et al.* (2017) para estimar Kc del maíz mostraron que el modelo bayesiano presentó incertidumbre menor en los TM y cuartiles empleados (Cuadro 4). En promedio, el EB presentó un valor medio de 0.10 y la desviación estándar de 0.01, mientras en el modelo de Toureiro el promedio fue 0.19 y 0.02 de desviación estándar. El Cv y los valores extremos fueron más bajos en el modelo bayesiano para todos los TM.

Los dos métodos mostraron que el primer cuartil obtuvo un Cv menor en comparación con el TM=5 y el segundo cuartil; sin embargo, el EB mostró certidumbre mayor. Esto es muy importante para los enfoques operacionales porque muestra que con cinco

Validation in maize and wheat data

The statistics associated with the Bayesian approach and the function by Toreiro *et al.* (2017) to estimate the Kc of maize showed that the Bayesian model displayed lower uncertainty in the TM and quartiles used (Table 4). On average, EB presented a mean value of 0.10 and a standard deviation of 0.01; whereas in the model by Toureiro, the average was 0.19 with a standard deviation of 0.02. The Cv and the extreme values were lower in the Bayesian model for all TM.

Both methods showed that the first quartile obtained a lower Cv in comparison with TM=5 and the second quartile, although EB displayed a higher certainty. This is very important for operational approaches, since it shows that with five records of Kc and NDVI obtained in the first stages of the established crop, it is possible to obtain reliable data to be able to estimate Kc. The results for wheat, between the Bayesian approach, and the function by Lei y Yang (2014) showed for 40 records (Table 5), that EB presented a Cv of 0.02%, whereas the model by Lei and Yang displayed a Cv of 5.45%.

For five records, EB displayed a Cv of 6.36% in comparison with 40% of the model by Lei and Yang. It is worth emphasizing that the EB with five records showed such a Cv value; whereas, in order to reach a similar variation, the model by Lei and Yang required 40 records. That is, nearly the 43 records which were the total size of the population used in that analysis. In the first quartile, the EB was 25.7% more homogenous than the method by Lei and Yang; whereas in the second quartile, it was 17.49%

Cuadro 4. Estadísticos obtenidos en tamaño de muestra diferente al evaluar la estimación de Kc para cultivos de maíz. 1Q (primer cuartil), 2Q (segundo cuartil).

Table 4. Statistics obtained in different sample sizes when evaluating the estimation of Kc for maize. 1Q (first quartile), 2Q (second quartile).

Método	Estimador	5	10	20	1Q	2Q
Enfoque bayesiano	Media	0.09	0.10	0.10	0.10	0.11
	SEE	0.02	0.01	0.00	0.01	0.01
	Max	0.18	0.13	0.10	0.13	0.15
	Cv	25.0	7.29	1.04	10.4	14.5
Toureiro <i>et al.</i> (2017)	Media	0.15	0.15	0.15	0.26	0.23
	SEE	0.04	0.02	0.00	0.01	0.02
	Max	0.26	0.22	0.17	0.28	0.29
	Cv	41.6	26.0	8.33	16.6	30.2

registros de Kc y NDVI obtenidos en las primeras etapas del cultivo establecido, se pueden obtener datos confiables para poder estimar Kc. Los resultados para el cultivo de trigo entre el enfoque bayesiano y la función de Lei y Yang (2014) mostraron en 40 registros (Cuadro 5) que el EB presentó un Cv de 0.02%, mientras que el modelo de Lei y Yang mostró un Cv de 5.45%.

El EB con cinco registros mostró un Cv de 6.36%, y es importante de resaltar comparado con el 40% del modelo de Lei y Yang; porque EB con solo cinco datos mostró un valor de Cv para el cual, el modelo de Lei y Yang necesitó 40 registros para alcanzar una variación similar. Esto representa, casi los 43 registros del total del tamaño de la población usada para ese análisis. En el primer cuartil, el EB fue 25.7% más homogéneo que el de Lei y Yang; mientras que en el segundo cuartil fue 17.49% más homogéneo. En el EB los Cv del primero y segundo cuartil fueron menores que 10%, lo cual indica que los errores fueron homogéneos. En el caso del modelo de Lei y Yang, los Cv del primero y segundo cuartil mostraron ser heterogéneos, con valores de 31.82 y 36.36%, respectivamente. Lo anterior confirma la robustez del EB para reducir la incertidumbre del error y evitar datos extremos que subestimen o sobrestimen el dato real de Kc.

Los resultados mostraron la misma tendencia que en el apartado anterior, el EB bayesiano presentó incertidumbre menor que las ecuaciones de Lei y Yang (2014) para trigo y Toureiro *et al.* (2017) para maíz, en cada TM. Por lo tanto, el EB se convierte en un método operativo que puede usar pocas imágenes satelitales dentro de un ciclo agrícola, ya que se pueden

more homogenous. In the EB, the Cv of the first and second quartiles were below 10%, indicating that errors were homogenous. In the case of the model by Lei and Yang, the Cv of the first and second quartile displayed heterogeneity, with values of 31.82 and 36.36%, respectively. This confirms the robustness of the EB to reduce the uncertainty of error, and to avoid extreme data that may underestimate or overestimate the real Kc data.

The results showed the same tendency than in the previous section; the EB showed a lower uncertainty than in the equations by Lei and Yang (2014) for wheat, and by Toureiro *et al.* (2017) for maize, in each TM. Therefore, the EB becomes an operative method that can use few satellite images within a single agricultural cycle, since images in some dates with cloudy skies can be omitted. The reduction of uncertainty signifies large savings of water volumes when irrigating, and becomes an inclusive methodology for precision agriculture (Calera *et al.*, 2017). In addition, EB can be applied in regions in which users lack financial resources to specifically calibrate the Kc-NDVI relation.

A limitation of this study is that the parameters used to analyze the Bayesian approach came from bibliographical sources that did not give details in the information on irrigation management, or on the color of soils in the experiments carried out to generate the parameters of the Kc(NDVI) functions. Due to this, it is unknown whether the crop coefficient was obtained under stress-free conditions (base crop coefficient or Kcb) or in field conditions (Kc) usually under stress. These crop management and field conditions should be considered when replicating the technique for operational use.

Cuadro 5. Estadísticos obtenidos en tamaño de muestra diferente para evaluar la estimación de Kc para datos de trigo.

Table 5. Statistics obtained in different sample sizes to evaluate the Kc estimation for wheat.

Método	Estimador	5	10	20	40	1Q	2Q
Enfoque bayesiano	Media	0.12	0.11	0.11	0.11	0.12	0.11
	StdDev	0.01	0.00	0.00	0.00	0.01	0.01
	Max	0.15	0.13	0.12	0.11	0.16	0.15
	Cv	6.36	3.64	3.64	0.02	8.18	6.36
Lei y Yang (2014)	Media	0.14	0.15	0.14	0.14	0.19	0.13
	StdDev	0.04	0.03	0.02	0.01	0.04	0.04
	Max	0.28	0.21	0.17	0.16	0.28	0.23
	Cv	40.00	26.36	14.55	5.45	31.82	36.36

omitir algunas fechas que presenten nubosidad. La reducción de la incertidumbre se traduce en ahorros grandes de volúmenes de agua al irrigar y se convierte en una metodología incluyente para la agricultura de precisión (Calera *et al.*, 2017). Además, EB se puede aplicar en regiones donde los usuarios no tienen los recursos financieros para poder calibrar la relación Kc-NDVI en específico.

Una limitante de este estudio es que los parámetros usados para analizar el enfoque bayesiano provinieron de fuentes bibliográficas que no indicaron información detallada sobre el manejo del riego ni el color del suelo en los experimentos realizados para generar los parámetros de las funciones Kc(NDVI). Por ello, no se conoce si el coeficiente del cultivo se obtuvo bajo condiciones sin estrés (coeficiente de cultivo basal o Kcb) o en condiciones de campo (Kc) usualmente bajo estrés. Estas condiciones de manejo del cultivo y de la parcela se deberán considerar al replicar la técnica para uso operacional.

CONCLUSIONES

El enfoque bayesiano utilizado redujo la incertidumbre asociada con la estimación del coeficiente de cultivo a través de índices de vegetación. Las simulaciones creadas permitieron observar el error de estimación en tamaño de muestra diferente. En cada una de ellas el enfoque bayesiano mostró ser la metodología con el error menor en comparación con el ajuste por mínimos cuadrados o por formulas generales. La comparación del enfoque bayesiano con ecuaciones específicas mostró incertidumbre menor para los cultivos de maíz y trigo, en todos los tamaños de muestra.

En la evaluación del primer y segundo cuartiles de los datos, el enfoque bayesiano demostró que con solo cinco datos aleatorios dentro del primer cuartil es suficiente para obtener estimaciones de incertidumbre baja. La información recopilada *a priori* permitió desarrollar ecuaciones nuevas que mejoran las estimaciones del coeficiente de cultivo en comparación con las estimaciones frecuentistas usadas de manera habitual.

LITERATURA CITADA

Allen, R. G., L. S. Pereira, T. A. Howell, and M. E. Jensen. 2011. Evapotranspiration information reporting: I. Factors governing measurement accuracy. *Agric. Water Manage.* 98: 899-920. DOI: 10.1016/j.agwat.2010.12.015.

CONCLUSIONS

The Bayesian approach reduced the uncertainty related to the estimation of the crop coefficient through vegetation indices. The created simulations helped to accurately observe the estimation error in different sample sizes. In each one of them, the Bayesian approach proved to be the methodology with the lowest error, in comparison with the adjustment by least squares or by general formulas. The comparison of the Bayesian approach with specific equations presented less uncertainty for maize and wheat crops in all sample sizes.

In the evaluation of the first and second data quartiles, the Bayesian approach proved that only five random data in the first quartile are enough to obtain low uncertainty estimations. The information gathered *a priori* helped to develop new equations that improve estimations of that coefficient in comparison with the frequentist estimations regularly used.

—End of the English version—

-----*

- Bausch, W. C. 1993. Soil background effects on reflectance-based crop coefficients for corn. *Rem. Sens. Environ.* 46: 213-222. DOI: 10.1016/0034-4257(93)90096-G.
- Campos, I., C. M. Neale, A. Calera, C. Balbontín, and J. González-Piqueras. 2010. Assessing satellite-based basal crop coefficients for irrigated grapes (*Vitis vinifera* L.). *Agric. Water Manage.* 98: 45-54. DOI: 10.1016/j.agwat.2010.07.011.
- Calera, A., I. Campos, A. Osann, G. D'Urso, and M. Menenti. 2017. Remote sensing for crop water management: from ET modelling to services for the end users. *Sensors* 17: 1104. DOI: 10.3390/s17051104.
- Colin, J., M. Menenti, E. Rubio, and A. Jochum. 2006. A Multi-Scales Surface Energy Balance System for Operational Actual Surface Evapotranspiration Monitoring. DOI: 10.1063/1.2349342.
- CONAGUA (Comisión Nacional del Agua). 2018. Estadísticas del agua en México: Comisión Nacional del Agua Ciudad de México, México. http://sina.conagua.gob.mx/publicaciones/EAM_2018.pdf (Consulta: agosto 2019).
- Choudhury, B. J., N. U. Ahmed, S. B. Idso, R. J. Reginato, and C. S. Daughtry. 1994. Relations between evaporation coefficients and vegetation indices studied by model simulations. *Rem. Sens. Environ.* 50: 1-17. DOI: 10.1016/0034-4257(94)90090-6.
- D'Urso, G., and A. Calera Belmonte. 2006. Operative approaches to determine crop water requirements from Earth Observation data: methodologies and applications. DOI: 10.1063/1.2349323.

- Duchemin, B., R. Hadria, S. Erraki, G. Boulet, P. Maisongrande, A. Chehbouni, and M. Kharrou. 2006. Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices. *Agric. Water Manage.* 79: 1-27. DOI: 10.1016/j.agwat.2005.02.013.
- González-Piqueras, J., A. Calera, M. A. Gilabert, A. Cuesta, and F. De la Cruz Tercero. 2004. Estimation of crop coefficients by means of optimized vegetation indices for corn. *Rem. Sens. Agric. Ecosys. Hydrol.* Vol. 5232- pp. 110-118. DOI: 10.1117/12.511317.
- Grzenda, W. 2015. The advantages of Bayesian methods over classical methods in the context of credible intervals. *Information Syst. Manage.* 4: 53-63.
- Hunsaker, D. J., P. J. Pinter, and B. A. Kimball. 2005. Wheat basal crop coefficients determined by normalized difference vegetation index. *Irrig. Sci.* 24: 1-14. DOI: 10.1007/s00271-005-0001-0.
- Kamble, B., A. Kilic, and K. Hubbard. 2013. Estimating crop coefficients using remote sensing-based vegetation index. *Rem. Sens.* 5: 1588-1602. DOI: 10.3390/rs5041588.
- Kruschke, J. 2014. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*: Academic Press. EE. UU: Elsevier. DOI: 10.1016/C2012-0-00477-2.
- Lei, H., and D. Yang. 2014. Combining the crop coefficient of winter wheat and summer maize with a remotely sensed vegetation index for estimating evapotranspiration in the North China plain. *J. Hydrol. Eng.* 19: 243-251. DOI: 10.1061/(ASCE)HE.1943-5584.0000765.
- Neale, C. M., W. C. Bausch, and D. F. Heermann. 1989. Development of reflectance-based crop coefficients for corn. *Trans. Am. Soc. Agric. Eng.* 32: 1891-1899. DOI: 10.13031/2013.31240.
- Odi-Lara, M., I. Campos, C. M. Neale, S. Ortega-Farías, C. Poblete-Echeverría, C. Balbontín, and A. Calera. 2016. Estimating evapotranspiration of an apple orchard using a remote sensing-based soil water balance. *Remote Sens.* 8: 253. DOI: 10.3390/rs8030253.
- Palacios-Vélez, E., J. E. Palacios-Sánchez, and L. A. Palacios-Sánchez. 2011. *Agricultura de riego asistida con satélites. Tecnol. y Cienc. del Agua* 2: 69-81. <http://www.scielo.org.mx/pdf/tca/v2n2/v2n2a5.pdf>
- Pôças, I., A. Calera, I. Campos, and M. Cunha. 2020. Remote sensing for estimating and mapping single and basal crop coefficients: A review on spectral vegetation indices approaches. *Agric. Water Manage.* 233: 106081. DOI: 10.1016/j.agwat.2020.106081.
- Rozenstein, O., N. Haymann, G. Kaplan, and J. Tanny. 2018. Estimating cotton water consumption using a time series of Sentinel-2 imagery. *Agric. Water Manage.* 207: 44-52. DOI: 10.1016/j.agwat.2018.05.017.
- Salvatiery, J., T. Wieckiã, and C. Fonnesbeck. 2016. *PyMC3: Python probabilistic programming framework*. ASCL: ascl-1610.
- Sileshi, G. W. 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. *Forest Ecol. Manage.* 329: 237-254. DOI: 10.1016/j.foreco.2014.06.026.
- Shirley, R., E. Pope, M. Bartlett, S. Oliver, N. Quadrianto, P. Hurley, and J. Bacon. 2020. An empirical, Bayesian approach to modelling crop yield: Maize in USA. *Environment. Res. Comm.* 2(2): 025002. DOI: 10.1088/2515-7620/ab67f0.
- Toureiro, C., R. Serralheiro, S. Shahidian, and A. Sousa. 2017. Irrigation management with remote sensing: Evaluating irrigation requirement for maize under Mediterranean climate condition. *Agric. Water Manage.* 184: 211-220. DOI: 10.1016/j.agwat.2016.02.010.
- Wang, M., Q. Liu, L. Fu, G. Wang, and X. Zhang. 2019. Airborne LIDAR-Derived aboveground biomass estimates using a Hierarchical Bayesian approach. *Remote Sens.* 11: 1050. DOI: 10.3390/rs11091050.
- Weiskittel, A. R., D. W. MacFarlane, P. J. Radtke, D. L. Affleck, H. Temesgen, C. W. Woodall, and J. W. Coulston. 2015. A call to improve methods for estimating tree biomass for regional and national assessments. *J. Forestry.* 113: 414-424. DOI: 10.5849/jof.14-091.
- Zapata-Cuartas, M., C. A. Sierra, and L. Alleman. 2012. Probability distribution of allometric coefficients and Bayesian estimation of aboveground tree biomass. *Forest Ecol. Manage.* 277: 173-179. DOI: 10.1016/j.foreco.2012.04.030.
- Zhang, D., X. Yan, Z. L. Yang, A. Wall, and J. Wang. 2013. Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. *Reliab. Eng. Syst. Safe.* 118: 93-105. DOI: 10.1016/j.res.2013.04.006.

