Journal of Hydrology: Regional Studies 44 (2022) 101255

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Water sufficiency for cacao production in the Sierra Nevada de Santa Marta (SNSM) region, Colombia

Jefferson Valencia [a,*], Jane Frankenberger [b], Keith Cherkauer [b], Javier M. Martín-López [a], Fredy Monserrate [a,c], Mayesse da Silva [a]

[a] International Center for Tropical Agriculture (CIAT), Km 17 Recta Cali-Palmira. Z.C. 763537 - A.A, 6713 Cali, Colombia
[b] Agricultural and Biological Engineering Department, Purdue University, West Lafayette, IN, USA
[c] School of Agriculture and Environment, Massey University, New Zealand

ARTICLE INFO

Keywords:
Water storage
Agriculture
Water balance
Farmers

ABSTRACT

Study region: Sierra Nevada de Santa Marta (SNSM) region of Colombia.
Study focus: This research was conducted as a case study to generate relevant, quantitative information to support cacao farmer decision-making processes concerning water management in the SNSM. It involved the development and evaluation of a spatial dataset of precipitation and temperature, integration of digital soil mapping with a modification of the Thornthwaite and Mather water balance model, and finally an assessment of water sufficiency for cacao production. We elaborated site-specific and spatially-distributed analyses to generate information that will be shared with technicians who assist cacao growers in the SNSM.
New hydrological insights for the region: Under the climate conditions for the analysis period (1989–2018), rainfall was not enough to prevent cacao yield losses for 10 out of the 27 farms evaluated. The location of farms in two departments with contrasting climate conditions showed the importance of spatial analysis of water availability when providing recommendations of management practices to cacao growers. The results revealed that farms facing less frequent water stress are characterized by higher rainfalls and lower temperatures, soils that contain more organic matter, and are located at higher elevations with steeper slopes. Temporally, water stress is highest in the months February-August, with special interest in March-April as the dry season ends and July-August just before the peak rainy season.

1. Introduction

Cacao (*Theobroma cacao* L.) probably originated in South America being spread to Central America where it has been cropped for around 1500 years (Motamayor et al., 2002). While cacao production in Colombia has a long history, it was only recently introduced as a commercial crop in the Sierra Nevada de Santa Marta (SNSM) region, on the north coast of the country, as a substitute for illicit crops after the implementation of “Plan Colombia” in 2000, which is a strategy of the U.S. and Colombian governments for fighting drug trafficking (DNP, 2006). The Cacao for Peace (CfP) program, funded by the U.S. Agency for International Development (USAID) and executed by the Foreign Agricultural Service of the United States Department of Agriculture (USDA-FAS), has supported local cacao farmers as part of its goal “to improve the cacao value chain in Colombia through cooperative research, technical assistance, and

Abbreviations: SNSM, Sierra Nevada de Santa Marta.
* Corresponding author.

https://doi.org/10.1016/j.ejrh.2022.101255
Received 17 February 2022; Received in revised form 1 November 2022; Accepted 2 November 2022

extension education" (USDA, 2016).

Cacao is extremely susceptible to water availability and, therefore, drought is a significant stress factor for its growth, development, and production. Medrano et al. (2015) found that cacao is highly prone to water stress early in its development; however, water requirements at this stage have not been clearly defined. Oppong et al. (1999) found in a multi-year trial that seedlings planted during the first three months of the rainy season improved their survival rate after the dry season by 20 %, in comparison with those planted in the last month. Rainy season onset is also the most important environmental factor in triggering flowering, which is critical for production of cacao pods and beans (Adjaloo et al., 2012; Carr and Lockwood, 2011). Climate change and variability, especially changes in temperature and water availability in the dry season, also impact cacao significantly (Schroth et al., 2016). This is reflected in the yield reductions of up to 89 % due to the severe droughts caused by El Niño-Southern Oscillation – ENSO (Gateau-Rey et al., 2018).

Despite the relevance of water in cacao cropping systems, farmers in developing countries normally do not have enough information concerning cacao water use and water sufficiency to make decisions about potentially implementing management practices to reduce water shortages. The goal of this case study is to generate relevant, quantitative information to support cacao farmer decision-making processes concerning water management with a focus on the SNSM region of Colombia, but being applicable to other regions with similar geographical and data conditions. It assesses cacao water use and its critical soil water thresholds in farms, as well as mapping the spatial variability of soil water in the region. The specific objectives of this research are to: (1) develop and evaluate a spatial dataset of precipitation and temperature; (2) assess soils and relevant soil properties through digital soil mapping; (3) quantify the long-term monthly water balance across the region; and (4) assess water sufficiency for cacao production and specifically for the target farms.

Strategic information and results obtained through this research will be shared with farmers and local technicians who provide assistance in the study area. Understanding cacao water requirements and knowing when to provide supplemental water to meet its critical needs can help to increase cacao production in the region and lead to better, more reliable incomes for farmers. As this research is conducted with the support of the CfP project, it will contribute toward capacity building for cacao farmers, project partners, related ministries, and institutions to meet the nation's long-term goal to increase quality cacao production, land use sustainability, and natural resources conservation and long-term food security in Colombia for future generations.

2. Materials and methods

2.1. Study area

This research takes place along the northern coast of Colombia in an area that encompasses 1610 km^2 of La Guajira and Magdalena departments (Fig. 1). This area is limited in the south by the SNSM mountains, in the north by the Caribbean Sea, in the west by a national protective forest reserve, and in the east by the Tapias River. The region contains nine national natural parks and is surrounded by two more, among them the renowned Tayrona Park in the northwest. Two strong relief types can be differentiated: a

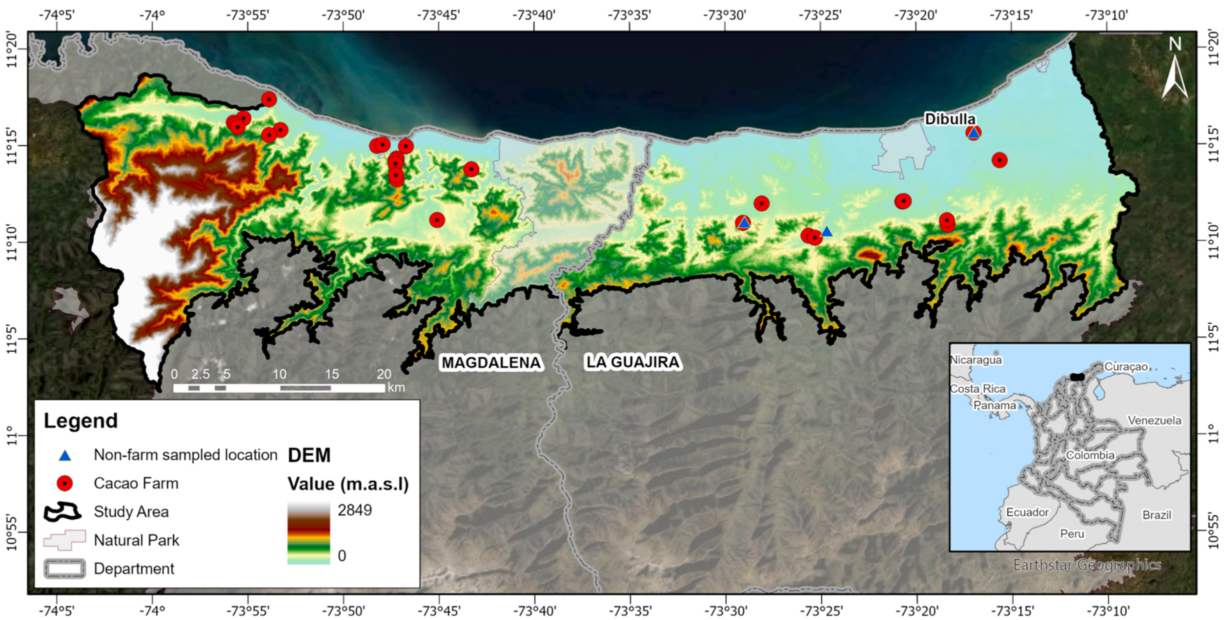

Fig. 1. Location of the study area within Colombia and split between the departments of Magdalena and La Guajira. Also marked are the study cacao farms and non-farm locations where soil characteristics were sampled. DEM: Digital Elevation Model. Refer to https://arcg.is/1HmGrL for better visualization of the cacao farms.

mountainous region located in the south and southwest with elevation ranging from 250 to 2849 m above sea level (m.a.s.l.), and a low region where wide plains predominate with an elevation range from 0 to 250 m.a.s.l. Average annual precipitation ranges spatially from 1300 to 4100 mm, with minimum temperatures from 4° to 21°C and maximum temperatures from 14° to 33 °C. This area experiences a bimodal rainfall behavior with a long dry period from December to April and a longer rainy season subdivided by a slight reduction in rains in July.

A water balance assessment was developed for 27 CfP assisted cacao farms located in the study area (see Fig. 1), 10 in the department of La Guajira and 17 in the department of Magdalena (see Table S1 in the Supplementary material for their general characteristics). La Guajira registers the lowest annual rainfall in the country, varying spatially from 300 to 1500 mm, while in Magdalena precipitation ranges from 1000 to more than 2500 mm (IDEAM, 2015). Study farms are located on both flat plains and mountainous relief with steep slopes. Two communal organizations of cacao producers are present in the region. Both groups focus on low-income producers that on average have one hectare of cacao in production. Most of these producers do not have irrigation systems and the majority of cacao systems are produced in combination with crops that provide shade to cacao including banana, citrus, and woody trees. Agronomic practices are implemented with expertise from the CfP, which has been helping cacao growers for many years to improve yields.

2.2. Climate data

Weather data for a 30-year climate period, from January 1, 1989, to December 31, 2018, were provided by the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM by its Spanish name). Stations within a buffer distance of 15 km of the study area were selected, resulting in 14 stations with daily precipitation data and seven with daily maximum and minimum temperature data that were subject to quality control and analyzed for data sufficiency.

2.2.1. Quality control (QC)

Data were cleaned to remove values out of range, statistical outliers, repetitive measurements during five days, and measurements that changed excessively between two consecutive days following the procedures of Esquivel et al. (2018). Minimum and maximum threshold values allowed for each variable were defined based on historic records (1981–2010) disaggregated by departments published on the Colombia Climatological Atlas (IDEAM, 2015). The precipitation range was defined as 0–200 mm/day, the minimum temperature range as 0–35 °C, and the maximum temperature range as 10–45 °C. Forty-five degrees Celsius is the historic maximum temperature reported for Colombia. Outliers, defined as values greater than a multiplier (20 for precipitation, 5 for temperature) of the interquartile range (IQR) were deleted. These multipliers were chosen to reflect the variability of the weather being higher for precipitation than for temperature. Repetitive or frequent values during five consecutive days, and values that changed by ten degrees Celsius or more between two consecutive days, were also removed in this process. Any value that failed any of the abovementioned checks was removed from the series and replaced with *NA*. Examples of cleaned datasets for two stations are provided in Fig. S1 (Supplementary material).

The resulting time series were analyzed in terms of missing values to determine which stations to use in the analysis. The mean percentage of missing values for minimum and maximum temperatures was almost 60 % while for precipitation it was 27 %. Due to the low density of stations in the study area, stations with percentages of missing data up to 43.3 % for precipitation and 67.7 % for temperature were kept. Two stations for temperature and two more for precipitation were discarded based on their missing value percentages for the analysis period (see Figs. S2-S4, Supplementary material). Table 1 presents the final selection of stations (10 for precipitation and 5 for temperature) that passed the quality checking process and were used in further analysis. One station (Parque. Tayrona) was discarded for precipitation but kept for temperature.

Table 1
Weather stations that passed the quality control process. Three stations in the region did not make this list. Elev.: elevation, Long.: longitude, Lat.: latitude, Prec.: daily precipitation, Min. Temp.: minimum daily temperature, Max. Temp.: maximum daily temperature, *: variable not measured at that station, and d: station discarded for that variable because of missing data.

Code	Station name	Elev. (m.a.s.l.)	Long.	Lat.	Missing values (%)		
					Prec.	Min. Temp.	Max. Temp.
15045010	Matitas	36	-73.03028	11.26389	15.1	37.7	52.2
15035020	Termoguajira	1	-73.41139	11.25250	43.3	60.0	57.9
15030020	Palomino	15	-73.57344	11.24472	42.8	*	*
15030010	Dibulla	10	-73.30639	11.27111	0.4	*	*
15015110	Alto.de.Mira	1025	-73.93239	11.09150	35.0	49.4	48.1
15015100	Parque.Tayrona	30	-73.91028	11.29167	*d*	53.3	67.7
15015060	San.Lorenzo	1749	-74.03361	11.08000	4.1	24.0	17.8
15010300	Guachaca	12	-73.83917	11.24750	1.1	*	*
15010040	Vista.Nieves	1535	-74.04222	11.07917	0.3	*	*
15010020	Buritaca	11	-73.76250	11.24944	0.0	*	*
15010010	Minca	614	-74.12000	11.14083	2.0	*	*

2.2.2. Data gap filling

Gaps in the daily weather data were filled using the nearest neighbor algorithm bounded by elevation (Moraes et al., 2022). This algorithm replaces missing values of a station with values from neighboring stations within an elevation restriction thus preventing the use of a closer station from a very different altitude. Two elevation restrictions, station elevations within 500 m and 1000 m, were used for temperature and precipitation, respectively. These values were determined based on elevation change with respect to climate within the study area, where larger changes in temperature than in precipitation were identified. In cases when a missing value could not be replaced from any of the nearest stations meeting the elevation restriction, a gap-filling iterative function was implemented to replace null values with mean daily values for the month calculated for the assessed station. The function checks other years for the same month until finding a non-null replacement value.

2.2.3. Generation of monthly climatic surfaces

After the data filling was completed, the gap-filled data was aggregated to a monthly time-step. This time scale was chosen as cacao is a perennial crop with a long-term life cycle of up to 40 years (Hernandez et al., 1989; Köhler et al., 2009), which implies that water demand is less temporally varying than it is with annual or transitory crops. Monthly surfaces were generated for each year through spatial interpolations at 1 km of resolution (approx. 30 arc-seconds) of monthly precipitation and minimum and maximum temperatures. This resolution was selected for consistency with widely used high spatial resolution climate datasets such as WorldClim (Hijmans et al., 2005). Navarro-Racines et al. (2020) also stated that this resolution allows for capturing climate patterns in climate data interpolations. For this purpose, the thin-plate spline smoothing method was used from ANUSPLIN version 4.4, which in turn, uses elevation as a covariate (Hutchinson and de Hoog, 1985). It was implemented a second-order spline interpolation and, consequently, created synthetic stations based on the gauging stations within the interpolation area (see Fig. 2), adding five stations for each variable. The resulting synthetic stations contain the same time series as the reference gauging stations and were created in locations with similar elevation and climate to theirs (Fig. 2). For this, a Digital Elevation Model (DEM) and the multi-year annual precipitation and mean temperature isohyets of Colombia for the period 1981–2010 (IDEAM, 2017a,b) were used. This process was carried out as the spline order required at least 10 stations per variable for its correct functioning. In total, 15 stations were used for precipitation and 10

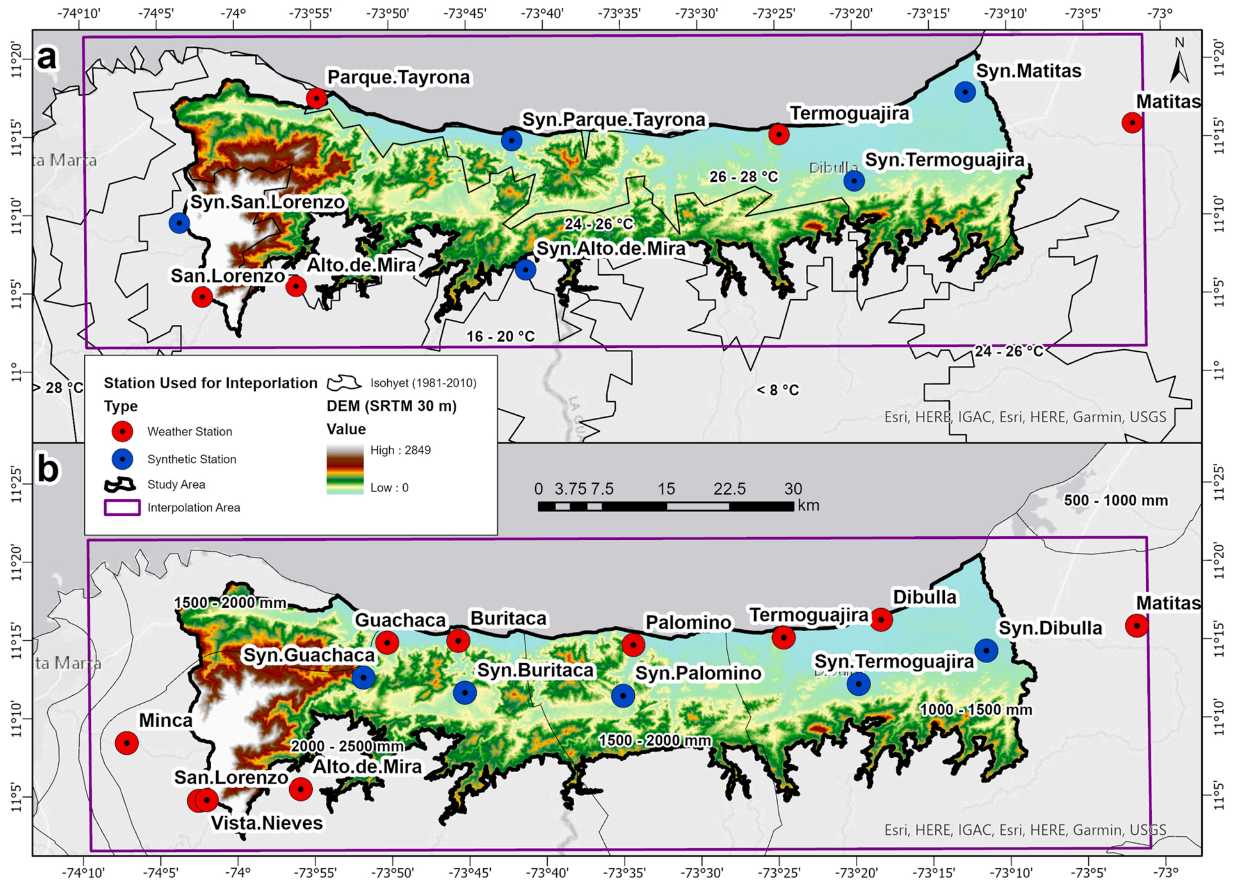

Fig. 2. Temperature (a) and precipitation (b) stations used for interpolation. Blue circles represent synthetic stations created based on gauging stations (red circles) and multi-year annual precipitation and mean temperature isohyets of Colombia for the period 1981–2010 (black contour lines).

for temperature. To carry out the interpolation process, the DEM, originally produced at 30 m, was resampled at a final spatial resolution of 1 km by using the bilinear resampling method.

To evaluate the performance of the interpolation algorithm, k-fold cross-validation (Fushiki, 2011) was executed using 75 % of data for training, 25 % for testing, and defining k = 25 folds (number of iterations). The resulting errors from interpolations were assessed by using the coefficient of determination (R^2) and the root mean square error (RMSE). Cross-validation results are presented in the Supplementary material (section S1.2). Based on the calculated R^2 values, the iterations whose medians were greater than the 90th percentile of all medians were identified and considered as the "best iterations". Their corresponding surfaces were then averaged by month, resulting in the final surfaces of the interpolation process. Values of the time-series climate characterization were restricted to the value ranges (thresholds) defined for the quality control process. Final surfaces were then clipped with the limit of the study area and projected to the local coordinate reference system "MAGNA-SIRGAS Colombia Bogota", which was used for all spatial information considered in this study. Finally, those surfaces were also aggregated and plotted on a multi-year monthly basis.

2.3. Digital soil mapping (DSM)

2.3.1. Soil data

Soils were sampled at twenty-seven farms and three non-farm locations (Fig. 1), resulting in seventy-three auger sites and a set of ten full pedon descriptions, seven of them at the farms. Sampling sites were selected based on representative soil units and landscapes of the study area, being the non-farm locations chosen for full pedon descriptions to cover the remaining soil units (Libohova et al., 2020). In total, 415 samples were taken through soil horizons down to depths of up to 2 m or until reaching the unconsolidated material in which the soil is formed. Effective soil depth (ESD) refers to physical restrictions such as high stoniness, compacted soil horizons, or unweathered parent materials (Table S1 in the Supplementary material shows the ESD at each sampling location within a farm). Soil properties such as texture (sand, clay, and silt) and organic matter (OM) were determined in the laboratory for every sample using the methods of Bouyoucos (1936) and Walkley and Black (1934), respectively. Available water content (AWC) was calculated by the difference between field capacity (33 kPa) and wilting point (1500 kPa) for all locations based on the pedotransfer functions (PTF) that in turn use sand, clay, and OM as input for characterizing hydrological soil properties (Saxton and Rawls, 2006).

An equal-area-spline function using the "GSIF" R-package (Hengl, 2020) was applied to transform measured data into specific depth layers (Odgers et al., 2012). Four layers were defined for this research based on cacao root interactions with soils and the maximum depth reached in the field sampling: 0–20 cm, 20–60 cm, 60–100 cm, and 100–200 cm. The first defined layer (0–20 cm depth; H1) includes the most active, fine cacao root system, which is essential in water and nutrient absorption (Arévalo-Gardini et al., 2017; Argüello et al., 2019). The second layer (20–60 cm depth; H2) contains up to 80 % of roots in cacao plantations (Asmamaw et al., 2017; de Almeida and Valle, 2007). Finally, since cacao roots can reach depths up to 1.5–2 m (Carr and Lockwood, 2011), two subsequent layers (60–100 cm; H3, 100–200 cm; H4) were also considered. At these last two depths, the cacao root system is dominated by the taproot, which is mainly in charge of anchoring the plant and from which other secondary roots sprout in the search for water and nutrients (Rajab et al., 2018).

2.3.2. Soil properties prediction

Soil texture, ESD, and AWC are important inputs for the water balance model. Thus, spatial distributions of the ESD (cm), soil texture (sand, silt, and clay in %), and AWC (%) for all soil depths (H1-H4) were mapped at 30 m of spatial resolution using a digital soil mapping (DSM) approach based on random forest. DSM predicts the spatial distribution of soil properties based on soil-forming factors, described by (Pendleton and Jenny, 1945), by combining field and laboratory data with a variety of environmental variables in a mathematical model (Ma et al., 2019; McBratney et al., 2003; Minasny and McBratney, 2016; Moore et al., 1993; Zhu, 1997; Zhu et al., 2001). Random forest is a machine learning method, widely used in DSM, that predicts soil properties from a large collection of non-correlated decision trees (Breiman, 2001), which are created from the spatial relationship between soil data and environmental variables. The following environmental variables at 30 m of spatial resolution were implemented as predictors in the DSM modeling: DEM (digital elevation model), channel distance, geomorphons (Jasiewicz and Stepinski, 2013), annual median NDVI (Normalized Difference Vegetation Index, obtained from Sentinel-2 spectral images for the years 2017, 2018, and 2019 using Google Earth Engine), land use, geology, slope, TWI (Topographic Wetness Index), relative slope position (RSP), soil classes, plan and profile curvatures, general ecosystems, MRRTF (multiresolution ridge top flatness), MRVBF (multiresolution valley bottom flatness), LS factor, valley depth, and TPI (topographic position index). More details on these environmental variables including their sources can be found in Table S2 (Supplementary material).

The data was randomly split into two sets: 70 % for model training, and 30 % for validation. The random forest was implemented through the "randomForest" R-package (Liaw and Wiener, 2003), using the following parameters: training set (composed of the estimated data and the extracted values of predictors at every sampling location), *ntree* = 1000, *mtry* = 6, and *nodesize* = *5*.

Model performance was assessed through the calculation of three statistical indices applied to the validation set: root mean square error (RMSE), mean absolute error (MAE), and nRMSE; which is the RMSE normalized by the data range (max-min). Results of model performance assessment are in the Supplementary material (section S3).

2.4. Land use/land cover (LULC)

The map used to represent the land use/land cover (LULC) was generated by IDEAM with the *CORINE Land Cover (CLC)* classification system (CEC, 1993) adapted for Colombia at a scale of 1:100,000 for 2017. The study area is composed of 30 LULC classes where cacao farms represent 0.02% (38.4 ha) of the total extent. These 27 cacao crop farms were inserted/imposed on the LULC map as a class named "Cacao". The classes in both Spanish (original) and English, other referenced parameters (i.e., Kc, p, and CN; which will be explained later in Section 2.5) and their sources, as well as the corresponding areas, are shown in Table S3 (Supplementary material).

2.5. Water balance model

The water balance was developed for the entire region by calculating surface runoff and effective precipitation, evapotranspiration, and percolation as described below.

2.5.1. Surface runoff calculation

Surface runoff was estimated using the Soil Conservation Service Curve Number (CN) method (USDA-SCS, 1986). For this, the hydrologic soil group (HSG) for each soil type was determined based on the texture of the first layer (H1) using the USDA-SCS (1986) procedure and the software SAGA GIS. Soils classified as HSG A have the smallest runoff potential while soils with HSG D have the greatest. The CN2 (average condition) values for all the classes of the LULC map were assigned based on multiple references (see Table S3, Supplementary material). The raw CN2 raster was calculated by executing the tool "Pick" of ArcGIS for Desktop using the calculated HSGs and the CN2s associated with the LULC classes that were rasterized at the same resolution of the DEM (30 m). This tool picks a value from a specific raster based on a value of another one (i.e., position raster, in this case, HSG), so it resulted in the association of the corresponding CN2 value according to the HSG. CNs are theoretically estimated for flat conditions (Neitsch et al., 2011) or plot slopes < 5 %, so CN2 for slopes greater than 5% were adjusted according to the procedure of Huang et al. (2006). CN2 values were not adjusted for antecedent moisture conditions, based on current recommendations in Hawkins et al. (2019).

To calculate runoff, daily precipitation values were first allocated to Thiessen polygons created for the stations used in the generation of the climatic surfaces (Fig. 2). By using these polygons and with the slope-adjusted CN2 raster, the runoff was calculated for each day of the series. Finally, the resulting daily runoff surfaces were aggregated on a monthly basis which served as an input to the water balance calculation.

2.5.2. Potential evapotranspiration

There are multiple methods used worldwide to calculate potential evapotranspiration based on weather variables such as

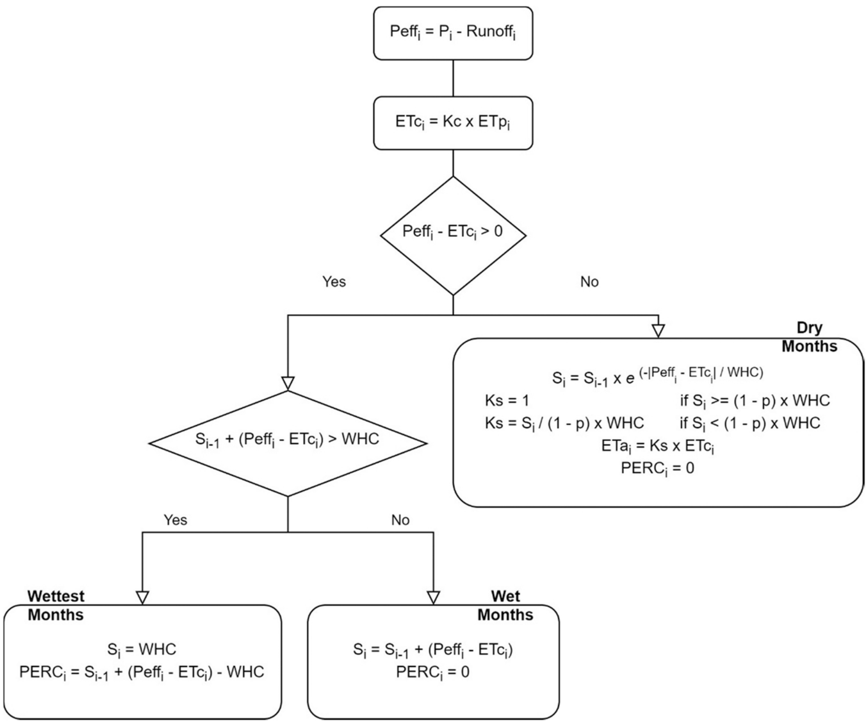

Fig. 3. Thornthwaite and Mather water balance scheme adjusted and used for monthly hydrological assessment. For month i, P_i is the precipitation [mm], $Runoff_i$ is the surface runoff [mm], $Peff_i$ is the effective precipitation [mm], WHC is the water holding capacity [mm], S_i is the soil water storage [mm], ETp_i is the potential evapotranspiration [mm], Kc is the crop evapotranspiration coefficient [dimensionless], ETc_i is the crop evapotranspiration [mm], p is the soil water depletion fraction for no stress [0 − 1], Ks is the water stress coefficient [dimensionless], ETa_i is the actual evapotranspiration [mm], and $PERC_i$ is the soil percolation [mm].

temperature, sunshine, wind speed, solar radiation, and humidity. The Hargreaves method (Hargreaves et al., 1985) has been implemented in areas with limited weather information demonstrating reasonable results (Allen et al., 1998), which makes it useful for the estimation of water balances in such locations. Droogers and Allen (2002) modified the standard equation of Hargreaves by including monthly rainfall as one of its components and concluded that, with this modification, it is more appropriate for use in areas with limited data availability of weather information than methods with higher data requirements like Penman-Monteith. For its implementation, solar radiation was calculated with equations 21–25 of the Food and Agriculture Organization – FAO-56 crop evapotranspiration guide (Allen et al., 1998).

2.5.3. Actual evapotranspiration

The crop evapotranspiration (ETc) of land uses/covers in the study area was determined based on potential evapotranspiration (ETp) estimates and a crop factor (Kc). Kc values were assigned to all the classes of the LULC map based on literature values (see Table S3, Supplementary material) and the resulting layer was rasterized. The Kc value assigned to the cacao farms (1.05) is for a mixed cacao system with total canopy cover (Allen et al., 1998). In consequence, the ETc of the farms should be close to that of a system with tall trees (Carr and Lockwood, 2011).

In dry months, ETc was reduced based on the calculated water stress to determine the actual evapotranspiration (ETa). The depletion fraction (p) of the water holding capacity represents the moisture that can be depleted from the root zone before moisture stress (reduction in ETc) occurs (Allen et al., 1998; Siebert and Döll, 2010). According to Allen et al. (1998), a value of 0.3 for p can be assigned to a cacao crop, which defines it as sensitive to water stress as stated by Raes (2002). For other (non-cacao) crops, a value of 0.5 was used (Allen et al., 1998) while for non-vegetated LULC classes a value of 0 was assigned (Table S3). A water stress coefficient (Ks) was calculated based on soil moisture, and ETc was reduced when the soil water storage (S) was below the threshold defined by p. No reduction was applied to ETc (Ks = 1) when S was at or above that threshold level.

2.5.4. Water balance calculation

The water balance was based on the modification by Ulmen (2000) of the widely-used Thornthwaite and Mather water balance model (Thornthwaite and Mather, 1955) following the scheme shown in Fig. 3. The starting value of S was assumed as 90% of the water holding capacity (WHC) following suggestions of Ulmen (2000). WHC was determined based on the available water content (AWC) and the effective soil depth (ESD) resulting from the digital soil mapping. Effective precipitation (Peff) is the remaining precipitation (P) after being reduced by surface runoff. Months were defined as dry (ETc > Peff) or wet (ETc ≤ Peff).

This model assumes that initial conditions are wet (water content is close to WHC), so the water balance was set to start in November, a month after the wettest month in the study area (October). Also, the first year (1989) was used as the initial/warm-up year of the water balance.

2.6. Water sufficiency for cacao production

2.6.1. Previous studies of effects of water deficit on cacao yields

Zuidema et al. (2005) found in a modeling study of 30 locations in 10 countries that rainfall of the two driest months and annual solar radiation could explain 70 % of cacao yield variations and that yield reductions up to 50 % could occur when rain is below 50 mm for two months. In an experiment implemented in Indonesia, Schwendenmann et al. (2010) found similar yield reductions (~45 %) when a severe drought period occurred during the initial stage of fruiting. They also noticed that slight drought levels during that stage decreased yield by 10 %. While cacao yield is severely reduced by drought, cacao trees have been shown to maintain minimal physiological activities to survive, and to suppress totally or partially their reproductive stage (Adjaloo et al., 2012; de Almeida et al., 2016; Moser et al., 2010). Köhler et al. (2010), Moser et al. (2010), and Schwendenmann et al. (2010) studied cacao water stress through physiological responses, cacao productivity, and their related soil moisture levels. They found that there are significant soil water thresholds that affect cacao yield and, therefore, can be considered to determine the water needs of a cropping system that would support increased productivity.

2.6.2. Soil water thresholds

Understanding and determining critical soil water thresholds of the cacao crop is of high relevance not only for the study area but also for other places where cacao cropping systems exist. The definition of these thresholds can provide a basis for understanding when a cacao system is likely facing water stress. Accordingly, these thresholds were determined based on the calculation of the Relative Extractable Soil Water – REW (Granier et al., 1999; Schwendenmann et al., 2010):

$$REW_i = \frac{S_i - S_{min}}{S_{max} - S_{min}}$$

where S_i is the soil water storage at month i [mm], and S_{min} and S_{max} are the long-term minimum and maximum values of S [mm]. Water stress can occur in cacao crops when the REW value is below 0.4, which can entail a yield reduction of up to 10% (Granier et al., 1999; Schwendenmann et al., 2010). REW values of 0.1 or less imply that cacao trees do not continue producing biomass, that is, they stop their vegetative growth and could be at the threshold of withering, generating yield reductions of up to 45% in the system (Granier et al., 1999; Moser et al., 2010; Schwendenmann et al., 2010).

Thus, the lower and upper soil water thresholds for cacao were defined as REW = 0.1 and REW = 0.4, respectively. The values of soil water storage (S) for those REW values were determined for each location in the study area and the difference for each month between the S_i and S at each threshold was calculated to determine if the location was experiencing water stress. Negative differences represent the supplemental soil water content in mm that would need to be added through the implementation of management practices such as irrigation to maintain a healthy cacao crop.

3. Results

3.1. Climate dataset

The results of the climate data processing and interpolation were (1) the cleaned and gap-filled time series for each station and variable and (2) the interpolated maps of precipitation, and minimum and maximum temperatures for each month. These results are available online at Valencia et al. (2022). Climate varies spatially, especially from the lowlands of the northeast to the high elevations of the southwest. As an overview of the resulting climatic surfaces, Fig. 4 shows the average annual precipitation and temperature extremes with the corresponding stations used for their generation. Rain is higher (~4100 mm/year) in the southwest region where the highest elevations predominate and where the lowest temperatures occur (~5 °C). The highest temperatures (~33 °C) and the lowest rains (~1300 mm/year) are in the lower altitudes, especially in the northeast (La Guajira department).

Temporal climate patterns in Fig. 5 illustrate that rainfall presents a bimodal behavior with predominant peaks in May and October. The dry season from December to April is critical for water availability for cacao. Temperature is quite consistent throughout the year, with slightly warmer temperatures from June to August. In February, the month with the lowest precipitation, average rainfall ranges across the region from 8 to 42 mm/month, minimum temperature from 4 to 21 °C, and maximum temperature from 14 to 33 °C. In October, the month with the highest precipitation, average rainfall varies spatially between 290 and 471 mm/month, minimum temperature between 6 and 23 °C, and maximum temperature between 13 and 33 °C.

3.2. Digital soil mapping (DSM)

Descriptive statistics of regional soil properties such as effective soil depth (ESD), soil texture (sand, silt, and clay), organic matter, and estimations of field capacity, wilting point, and available water content (AWC) are summarized in Table 2. Loamy soil textures predominate in all layers (H1-H4), followed by clay loam, sandy loam, and sandy clay loam. Mean values of sand, silt, and clay remain similar through all depths, while the mean organic matter value is higher for shallower soils, due to the higher accumulation of this property in the first centimeters of the soil than deeper in the soil profile. This is attributed to more biological activity and a greater presence of litter, vegetation residuals, and root systems in the topsoil layer. Together these properties result in a slightly higher capacity of the soil at H1 to hold water at tensions of 1500 kPa (WP) and 33 kPa (FC), and a greater mean value of AWC at this depth.

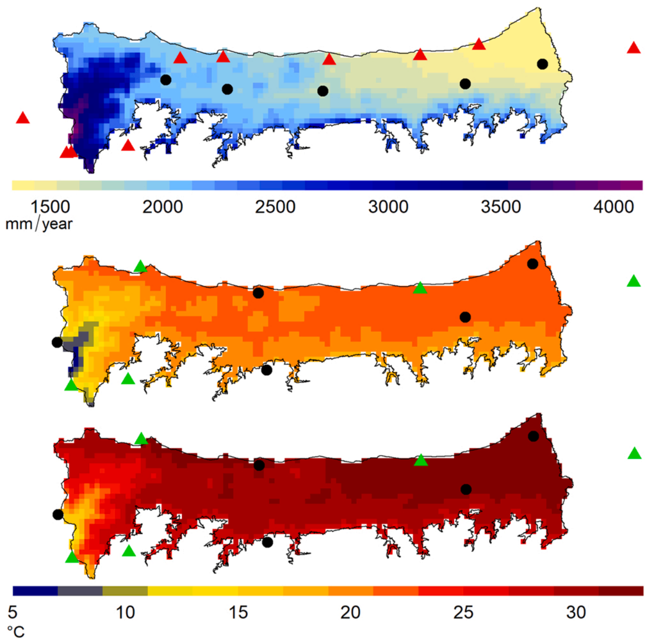

Fig. 4. Climatic surfaces masked to the study area and generated through interpolation of weather station time-series. Long-term (1989–2018) annual precipitation (top) and minimum (middle) and maximum (bottom) air temperatures with corresponding stations. Triangle marks (red for precipitation and green for temperature) define the locations of gauging stations while black dots mark the locations of synthetic stations used to improve the interpolation results. Values in mm/year for precipitation and degrees Celsius for temperature.

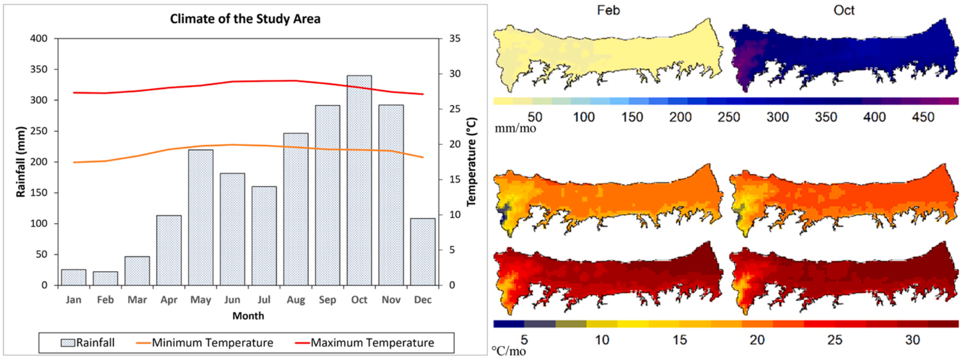

Fig. 5. Temporal climate pattern of the study area (left, mean values over the entire study area) and spatial distribution of average monthly precipitation (top-right), and average monthly minimum (middle-right) and maximum (bottom-right) temperatures for February (Feb) and October (Oct), the months with the lowest and highest precipitation, respectively. Values in mm/month for precipitation and degrees Celsius for temperature.

Table 2
Descriptive statistics of soil properties mapped using DSM at each soil depth: sand, silt, clay, organic matter (OM), soil texture, field capacity (FC), wilting point (WP), available water content (AWC), and effective soil depth (ESD).

Soil property	H1: 0–20 cm	H2: 20–60 cm	H3: 60–100 cm	H4: 100–200 cm
Sand (%)	46 ± 13	45 ± 14	46 ± 16	47 ± 18
Silt (%)	34 ± 12	34 ± 12	32 ± 13	34 ± 14
Clay (%)	21 ± 7	21 ± 8	21 ± 10	19 ± 8
Soil texture (proportion)	loam: 73 %, clay loam: 15 %, others: 12 %	loam: 89 %, sandy clay loam: 5 %, others: 6 %	loam: 58 %, clay loam: 32 %, others: 10 %	loam: 66 %, sandy loam: 17 %, others: 17 %
OM (%)	3.2 ± 1.4	1.29 ± 0.6	0.7 ± 0.4	0.6 ± 0.4
WP (% vol)	15.0 ± 4.2	13.4 ± 5.0	12.7 ± 5.1	13.0 ± 6.6
FC (% vol)	28.3 ± 5.3	25.9 ± 6.3	24.7 ± 6.7	24.3 ± 8.7
AWC (% vol)	13.3 ± 2.6	12.4 ± 2.5	12.0 ± 2.8	11.9 ± 3.2
ESD (cm)	99.2 ± 33.5			

Regarding ESD, a predominance of moderate (50–100 cm) and deep (100–150 cm) soils was identified at sampled locations, while shallow (< 50 cm) and very deep (> 150 cm) soils were not common (see Table S1 in the Supplementary material).

The spatial distribution of AWC predicted for each soil depth used in this study (H1 to H4) is presented in Fig. 6. Higher values of AWC at depths H1 and H4 were predicted in the northern plains, likely due to the presence of soils derived from sedimentary rocks, which are characterized by high organic matter contents and finer soil textures. Finer soil textures promote soil aggregation, OM protection, and soil moisture (Augustin and Cihacek, 2016; Xie et al., 2021; Yang et al., 2016), enhancing thus the water holding capacity and plant water availability of the soil. Higher values of AWC at depths H2 and H3 were predicted in the soils formed mainly from metamorphic and igneous rocks, located in mountainous landscapes characterized by higher OM and fine textures contents. Spatial distributions of soil texture at depth H1 and ESD as well as soil sampling data and other soil property maps can be visualized at

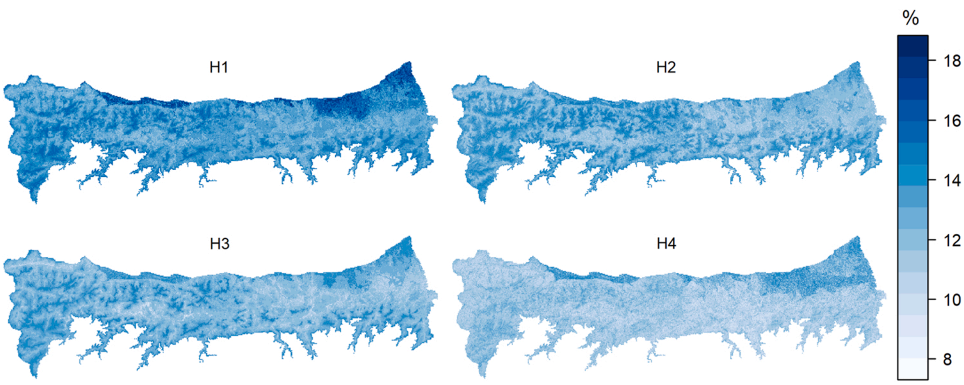

Fig. 6. Soil available water content (AWC) as a percentage, predicted for the study area at four different depths: H1 (0–20 cm), H2 (20–60 cm), H3 (60–100 cm), and H4 (100–200 cm).

the Cacao for Peace GIS Platform - https://arcg.is/1HmGrL.

3.3. Water balance assessment

The long-term (30 years) annual water balance developed for the study area determined that the average annual rainfall of 2041 mm is distributed as 234 mm of surface runoff (11% of rainfall), 831 mm of evapotranspiration (41% of rainfall), and 987 mm of percolation (48 % of rainfall). As no streamflow measurements were available for the study area for validation of the water balance results, the GSCD – Global Streamflow Characteristics Dataset (Beck et al., 2015, 2013) at a spatial resolution of 0.125 degree (approx. 13.5 km at the equator) was used as an approximate quantitative assessment, that is appropriate for ungauged catchments, to evaluate the simulated values. The water yield (runoff + percolation) ratio of the developed water balance (59.8 %) compared well with that provided by GSCD (60.3%) for the study area. This small difference in values allowed us to conclude that, in general terms, the model used in this research correctly balances water long-term for the study area and can be used.

Our evapotranspiration estimates also compared well to values from the literature for cacao. Computed ETa of the 27 cacao farms from our water balance model ranged between 776 and 1130 mm/year with a mean value of 968 mm/year. These results were found to be similar to those reported by Radersma and de Ridder (1996), who provided a range of annual ETa for cacao between 878 and 1074 mm, where the lower number was for cacao under drought stress.The spatially-averaged monthly water balance for the region is presented in Fig. 7 and complete monthly results are in Table S4 (Supplementary material). During dry months (January to March), ETa was greater than rainfall, meaning that evaporative demand of plants is drawn from soil moisture and that there is an increased potential for water stress at the end of this dry season.

3.4. Water sufficiency for cacao

The method developed allowed the evaluation of the spatial distribution of soil water availability and its relationship to the critical thresholds for cacao on a monthly basis (Fig. 8). The maps identify locations in the region where cacao is more or less susceptible to water stress based principally on climate and soil. Analysis of the resulting climate datasets shows that those farms (10 out of 27) located in the department of La Guajira are more vulnerable to water shortage and, therefore, their corresponding cacao systems are likely to experience water stress more frequently.

Most of the cacao farms would require supplemental water to not fall below the upper critical threshold where cacao drought stress can occur (see top-left map in Fig. 8), The supplemental water demand is larger for those farms located on the east side of the study area, with an average monthly demand of up to 50 mm in February. For this month, multi-year average monthly soil water storage for none of the 27 cacao farms fell below the lower critical threshold (10 %).

For the month with the highest precipitation, October, the average monthly soil water storage exceeds both critical thresholds for the entire study area, which indicates that the cacao farms do not need supplemental water during this month and do not experience drought stress. However, this analysis is based on long-term average surfaces for the study area and, therefore, extreme values or critical moments during the 30 years of water balance modeling could be smoothed or hidden. For this reason, specific analyses by farm were performed considering the whole time-series which allowed us to make comparisons among farms and identify the critical times of the year for cacao in the study area.

The percentage of all months over the 30-year simulation period that the soil water storage was below both critical thresholds is shown in Fig. 9 for each of the 27 farms, ranked by the percentage below the lower (10 %) critical threshold. When soil water storage is below the upper critical threshold, the cacao crop starts experiencing drought stress, and when it is below the lower critical threshold it no longer produces biomass, stops all vegetative growth, and is at the limit of withering and permanent damage. Farms located on the

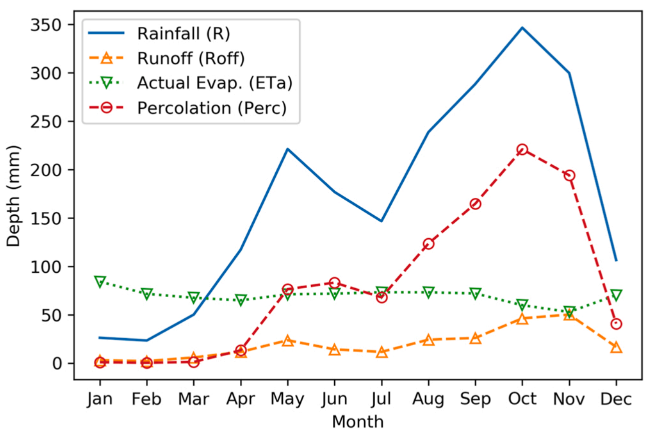

Fig. 7. Long-term (1989–2018) monthly water balance components (rainfall, surface runoff, actual evapotranspiration, and percolation) presented as mean values for the entire study area.

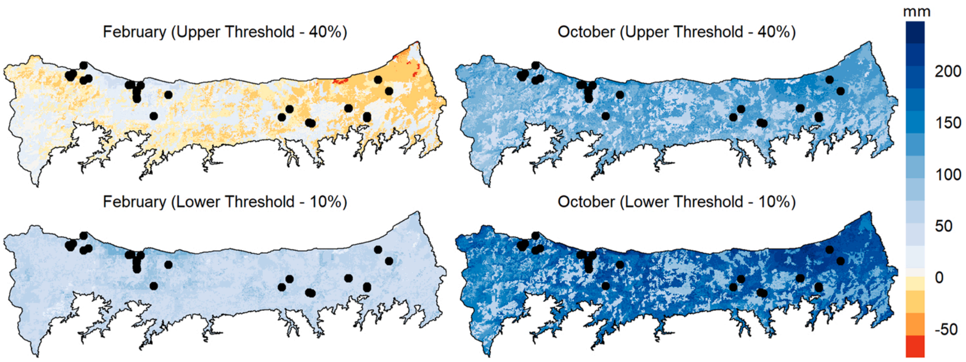

Fig. 8. Soil water storage (in mm) in excess of the upper (40% - top) and lower (10% - bottom) critical thresholds for the months with the lowest (Feb) and highest (Oct) precipitation. Negative values (oranges) indicate the amount of supplemental water that would need to be added to the soil to reach the specific threshold; while positive values (blues) identify areas that exceed the specific threshold under natural rainfall conditions. Black dots indicate the locations of study cacao farms.

left side of the chart, mostly located in the department of Magdalena, experience less water stress than those on the right, which are mostly in the department of La Guajira. Both critical thresholds should be analyzed together to have a realistic outlook of the potential for water stress at a farm, as the crop is suffering from drought impacts while below the upper critical threshold but in danger of permanent damage when below the lower critical threshold. For example, farms F56 and F9, both in the La Guajira department, experience nearly the same percentage of months below the upper critical threshold (33 % and 34 %, respectively), but F9 is below the lower critical threshold for 5 % of all months while F56 is never below the lower critical threshold. This indicates that both farms suffer from drought a similar amount of time, but that drought conditions are more severe at F9.

Correlation between the percentage of months that soil water storage was below either of the critical soil water thresholds for the 30 years of analysis and variables describing location, soil, and climate of individual farms were evaluated to identify potential predictors of farms that suffer more from water scarcity (see Table S1 and Fig. S6 in the Supplementary material for the corresponding scatter plots). Variables such as annual precipitation, longitude, organic matter, and mean temperature are statistically significant and highly correlated to the threshold percentages (Table 3), while elevation and slope were found to be less strongly correlated. Other variables such as latitude, effective soil depth, and soil texture (sand, clay, and silt) were not statistically significant. In summary, farms facing less frequent water stress conditions for cacao production are characterized by higher rainfalls and lower temperatures, are located more to the west at higher elevations, present greater slopes, and their soils contain higher organic matter, which facilitates soil water retention and, consequently, increases soil moisture. Although these farms are in a steep relief, their dense coverage of associated shading trees and litter are efficient in controlling runoff.

The approaches and methods used in this research allowed for the discrimination of farms in terms of water sufficiency even when they are in close spatially proximity. For example, three farms (F40, F39, and F38) were very close to each other spatially (approx. 250 m apart), but were found to have very different water balance characteristics. F40 experiences considerably less water stress than farms F38 and F39 (see Fig. 9, and Table S1 in the supplementary material). Farm F40 is at a higher elevation (358 m.a.s.l.), with less soil clay content (22 %) and a higher infiltration capacity (HSG: B and CN2: 69), in contrast to F38 and F39 (279 and 315 m.a.s.l., respectively, clay: 26 %, HSG: C, and CN2: 79). Furthermore, as revealed through the soil sampling campaign, F40 presented a denser coverage of associated shading trees.

3.5. Individual farm assessments

The tool developed as part of this research can produce summaries of water availability and stress level for each of the farms. To illustrate this point, two farms have been selected for additional discussion. Farm F0, located in the La Guajira department, is an example of a farm showing significant water stress as its soil water storage was below the upper critical threshold 65 % of the time, and below the lower critical threshold 20 % of all months (see Fig. 9). Meanwhile, farm F29, located in the Magdalena department, experienced almost no water stress with soil water storage below the upper critical threshold only 2 % of the time, and never below the lower critical threshold. As examples of the specific analyses that were carried out for each farm, the water balance in terms of soil water storage for both average conditions and extreme climatic years together with the soil water thresholds relevant to cacao are presented in Fig. 10 for these two contrasting farms.

Soil water storage at Farm 0 peaks in October and November along with annual average rainfall, but by December soil storage has already begun to drop with the end of the rainy season. Storage is below the upper critical threshold for February, March, and April, with soil moisture storage reaching a minimum in April even as the new rainy season gets started. Farm soil moisture storage does not recover and exceeds the upper critical threshold until August and September when the second, wetter part, of the rainy season begins. This farm is highly dependent on the timing of the start of the wet season and on the quantity of rainfall delivered. It is a location that would benefit significantly from irrigation to reduce the impacts of drought.

Farm 29, on the other hand, hardly experiences any water stress in most years. Storage falls below the upper critical thresholds only

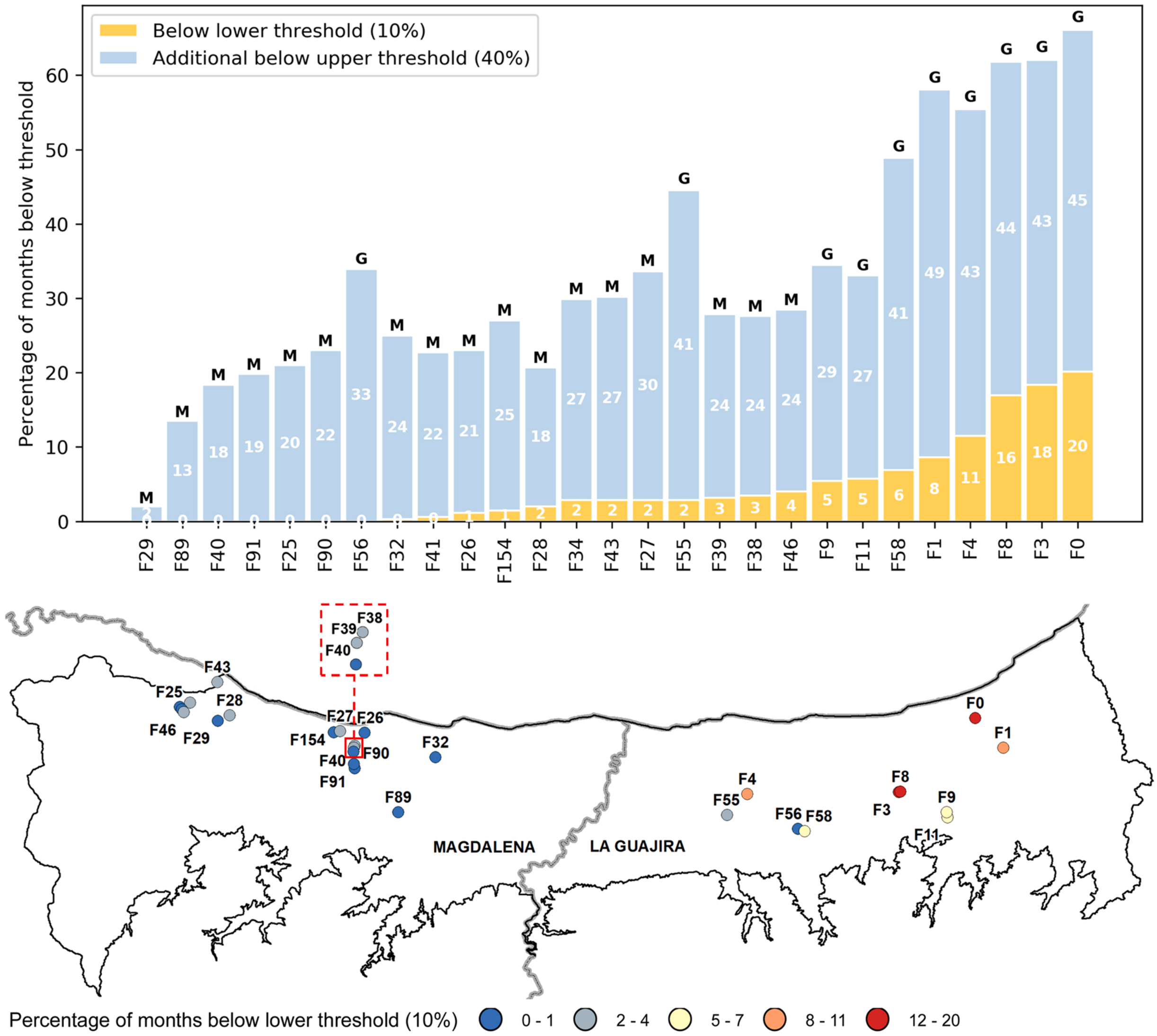

Fig. 9. Percentage of months during the analysis period (1989–2018) where soil water storage was below the upper (40 %) and lower (10 %) critical thresholds by farm. Number of months falling below the lower threshold is indicated on the yellow bar of the chart while the total height of a bar (blue + yellow) represents the total percentage of months below the upper threshold. Farms are ranked by the percentage below the lower threshold on both the chart and the map. G: La Guajira department, M: Magdalena department.

17 % of the time in March, and only 7% of the time in April. This farm has never come close to falling below the lower threshold. The need for irrigation at this farm is small. Future farms placed in similar locations would likely not need supplemental irrigation.

4. Discussion

The new climate dataset described here, which includes spatially varying precipitation and temperature, provides an improved basis for analyzing water sufficiency in the study region. The 1-km spatial resolution used for the interpolation of climate surfaces allowed us to capture the weather variations of the study region that could be lost at lower resolutions, particularly in mountainous landscapes. Existing climate datasets such as the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS, Funk et al., 2014) or the Global Precipitation Measurement (GPM, Huffman et al., 2019) are either at coarser resolution or do not focus on this region.

The newly-developed region-specific digital soil maps also provide advantages over currently available global digital soil maps (e.g., SoilGrids, Hengl et al., 2017) since they were developed with more consistent samples at a greater scale of detail using a sampling scheme developed to characterize the local soil characteristics for model training.

The use of these new datasets, together with the modified Thornthwaite and Mather model to calculate a spatially-varying monthly water balance, provides an improved method for determining water sufficiency for cacao production in the region. Existing studies

Table 3
Correlation between the percentage of months below critical soil water thresholds and farm-related variables. Statistically significant (p-value ≤ 0.05) variables are marked with asterisk.

Variable	Critical threshold	
	Lower (10 %)	Upper (40 %)
Precipitation (mm/year)	-0.79*	-0.88*
Organic Matter (%)	-0.61*	-0.60*
Elevation (m.a.s.l.)	-0.41*	-0.44*
Slope (%)	-0.41*	-0.42*
Clay (%)	-0.35	-0.32
Latitude	-0.20	-0.33
Curve Number (average condition)	0.03	0.0.4
Sand (%)	0.05	0.03
Silt (%)	0.09	0.11
Effective Soil Depth (cm)	0.18	0.19
Mean Temperature (°C)	0.57*	0.63*
Longitude	0.70*	0.79*

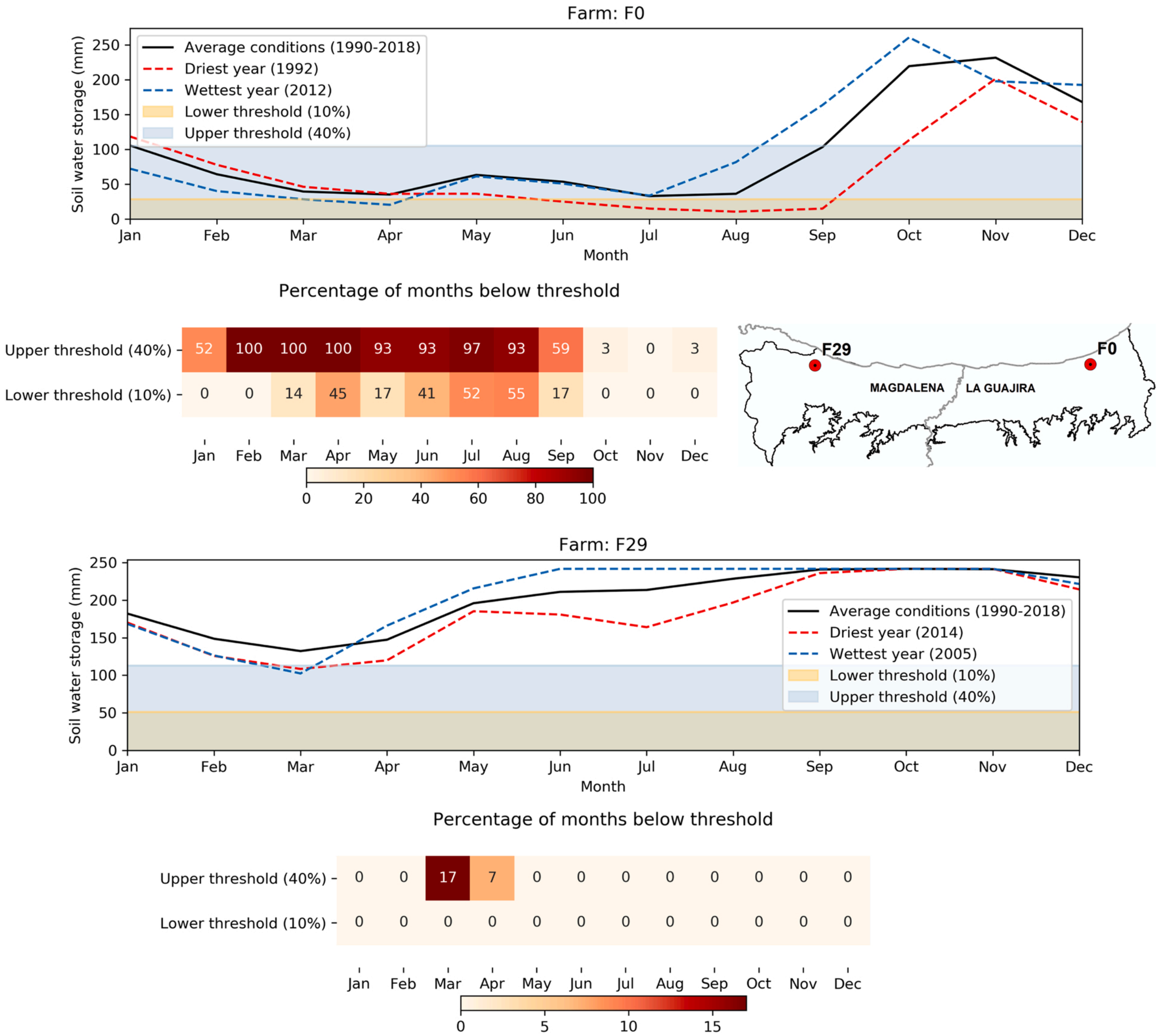

Fig. 10. Soil water storage and thresholds for farms F0 (with significant water stress - top) and F29 (experiencing almost no water stress - bottom).

either focused on general climate variables such as precipitation and temperature without explicitly calculating a water balance (e.g., Singh et al., 2021) or calculated the water balance at the plot scale without a broader spatial analysis (Köhler et al., 2014, 2010).

Comparison to thresholds for cacao yield reductions used in previous research showed that the location of farms needs to be carefully considered when providing recommendations of management practices to cacao growers. Granier et al. (1999) and Schwendenmann et al. (2010) found that yield reductions of up to 10% could be expected when soil moisture falls below the upper threshold, a situation that occurred regularly especially in the eastern part of the study area without the application of supplemental water. However, soil moisture below the lower threshold was rare. This is important because if water storage drops below this threshold it can lead to yield reductions of up to 45% (Granier et al., 1999; Moser et al., 2010; Schwendenmann et al., 2010), and means that the trees could be at the limit of withering. Extended periods below the lower critical threshold could result in die-off in the cacao plantation.

This study, like others, found that precipitation is a key factor for cacao throughout its development as this crop is highly prone to drought (Adjaloo et al., 2012; Medrano et al., 2015). Temperature also impacts cacao considerably as also shown by Schroth et al. (2016), especially in extremely dry periods where reductions in yields have been observed (Gateau-Rey et al., 2018). Key soil properties identified in this research as well as in previous research include soil organic matter and soil water retention capacity (Gusli et al., 2020; Niether et al., 2017). Although this study used spatial data products generated from local climate and soil information and well-validated modeling approaches, there are limitations. The weather stations from which the climate data were generated are at a low density and the time series at several stations contained numerous gaps. Soil sampling sites were selected within farms assisted by the CfP project, with the exception of the three non-farm locations, and natural parks were excluded as cacao production is not allowed there. The total number of sites and samples were limited due to the resources available for the CfP project. However, the locations of the sampling sites were selected considering their representativeness in terms of soils and landscapes both at the farm and regional levels, a point of great importance in digital soil mapping. There are errors and uncertainties in the resulting digital soil maps due to errors in sampling, sampling density, the quality of the covariates used, and the techniques or algorithms used for the mapping process. Each of these datasets and their associated uncertainties were combined in the water balance model to produce estimates of water availability at a monthly time scale, leading to further uncertainty in the results. There is also a lack of independent data for validation of the modeling results, however, the methods employed passed the validation checks that we were able to conduct and serve as a model for how the water sufficiency of cacao plantations could be used by other researchers in areas with similar levels of data scarcity.

For future analysis, limitations with the soil mapping could be reduced by doing more sampling if more funding were available. In-field validation can be implemented by comparing each farm's characterization based on soil moisture measures with the results obtained in this research. Further work could also include validating the assumptions and results with local technicians, whose experience with cacao management and onsite knowledge of the study farms could help to confirm whether the modeled properties of the farms are representative of the real conditions in the region.

5. Conclusions

Despite the scarcity of weather stations within and around the study area and the high occurrence of missing data, interpolated surfaces of precipitation and temperature were developed for quantifying water scarcity for cacao production. Similarly, testing of the digital soil mapping validated the fundamental inputs for the implementation of site-specific and spatially-distributed water balance, which was consistent with a long-term water balance from a global dataset for the study area. Evapotranspiration estimates also compared well to values from the literature for cacao.

The spatially-distributed water balance allowed us to assess water sufficiency for cacao in the SNSM showing that, under the climate conditions for the analysis period (1989–2018), rainfall was not enough to prevent cacao yield losses for 10 out of the 27 farms evaluated. This suggests that supplemental water would need to be provided or alternative management practices implemented to increase soil water content and, thus, reduce cacao yield losses and avoid its withering. The location of farms in two different departments (La Guajira and Magdalena) with contrasting climate conditions needs to be carefully considered when providing recommendations of management practices to cacao growers. Farms facing less severe water stress conditions for cacao production are characterized on average by higher rainfalls and lower temperatures and are located in the west of the study region at higher elevations with steeper slopes and soils that contain more organic matter, which increases soil water retention.

Temporally, attention should be focused on the months between February and August when the region transitions from the dry season into the first part of the wet season. Water management is most important in March-April as the dry season gives way to the new rainy season, and July-August just before the wettest part of the rainy season. Results obtained through this case study will be shared with technicians who assist cacao growers in the SNSM. As the methods followed in this case study are applicable to other cacao regions, this can help to increase cacao production worldwide and lead to better, more reliable incomes for farmers.

CRediT authorship contribution statement

Valencia Jefferson: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Frankenberger Jane:** Conceptualization, Methodology, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Supervision, Project Administration, Funding Acquisition. **Cherkauer Keith:** Conceptualization, Methodology, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing, Supervision, Project Administration, Funding Acquisition. **Martín-López Javier M.:** Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Monserrate Fredy:** Conceptualization, Methodology, Investigation, Writing - Original Draft. **Da Silva Mayesse:** Conceptualization, Methodology, Resources, Writing - Original Draft, Writing - Review & Editing, Supervision, Funding Acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This research was conducted through the Borlaug International Agricultural Science and Technology Fellowship Program promoted and funded by the U.S. Department of Agriculture - Foreign Agricultural Service (USDA-FAS) under Agreement No. FX19BF-10777R029. The authors thank the U.S. Agency for International Development (USAID) and the USDA-FAS for providing funds and access to valued data for this research study through the Cacao for Peace program.

Disclosure statement

No potential conflict of interest was reported by the authors. Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Agriculture.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2022.101255.

References

Adjaloo, M.K., Oduro, W., Banful, B.K., 2012. Floral phenology of upper amazon cocoa trees: implications for reproduction and productivity of cocoa. ISRN Agron. 2012, 1–8. https://doi.org/10.5402/2012/461674.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Rome, Italy.

de Almeida, A.-A.F., Valle, R.R., 2007. Ecophysiology of the cacao tree. Braz. J. Plant Physiol. 19, 425–448. https://doi.org/10.1590/S1677-04202007000400011.

de Almeida, J., Tezara, W., Herrera, A., 2016. Physiological responses to drought and experimental water deficit and waterlogging of four clones of cacao (Theobroma cacao L.) selected for cultivation in Venezuela. Agric. Water Manag. 171, 80–88. https://doi.org/10.1016/j.agwat.2016.03.012.

Arévalo-Gardini, E., Arévalo-Hernández, C.O., Baligar, V.C., He, Z.L., 2017. Heavy metal accumulation in leaves and beans of cacao (Theobroma cacao L.) in major cacao growing regions in Peru. Sci. Total Environ. 605–606, 792–800. https://doi.org/10.1016/j.scitotenv.2017.06.122.

Argüello, D., Chavez, E., Lauryssen, F., Vanderschueren, R., Smolders, E., Montalvo, D., 2019. Soil properties and agronomic factors affecting cadmium concentrations in cacao beans: a nationwide survey in Ecuador. Sci. Total Environ. 649, 120–127. https://doi.org/10.1016/j.scitotenv.2018.08.292.

Asmamaw, L.B., Mohammed, A.A., Duan, X., Shi, X., Li, Y., Rong, L., Fen, D., Tamene, L., Adimassu, Z., Aynekulu, E., Yaekob, T., Ellison, J., Yaekob, T., Woldearegay, K., Mekonnen, K., Thorne, P., Le, Q.B., Matthews, J.A., Conklin, H.E., Duan, X., Xie, Y., Ou, T., Lu, H., Fenta, A.A., Tsunekawa, A., Haregeweyn, N., Poesen, J., Tsubo, M., Borrelli, P., Panagos, P., Vanmaercke, M., Broeckx, J., Yasuda, H., Kawai, T., Kurosaki, Y., FAO, 2017. Land susceptibility to water and wind erosion risks in the East Africa region. In: Sci. Total Environ, 703, p. 100. https://doi.org/10.1017/CBO9780511807527.014.

Augustin, C., Cihacek, L.J., 2016. Relationships between soil carbon and soil texture in the Northern Great Plains. Soil Sci. 181, 386–392. https://doi.org/10.1097/SS.0000000000000173.

Beck, H.E., de Roo, A., van Dijk, A.I.J.M., 2015. Global maps of streamflow characteristics based on observations from several thousand catchments. J. Hydrometeorol. 16, 1478–1501. https://doi.org/10.1175/JHM-D-14-0155.1.

Beck, H.E., van Dijk, A.I.J.M., Miralles, D.G., de Jeu, R.A.M., Sampurno Bruijnzeel, L.A., McVicar, T.R., Schellekens, J., 2013. Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. Water Resour. Res. 49, 7843–7863. https://doi.org/10.1002/2013WR013918.

Bouyoucos, G.J., 1936. Directions for making mechanical analyses of soils by the hydrometer method. Soil Sci. 42, 225–230. https://doi.org/10.1097/00010694-193609000-00007.

Breiman, L., 2001. Random forests. Mach. Learn 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Carr, M.K.V., Lockwood, G., 2011. The water relations and irrigation requirements of cocoa (Theobroma cacao L.): a review. Exp. Agric. 47, 653–676. https://doi.org/10.1017/S0014479711000421.

CEC, 1993. CORINE Land Cover — Guide Technique (No. Report EUR 12585 EN). Office for Official Publications of the European Communities, Luxembourg.

DNP, 2006. Balance Plan Colombia 1999–2005.

Droogers, P., Allen, R.G., 2002. Estimating reference evapotranspiration under inaccurate data conditions. Irrig. Drain. Syst. 16, 33–45. https://doi.org/10.1023/A:1015508322413.

Esquivel, A., Llanos-Herrera, L., Agudelo, D., Prager, S.D., Fernandes, K., Rojas, A., Valencia, J.J., Ramirez-Villegas, J., 2018. Predictability of seasonal precipitation across major crop growing areas in Colombia. Clim. Serv. 12, 36–47. https://doi.org/10.1016/j.cliser.2018.09.001.

Funk, C.C., Peterson, P.J., Landsfeld, M.F., Pedreros, D.H., Verdin, J.P., Rowland, J., Romero, B.E., Husak, G.J., Michaelsen, J.C., Verdin, A.P., 2014. A quasi-global precipitation time series for drought monitoring. U. S. Geol. Surv. Data Ser. https://doi.org/110.3133/ds832.

Fushiki, T., 2011. Estimation of prediction error by using K-fold cross-validation. Stat. Comput. 21, 137–146. https://doi.org/10.1007/s11222-009-9153-8.

Gateau-Rey, L., Tanner, E.V.J., Rapidel, B., Marelli, J.-P., Royaert, S., 2018. Climate change could threaten cocoa production: effects of 2015-16 El Niño-related drought on cocoa agroforests in Bahia, Brazil. PLOS One 13, e0200454. https://doi.org/10.1371/journal.pone.0200454.

Granier, A., Bréda, N., Biron, P., Villette, S., 1999. A lumped water balance model to evaluate duration and intensity of drought constraints in forest stands. Ecol. Modell. 116, 269–283.
Gusli, S., Sumeni, S., Sabodin, R., Muqfi, I.H., Nur, M., Hairiah, K., Useng, D., van Noordwijk, M., 2020. Soil organic matter, mitigation of and adaptation to climate change in cocoa-based agroforestry systems. Land 9. https://doi.org/10.3390/LAND9090323.
Hargreaves, G.L., Hargreaves, G.H., Riley, J.P., 1985. Agricultural benefits for Senegal river basin. J. Irrig. Drain. Eng. 111, 113–124. https://doi.org/10.1061/(ASCE)0733-9437(1985)111:2(113).
Hawkins, R.H., Theurer, F.D., Rezaeianzadeh, M., 2019. Understanding the Basis of the Curve Number Method for Watershed Models and TMDLs. J. Hydrol. Eng. 24, 06019003. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001755.
Hengl, T., 2020. GSIF: Global Soil Information Facilities.
Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: global gridded soil information based on machine learning. PLOS One 12, e0169748. https://doi.org/10.1371/journal.pone.0169748.
Hernandez, A., Cock, J.H., El-Sharkawy, M.A., 1989. The responses of leaf gas exchange and stomatal conductance to air humidity in shade-grown coffee, tea, and cacao plants as compared with sunflower. Rev. Bras. Fisiol. Veg. 1, 155–161.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Clim. 25, 1965–1978.
Huang, M., Gallichand, J., Wang, Z., Goulet, M., 2006. A modification to the Soil Conservation Service curve number method for steep slopes in the Loess Plateau of China. Hydrol. Process. 20, 579–589. https://doi.org/10.1002/hyp.5925.
Huffman, G.J., Stocker, E.F., Bolvin, D.T., Nelkin, E.J., Tan, J., 2019. GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06. https://doi.org/10.5067/GPM/IMERG/3B-HH/06.
Hutchinson, M.F., de Hoog, F.R., 1985. Smoothing noisy data with spline functions. Numer. Math. 47, 99–106. https://doi.org/10.1007/BF01389878.
IDEAM, 2015. Atlas Climatológico de Colombia.
IDEAM, 2017a. Precipitación Media Total Anual Promedio Multianual durante el periodo 1981–2010.
IDEAM, 2017b. Temperatura Media Anual Promedio Multianual durante el periodo 1981–2010.
Jasiewicz, J., Stepinski, T.F., 2013. Geomorphons-a pattern recognition approach to classification and mapping of landforms. Geomorphology 182, 147–156. https://doi.org/10.1016/j.geomorph.2012.11.005.
Köhler, M., Schwendenmann, L., Hölscher, D., 2010. Throughfall reduction in a cacao agroforest: tree water use and soil water budgeting. Agric. . Meteorol. 150, 1079–1089. https://doi.org/10.1016/j.agrformet.2010.04.005.
Köhler, M., Dierick, D., Schwendenmann, L., Hölscher, D., 2009. Water use characteristics of cacao and Gliricidia trees in an agroforest in Central Sulawesi, Indonesia. Ecohydrology 2, 520–529. https://doi.org/10.1002/eco.67.
Köhler, M., Hanf, A., Barus, H., Hendrayanto, Hölscher, D., 2014. Cacao trees under different shade tree shelter: effects on water use. Agrofor. Syst. 88, 63–73. https://doi.org/10.1007/s10457-013-9656-3.
Liaw, A., Wiener, M., 2003. Classification and regression by randomForest. R News 2. R News 3, 18–22.
Libohova, Z., Martín-López, J.M., da Silva, M.A., Lagoueyte, C., Cruz, J., Drohan, P., Maximova, S., Guiltinan, M., Ferruzzi, M.G., Guarín, D., Reich, P., Kome, C., Zapata, Y.P., Gallego-Sánchez, G., Quintero, C., Botero, C., Winters, N.P., Robotham, M., 2020. Soil and cacao genomics survey of Sierra Nevada de Santa Marta Region, Colombia.
Ma, Y., Minasny, B., Welivitiya, W.D.D.P., Malone, B.P., Willgoose, G.R., McBratney, A.B., 2019. The feasibility of predicting the spatial pattern of soil particle-size distribution using a pedogenesis model. Geoderma 341, 195–205. https://doi.org/10.1016/j.geoderma.2019.01.049.
McBratney, A., Mendonça Santos, M., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.
Medrano, H., Tomás, M., Martorell, S., Flexas, J., Hernández, E., Rosselló, J., Pou, A., Escalona, J.-M., Bota, J., 2015. From leaf to whole-plant water use efficiency (WUE) in complex canopies: limitations of leaf WUE as a selection target. Crop J. 3, 220–228. https://doi.org/10.1016/j.cj.2015.04.002.
Minasny, B., McBratney, A.B., 2016. Digital soil mapping: a brief history and some lessons. Geoderma 264, 301–311. https://doi.org/10.1016/j.geoderma.2015.07.017.
Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57, 443–452. https://doi.org/10.2136/sssaj1993.03615995005700020026x.
Moraes, A.G. de L., Bowling, L.C., Zeballos-Velarde, C.R., Daneshvar, F., Watkins, A.H., Cherkauer, K.A., 2022. Terrain sensitive climate mapping for the Arequipa Department in Peru. Int. J. Climatol. 1–23. https://doi.org/10.1002/joc.7730.
Moser, G., Leuschner, C., Hertel, D., Hölscher, D., Köhler, M., Leitner, D., Michalzik, B., Prihastanti, E., Tjitrosemito, S., Schwendenmann, L., 2010. Response of cocoa trees (Theobroma cacao) to a 13-month desiccation period in Sulawesi, Indonesia. Agrofor. Syst. 79, 171–187. https://doi.org/10.1007/s10457-010-9303-1.
Motamayor, J.C., Risterucci, A.M., Lopez, P.A., Ortiz, C.F., Moreno, A., Lanaud, C., 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity 89, 380–386. https://doi.org/10.1038/sj.hdy.6800156.
Navarro-Racines, C., Tarapues, J., Thornton, P., Jarvis, A., Ramirez-Villegas, J., 2020. High-resolution and bias-corrected CMIP5 projections for climate change impact assessments. Sci. Data 7, 7. https://doi.org/10.1038/s41597-019-0343-8.
Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009. College Station, USA.
Niether, W., Schneidewind, U., Armengot, L., Adamtey, N., Schneider, M., Gerold, G., 2017. Spatial-temporal soil moisture dynamics under different cocoa production systems. Catena 158, 340–349. https://doi.org/10.1016/j.catena.2017.07.011.
Odgers, N.P., Libohova, Z., Thompson, J.A., 2012. Equal-area spline functions applied to a legacy soil database to create weighted-means maps of soil organic carbon at a continental scale. Geoderma 189–190, 153–163. https://doi.org/10.1016/j.geoderma.2012.05.026.
Oppong, F.K., Opoku-ameyaw, K., Osei-bonsu, K., Amoah, F., Brew, K.M., Oppong, R., 1999. The effect of time of planting at stake on cocoa seedling survival. Ghana Jnl agc. Sci 32, 79–86.
Pendleton, R.L., Jenny, H., 1945. Factors of Soil Formation: A System of Quantitative Pedology. Geogr. Rev. 35, 336. https://doi.org/10.2307/211491.
Radersma, S., de Ridder, N., 1996. Computed evapotranspiration of annual and perennial crops at different temporal and spatial scales using published parameter values. Agric. Water Manag. 31, 17–34. https://doi.org/10.1016/0378-3774(95)01235-4.
Raes, D., 2002. BUDGET - A soil water and salt balance model. Reference Manual.
Rajab, Y.A., Hölscher, D., Leuschner, C., Barus, H., Tjoa, A., Hertel, D., 2018. Effects of shade tree cover and diversity on root system structure and dynamics in cacao agroforests: The role of root competition and space partitioning. Plant Soil 422, 349–369. https://doi.org/10.1007/s11104-017-3456-x.
Saxton, K.E., Rawls, W.J., 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. Soil Sci. Soc. Am. J. 70, 1569–1578. https://doi.org/10.2136/sssaj2005.0117.
Schroth, G., Läderach, P., Martinez-Valle, A.I., Bunn, C., Jassogne, L., 2016. Vulnerability to climate change of cocoa in West Africa: patterns, opportunities and limits to adaptation. Sci. Total Environ. 556, 231–241. https://doi.org/10.1016/j.scitotenv.2016.03.024.
Schwendenmann, L., Veldkamp, E., Moser, G., Hölscher, D., Köhler, M., Clough, Y., Anas, I., Djajakirana, G., Erasmi, S., Hertel, D., Leitner, D., Leuschner, C., Michalzik, B., Propastin, P., Tjoa, A., Tscharntke, T., van Straaten, O., 2010. Effects of an experimental drought on the functioning of a cacao agroforestry system, Sulawesi, Indonesia. Glob. Chang. Biol. 16, 1515–1530. https://doi.org/10.1111/j.1365-2486.2009.02034.x.
Siebert, S., Döll, P., 2010. Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. J. Hydrol. 384, 198–217. https://doi.org/10.1016/j.jhydrol.2009.07.031.
Singh, K., Fuentes, I., Fidelis, C., Yinil, D., Sanderson, T., Snoeck, D., Minasny, B., Field, D.J., 2021. Cocoa suitability mapping using multi-criteria decision making: An agile step towards soil security. Soil Secur. 5, 100019 https://doi.org/10.1016/j.soisec.2021.100019.
Thornthwaite, C.W., Mather, J.R., 1955. The water balance. Publ. Climatol. 8.
Ulmen, C., 2000. Report No. 26 Modelling raster-based monthly water balance components for Europe, GRDC No. 26.

USDA, 2016. Cacao for Peace (CfP) Overview.
USDA-SCS, 1986. Urban Hydrology for Small Watersheds TR-55. U.S. Dept. Agric. Nat. Resour. Conserv. Serv. Conserv. Eng. Div. 164.
Valencia, J., Frankenberger, J., Cherkauer, K., Martín-López, J.M., Monserrate, F., da Silva, M.A., 2022. Post-processed climate dataset for the Sierra Nevada de Santa Marta (SNSM) Region in Colombia. Harvard Dataverse, V1. https://doi.org/10.7910/DVN/YNNXO1.
Walkley, A., Black, I.A., 1934. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Sci. 37, 29–38. https://doi.org/10.1097/00010694-193401000-00003.
Xie, E., Zhang, Y., Huang, B., Zhao, Y., Shi, X., Hu, W., Qu, M., 2021. Spatiotemporal variations in soil organic carbon and their drivers in southeastern China during 1981-2011. Soil Tillage Res. 205, 104763 https://doi.org/10.1016/j.still.2020.104763.
Yang, R., Liu, F., Zhang, G., Zhao, Y., Li, D., Yang, J., Yang, Fei, Yang, Fan, 2016. Mapping soil texture based on field soil moisture observations at a high temporal resolution in an oasis agricultural area. Pedosphere 26, 699–708. https://doi.org/10.1016/S1002-0160(15)60078-9.
Zhu, A.-X., 1997. A similarity model for representing soil spatial information. Geoderma 77, 217–242.
Zhu, A.-X., Hudson, B., Burt, J., Lubich, K., Simonson, D., 2001. Soil mapping using GIS, expert knowledge, and fuzzy logic. Soil Sci. Soc. Am. J. 65, 1463–1472. https://doi.org/10.2136/sssaj2001.6551463x.
Zuidema, P.A., Leffelaar, P.A., Gerritsma, W., Mommer, L., Anten, N.P.R., 2005. A physiological production model for cocoa (Theobroma cacao): model presentation, validation and application. Agric. Syst. 84, 195–225. https://doi.org/10.1016/j.agsy.2004.06.015.