begindocument/before

Article

A modelling strategy to improve cacao quality and productivity

Angela P. Romero V. [1,‡,*], **Anyela V. Camargo R.** [1,‡], **Oscar D. Ramirez** [3], **Paula A. Arenas V** [3] **and Adriana M. Gallego** [2]

[1] The John Bingham Laboratory, NIAB, 93 Lawrence Weaver Road, Cambridge CB3 0LE, UK 1
[2] Grupo BIOS, Centro de Bioinformática y Biología Computacional de Colombia - BIOS, Manizales, Caldas, Colombia, South America.
[3] Federación Nacional de Cafeteros FEDECACAO, Colombia, South America.
* Correspondence: angela.romerovergel@niab.com
‡ These authors contributed equally to this work.

Abstract: Cacao production systems in Latin America have a high importance over social and economic development, facing the fight against hunger and poverty. Although Colombian cacao has the potential to be in the high value markets for fine flavour, the lack of expert support as well as the use of traditional, and oftentimes suboptimal, technologies makes cocoa production negligibly. For example, traditionally cacao harvest takes place at 5 or 6 months after flowering, other environmental parameters that have more association with pod maturation speed are not taken into account. Cocoa fruits development can be considered as the result of a number of physiological and morphological processes that can be described by mathematical relationships even under uncontrolled environments. In this context, crop models are useful tools to simulate and predict crop development over time and under multiple environmental conditions. Since, harvesting at the right time can yield high quality cacao, we parametrised a crop model to predict the best time for harvest cocoa fruits in Colombia. Our aim was to develop a practical tool that supports cacao farmers in the production of high quality cacao. When comparing simulated and observed data, our results showed an RRMSE of 7.2% for the yield prediction, while the simulated harvest date varied between +/- 2 to 20 days depending on the temperature variations of the year between regions. This crop model contributed to understand and predict the phenology of cacao fruits for two key cacao varieties ICS95 y CCN51.

Keywords: ICS95; CCN51; thermal time, flowering date

Citation: Romero, A.P.V; Camargo, A.; Arenas P.; Ramirez O. D.; Gallego, A.; Title. *Journal Not Specified* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

1. Introduction

Cocoa (*Theobroma cacao* L.) is an important worldwide perennial tropical crop endemic to the South American rainforests [1–4]. Cacao plant is a member of the Malvaceae (formerly Sterculiaceae) botanical family such as cotton *Gossypium hirstium* [5]. Coton has been is modeled in SIMPLE model [6]. Cocoa is grown for its fruits, known as cacao pods [7,8]. Only the 5% of the world cocoa yield is destined for Fine-cocoa production due to the low productivity through the traditional crop management [3]. In Colombia, cocoa is traditionally consumed as a beverage. It is one of the crops promoted by the Colombian government in the social and agricultural development programs aimed at favouring peace in post-conflict regions [4,9]. This crop is grown by approximately 52.000 families [10] and 98% of production being carried out by small and medium-sized producers [11,12]. Colombia registered an increase of 3.750 tons in production in 2020 compared to the previous year [13].

Although Colombian cocoa has the potential to be in the high value markets for fine flavour [12], the lack of expert support as well as the use of traditional, and often times suboptimal, technologies makes cocoa production negligibly. For example, the farmer practice of empirically harvesting from 5 to 6 months or 180 days after flowering date (DAF), produced a mix of quality of cocoa beans that are then be fermented. This practice produces heterogeneous characteristics between each fermentation batch which potentially diminish the quality of the end cocoa product [14]. To identify the best moment to harvest is important to consider physiological responses affected by climate variables such as rain, solar radiation and wind. Thus, for cocoa in Colombia, physiological simulation models may be valuable to identify the best moment to harvest cocoa considering variable weather conditions, soil types and cultivar specifications.

Crop models represent a quantitative assumption of plant growth depending on sunlight interception efficiency values and climate data supported by a large amount of empirical and ground data. [15]. Physiological crop models have shown to be very useful tool for provided agronomical advices and improvements of the cropping systems of annual crops mainly. Recently crop modelling studies are focusing on perennial crops production [1,6,16,17]. However, the information reported is be less than for annual crops due to the lack of field data available , relatively high research costs and the difficulties of accumulated errors in long-term simulations [1]. For cacao there approaches to predict yield mainly using algorithms of machine learning [13] and just one mechanistic model simulates physiological cocoa performance "SUCROS-cocoa" [1]. This crop model calculates light interception, photosynthesis, maintenance respiration, evapotranspiration, biomass production and cocoa yield. It can be parametrised having data on cocoa physiology and morphology [1]. However, there is not specific cocoa physiology data available from small and medium-sized producers. Thus, we adapted the simple generic crop model (SIMPLE) that could be easily modified for any crop to simulate development, crop growth and yield using few parameters such as weather and cultivar specification [6].

In this paper, we present a physiological parametrisation of SIMPLE crop model for cocoa to predict best harvest time and overall yield. We parameterised the SIMPLE crop model [6] because it had already been successfully fitted to other tree crops in south America. The model simulates crop development, growth and yield, and predicts the maturation day when the fruit is likely to be ready for harvest. Our aim was to develop a practical tool that supports cacao farmers in the production of high quality cacao. Thus, making Colombian cacao more competitive in the Fine-cocoa market.

2. Materials and Methods

2.1. Floral Phenology of Cocoa

Usually, the phenological stages of a cacao tree are divided in two main phases: vegetative and reproductive. Defining the reproductive phase (fig. 1, a) as described by the floral phenology from the date of inflorescence emergence (BBCH scale 5) (fig. 1,b) to the date of ripening of fruit and seed (BBCH 8) (fig. 1,c) [7]. The reproductive phase in the Andean region of Colombia is cyclically fulfilled during two annual cycles which goes through the following phases: inflorescence emergence, flowering, pollination, fruit development and harvest. Therefore, for modelling parameterisation the crop cycle of cocoa as perennial plant does not start at the plantation date such as annual crops systems. Instead, the start point of the cocoa crop cycle is the inflorescence emergence date (fig. 1,b). Consequently , the growth period of the fruit can vary from 110 to 150 daa (days after anthesis) [18] when cacao fruits reach the physiological maturity, but it can be harvested at 170 days daa [7] for quality purposes.

Cocoa is a cauliflorous plant, which means that flowers grow on the trunk and branches. Cocoa trees usually produce up to 10000 flowers per tree each year, 50 % of them do not develop into ripe fruits (Personal communication with Fedecacao). Flower development takes approximately 30 days across 12 micro-stages, from meristem development (stages 1 to 6) to the fully developed flower (Stages 7 to 12) [18,19] when it is ready to be pollinated. The opening of flowers or anthesis occurs over a 12-hour period during the night and it is synchronised between the groups of mature flowers [7]. However, the live of a flower can last approximately 1 day after the opening falling form the trunk if it is unfertilised [7,20]. Subsequentially, after anthesis, fruits growth for approximately 150 days until the maturation, mucilage. Therefore, the complete maturation process of the fruit, from the pollination to fully mature fruit, takes 160- 210 days [21]. The accumulation of lipids, storage proteins and anthocyanin start about 85 days after pollination when fruits have an active metabolism and seeds moisture content decreases up to 30% [7,22]. During this phase the quality of cocoa seeds is defined.

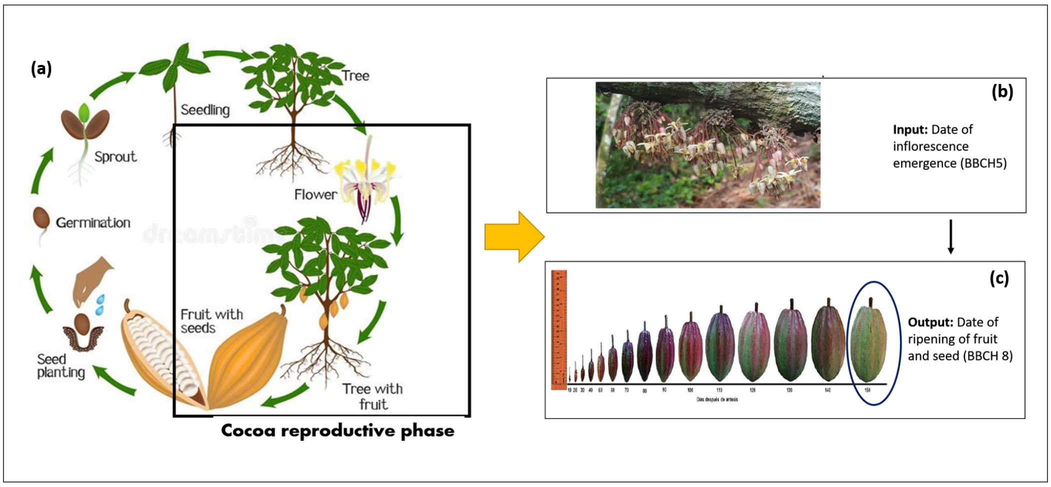

Figure 1. Phenology of cocoa in Colombia for crop modelling.

Source: Taken from from Dreamstime.com, phys.org [23] and [18]

2.2. Inputs and Data Acquisition

Input variables required to run the cocoa model included the flowering date and daily weather of solar radiation (SRAD), maximum and minimum temperature (TMAX, TMIN) and rain. Therefore, weather data as csv file was downloaded from the POWER Data Access Viewer [24] from January 01 of 2018 to December 31 of 2020 for the five locations in Colombia (Fig.2, b). The csv file had to be transformed to .WHT file using R 1.4 version [25].

Field data for this study was provide by Fedecacao (National Cacao Producers Federation). Reports contained information of 23 flowering dates from 12-July-2019 to 23-June-2020. Each of these dates had their corresponding date of harvest after at exactly 180 DAF, the age of the trees, plant density (trees ha^{-1}), yield (dry beans kg ha^{-1}) and number of fruits harvested per hectare. Thus, 23 flowering dates for five farms gave us 115 samples in total.

2.3. Test Site and Yield Production

Five farms in different regions were tested : Saravena (Arauca), Rionegro (Santander), Cali (Valle del Cauca), Apartado (Antioquia) and Manizalez (Caldas). These were considered because had the data available of production in the reports of Fedecacao. Yield depend on the successfully development of flowers to form ripe pods. According to personal communication from farmers, the highest flowering season occurs in September and January suggesting that harvest occurs in March and July. However, the data collected from the farms suggested that flowering and pod production were not constant for all the locations. For example, pod harvest in Caldas farm increased in May and from October to December, and decreased from January to March. Meanwhile, the Arauca and Apartado farms reported the highest yield in the months January, July, November and December. Santander farm had picks of production in March, May and September (Fig.2, a). However, cacao is cultivated in 30 regions of Colombian which accounts for approximately about 147,000 ha [26]. Therefore, the national production of cocoa beans in 2020 was 63.416 ton according to information published by Fedecacao ([27]). The highest yield was reposted by Santander region with 26,315, following in

124 descending order by Antioquia with 5.974 tons, Arauca with 5.082 tons, Caldas with
125 1.343 tons and Valle del Cauca with 339 tons.

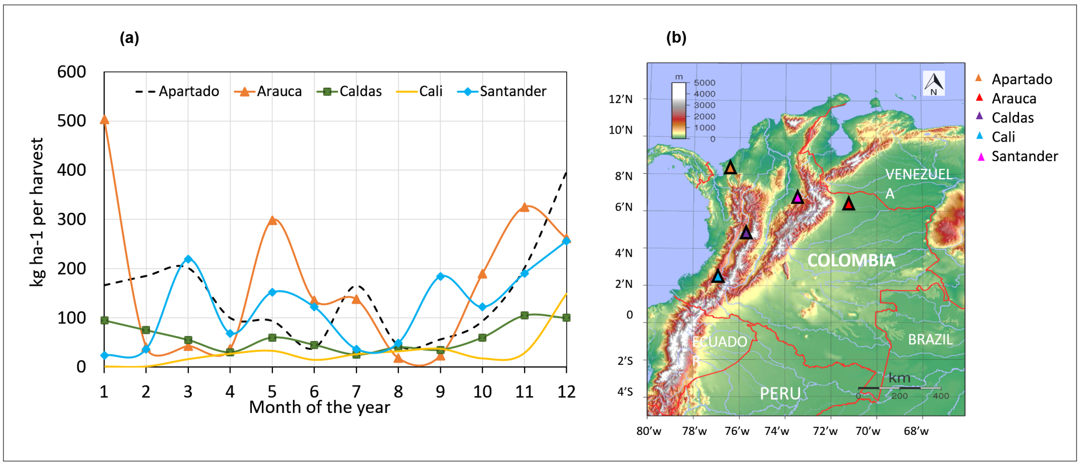

Figure 2. Cocoa production of five farms in different regions.

(a) Production per month (2019 - 2020) of five farms. (b) Map showing the regions where farms are located.

2.4. Thermal Time for Pod Harvest Date Identification

The cumulative sum of daily temperature from a reference day 0 is defined as *Thermal time* and its units of measurement are in days degrees (days °C). That starter point of 0 days °C generally is the planting date [28] but for cocoa crop simulation model it is the flowering date. Thermal time of a cultivated plant may consider the base temperature (T_b), which is the minimum temperature required by cocoa plant to grow. T_b an vary between cultivars [29,30]. For cocoa the vegetative growth T_b has been reported between 18.6 and 20.8 °C [30]. Nevertheless, the pod growth has a lower T_b which range between 9 and 12.9 °C [30,31]. We calculated the cocoa thermal time with a pod growth T_b of 10 °C because it is the absolute minimum temperature for cocoa growing in South America reported by [32], in [31].

The thermal time required for the crop model was characterised for each location starting from flowering date (0 days °C) to harvest date (180 after flowering = 6 months), as farmers used to harvest by calendar days. Thus, the cocoa model predicts the maturation day to harvest pods. It can vary depending on temperature variations. Thermal time was calculated using the equation 1. Where tt is the cumulative sum of the daily temperature (T_i) and T_b for cocoa is 10°C.

$$Thermal\ time\ (tt) = \begin{cases} \sum_{i=1}^{n} T_i - T_b \\ \\ 0, \ \ Flowering\ date \end{cases} \tag{1}$$

2.5. Model Calibration

Calibration of crop models are conducted typically for particular cultivars and require site specific inputs of weather [33]. The procedure for the SIMPLE model [6] calibration was a sequential process of modifying physiological variables specific for cocoa in the inputs files and adding the appropriate files of weather for each region tested. In the SIMPLE model has seven input files where new cocoa crop data should be provided. Files in the list below can be edited to define the features of the new cultivars or experiments. Cocoa crop has not been simulated with SIMPLE model, hence consider previous cocoa studies, values such as leaf area index (LAI) [34] and Harvest index (HI) [35] where modified. Radiation Use Efficiency (RUE) was calibrate according to the perennial crops (banana and cotton) that had been calibrated previously in SIMPLE model [6] with a RUE of 0.8 and 0.85 respectively. As cocoa trees are under shadow the RUE was lower with values between 0.7 and 0.5 g MJ^{-1} m^2 (table 1). Once the physiological parameters were calibrated to the simulated yields for cocoa were reasonably close to the observed yield. The 23 flowering from Fedecacao reports were introduced in the treatment file, running the program and saving the results.

2.6. Parameters

This cocoa model has were three parameters which vary by region (table 1): The thermal time required for harvest after the flowering date (Tsum) , the Radiation Use Efficiency (RUE) and yield observed on field. Physiological parameters in table 2 are common for all the regions studied. These parameters were calibrate for cultivars ICS95 and CCN51 considering a range of time of 200 DAF from flowering date to harvest day, even thought farmers collect the pod at 180 DAF. Heat and water stress parameters were not considered.

Table 1: Cocoa crop parameter values used per region.

Region	Tsum	RUE	Yield*
Apartado	2906	0.6	3378
Arauca	2764	0.7	3981
Santander	2016	0.6	2687
Cali	1912	0.5	1900
Caldas	1192	0.6	740

RUE Radiation use efficiency (above ground only and without respiration)g MJ^{-1} m^2
* Yield observed kg ha^{-1} per year.

Table 2: Parameter values used to run SIMPLEcocoa model.

File	Variable name	Value
	SoilName	Loamy sand4
	InitialFsolar	0.01
Treatment	Weather	KOKO (.WTH file name)
	CO_2	400 ppm
	SowingDate	Flowering date
	Crop cycle DAP	200 days
	LAI	1.8
Observation	FSolar	0.70
	Biomass	40kg dry mass per plant
	Harvest index	0.3
Cultivar	150A	680 °C day
	150B	680 °C day
	Tbase	10°C
	Topti	26°C
Species	MaxT	35°C
	ExtremeT	40°C
	CO_2RUE	0.09°C
	S-water	0 ARID index

S-water is associated drought stress evaluations ranging from 0 (no water shortage) to 1 (extreme water shortage) [6]

2.7. Evaluation of Model Performance

The cocoa model performance was evaluated by comparing simulated values cocoa yield with those reported by Fedecacao from cocoa plantations, using the statistical indice of relative root mean square error (RRMSE) (Equ. 2) [6,16].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)^2} \tag{2}$$

3. RESULT

3.1. Weather Conditions Over Flowering Time

The weather data from NASA platform is showed in the figure 3. Data from 2018 to 2020 were analysed to show the tendencies by two years when the field data was reported. Looking at the data, Santander and Caldas had the biggest variability and the maximum of solar radiation values over 20 MJ $m^2 day^{-1}$. In contrast , Cali and Arauca presented the lowest values of PAR below 5 MJ $m^2 day^{-1}$. (Fig.3, a). Even though, Cali and Santander had contrasting PAR conditions, they regions presented have similar temperature during 2020 (Fig.3, b). The temperature ranges from 16 to 28 °C and it is relatively constant for each region. However, Arauca presented the biggest variability with hotter months during the first half of the year 2019 and 2020. Apartado was found as the hottest region studied with 26 °C and Caldas as the coldest site with 18 °C. Precipitation in Colombia is presented in two seasons per year from February to April and from October to November, while the relative humidity remain constant over 80% (Fig.3,c). In general, the coldest regions tested (Caldas and Santander) had the maximun values of solar radiation available for photosynthesis.

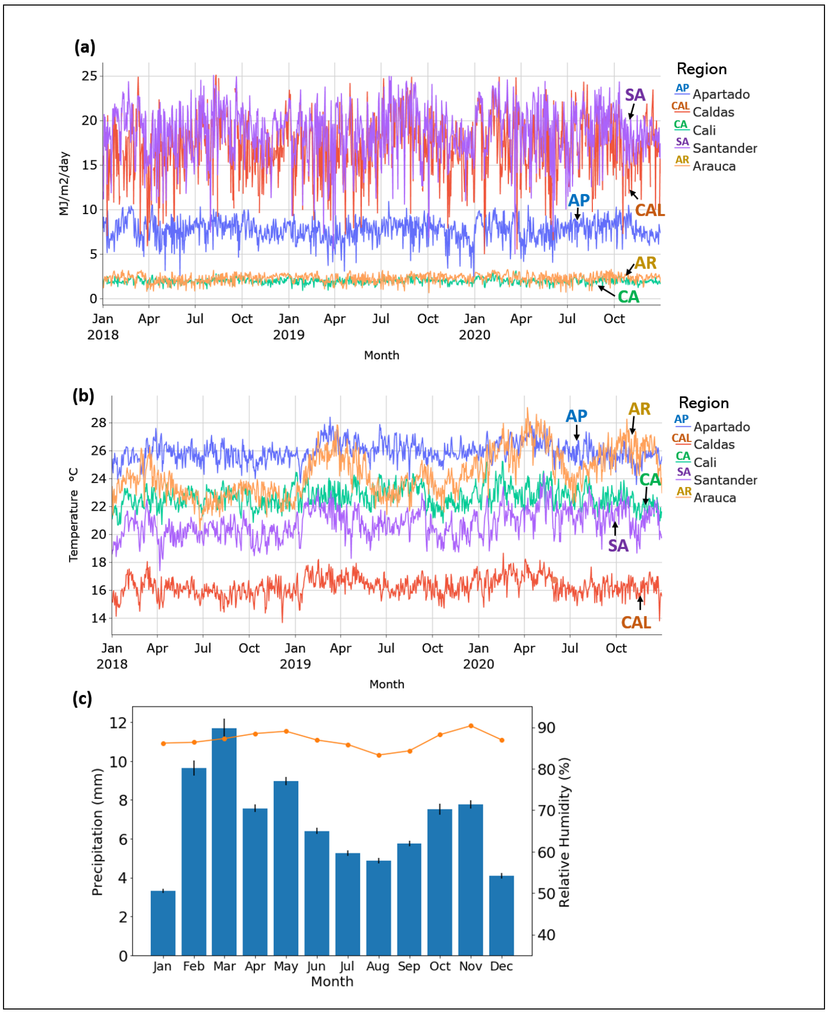

Figure 3. Colombian weather conditions.

(a), Available photosynthetic solar radiation (PAR). (b), Daily average temperature (c), Monthly average of precipitation (bars) and relative humidity (dotted line) from 2018 to 2021

Figure 4 shows the pearson correlation to study the weather data of the flowering time over the months of flowering (monthF), month of harvest (monthH) and their final yield (fruit_kg). The results showed that thermal time T_b (ttb) is 0.52 correlated with daily average temperature and maximum temperature (TMAX) and temperature minimum (TMIN) and Dew Frost Point at 2 meters (T2MDEW) with a correlation coefficient of 0.60. The wind (WS2M) is correlated with ttb with 0.57. However, less clear correlations were found of monthF with T2MDEW, relative humidity (RH), WS2M and rain.

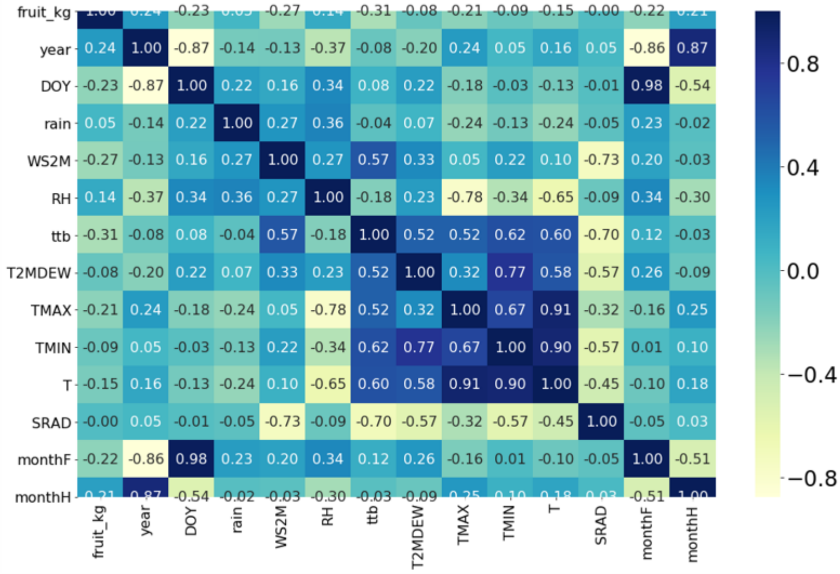

Figure 4. Pearson correlation average weather variables and flowering dates for five locations in Colombia. Numbers in the squares are the correlation coefficients

3.2. Thermal Time

Thermal time characterisation was made considering 180 DAF for each location. The boxplot in the figure 5 shows the data distribution where boxes indicate the range of the central 50% of the total data per region, the central line in the box is marking the median value and lines draw out from each box mean the range of the remaining data. Therefore, this boxplot shows differences between locations as was expected following the tendencies of the temperature (fig. 3,c). Apartado and Arauca had the highest temperatures hence, the highest thermal time values with 2909 and 2764 days°C respectively. Caldas had the lowest values with 1173 days°C. Meanwhile, Cali and Santander presented similar thermal time around 2000 °C (table 1). The accumulated temperature during the pod development (fig.5) depends on the region where cocoa is cultivated, and the variety planted. Thermal time values are also proportional to the yield reported on field (table 1).

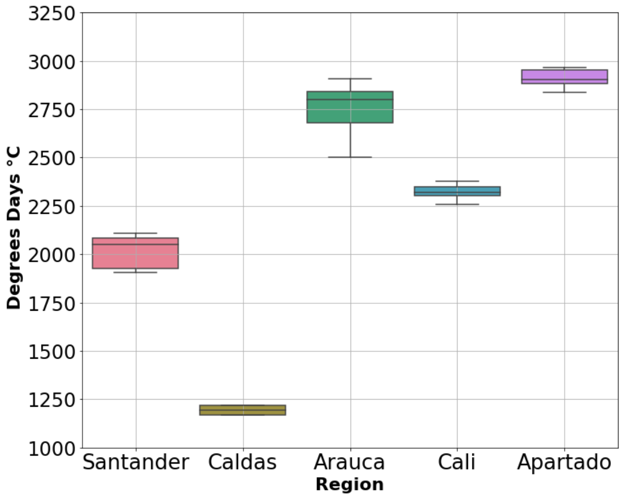

Figure 5. Cocoa yield and thermal time characterisation at 180 days after flowering.

3.3. Model Validation

Biomass production of the aerial part of the cocoa plant was simulated which include every organ of the plant that is over the soil surface. It is important to calculated how much biomass form the aerial part belongs to the the pods according to the harvest index (HI) (table 2), hence cocoa seeds. The figure 6 (a) shows the daily biomass growth rate. Biomass simulation is affected by the solar radiation, RUE, daily temperature, atmospheric CO_2 concentration (ppm) and the fraction of solar radiation intercepted by a tree of cocoa during the fruit development (fSolar) (fig. 6, b).There were not field data of absorption of solar radiation by the plants or biomass production then there is not observed data to compare the simulated data. WHAT KIND OF DATA?? I THOUGHT YOU DOWNLOADED DATA FROM NASA. AR: Just weather data is from NASA WHAT DO YOU MEAN BIOMASS RESPONDED? AR: I meant that it is proportional to the yield. I DIDN'T UNDERSTAND HERE, DROP THE THE FROM BIOMASS BECAUSE WE ARE NOT TALKING ABOUT A SPECIFIC BIOMASS AR: it is biomass of the aerial part (every organ of the plant that is over the soil surface). However, the daily biomass growth rate increased proportionally with the yield production (Fig. 6, c) . Thus, it was possible to calculated the final yield as the product of accumulated biomass at the harvest day and HI = 0.3 (Equ.3).

$$Cocoa\ yield = Accumulatedbiomass\ X\ HI \tag{3}$$

(HI) (Eq. (5))

The fraction of solar radiation intercepted cocoa trees during the fruit development (fSolar) was also simulated during the fruit development cycle. The results in the figure 6 (b) showed that all regions had the maximum fSolar at 0.94, except Caldas where crops reached 0.76. Apartado and Arauca reached faster this fSolar-max. These high values of solar radiation intercepted for photosynthesis, lasted differently depending on the region and their solar radiation (Fig.3, a) : Apartado 66 days from 69 to 135 DAF, Arauca 61 days from 72 to 133 DAF, Cali 38 days from 83 to 121 DAF, Santander 24 days from 92 to 116 DAF and Caldas 2 days at 125 DAF. fSolar declined in the interception of solar radiation until the pod harvest day.

Cocoa yield simulation (kg ha^{-1} per year) was validated using observed data (table 1) showing a final RRMSE of 7.2% (fig. 6, c). Individual errors per region are presented in table 3, where the best fit of the calibration model for yield prediction was for crops in Apartado and the highest error was calculated for Caldas crops. The model responded to the variations of temperature and solar radiation. Therefore, the highest yield values simulated were obtained for Arauca over 4000 kg ha^{-1}, followed by Apartado Santander with yields over 2000 kg ha^{-1}. The lowest yield was simulated for Caldas region with less of 1000 kg ha^{-1} . Final yield in the model was is calculated as the product of biomass of aerial part and harvest index (HI) [6,36], where the HI is similar to the CropSyst [37] and AquaCrop [38].

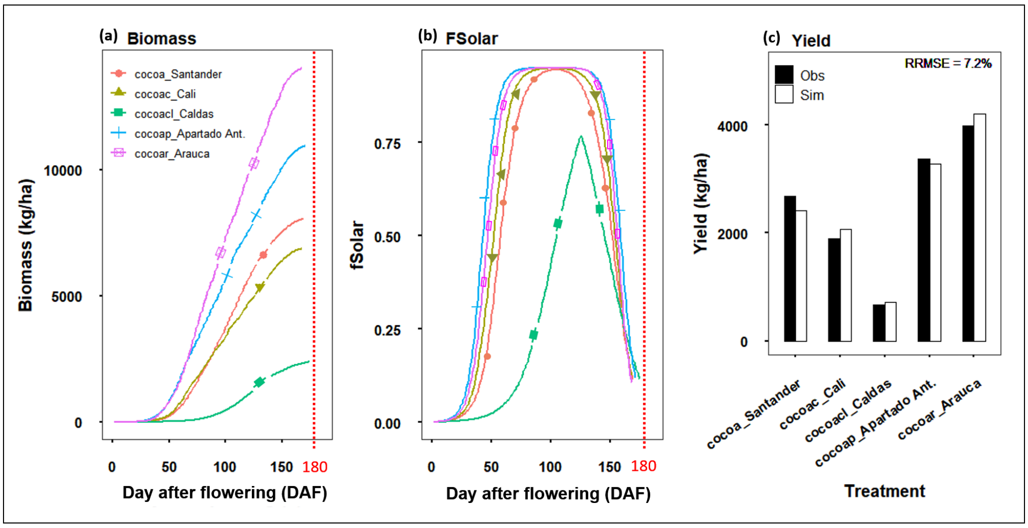

Figure 6. Model predictions. (a) Biomass aerial part. (b) Interception of solar radiation. (c) Yield. Crop cycle close to 180 DAF (vertical red line) base on figure 5.

Table 3: Summary of relative root mean square error RMSE for yield prediction using The cocoa model.

Region	Apartado	Arauca	Santander	Cali	Caldas	Overal
RMMSE %	3	6.05	10.06	8.5	14.90	7.2

3.4. Predicting Pod Harvest Day

The cocoa model predict the day when pods can be harvested (predicted) and the results were compared with the 180 counting after flowering (observed) because it is the empirical date of harvest. The figure 7 present that observe day of harvest (180 DAF) independently of the region. All the regions except Santander, presented the earliest predicted harvest when the flowering date was between December and February. The fruit can be ready to harvest before or after 180 DAF depending on the environmental condition per region. The results demonstrated that the traditional way to harvest which is always at 180 DAF, it is not having account physiological and environmental conditions that can be affecting the pod maturation. Thus, when the fruit development was simulated the maturity day in Cali and Caldas were from 3 to 12 and 4 to 23 days before 180 DAF (Fig. 7, c and f respectively). Apartado presented the most similar predicted dates of harvest to 180 DAF with 170 to 182 DAF (Fig. 7, d). The pod may be harvest in Arauca (Fig.7, b) between 165 and 193 DAF, Santander (Fig.7, a) between 165 and 183 DAF, ten days less than in Arauca for the same months of flowering of July and August of 2019 . Only Arauca presented longer crop cycles when the flowering was between during that period of time. This means that Arauca had bigger variation of temperature between months (Fig. 3, b). To summarize, depending on which month of the year the trees are flowering, the number of days to reach the harvest of ripe pods vary more or less 180 days (table 4).

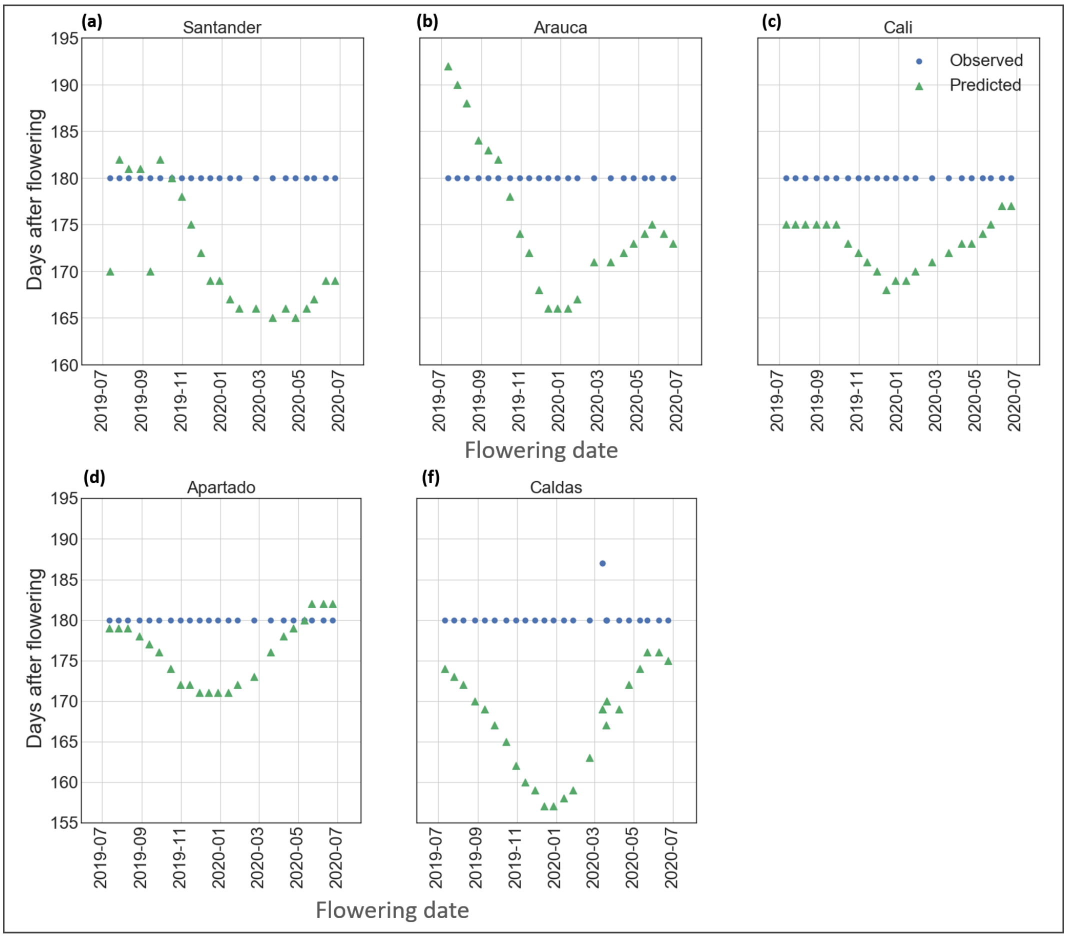

Figure 7. Harvest day prediction from flowering date for (a) Santander, (b) Arauca, (c) Cali, (d) Apartado,(f) Caldas.

.

Table 4: Average of days to harvest according to the month of flowering.

Month	Santander	Arauca	Cali	Apartado	Caldas
January	166.5	166.5	169.5	171.5	158.5
February	166	171	171	173	163
March	165	171	172	176	167
April	165.5	172.5	173	178.5	170
May	166.5	174.5	174.5	181	175
June	169	173.5	177	182	175.5
July	176	191	175	179	173.5
August	181	186	175	178.5	171
September	176	182.5	175	176.5	168
October	179	176	172.5	173	163.5
November	173.5	170	170.5	171.5	159.5
December	169	166	168.5	171	157

Days to harvest cocoa after flowering are approximate, as a result of the cocoa model simulations. Calibration was based on FEDECACAO reports from 2018 to 2020.

4. Discussion

4.1. Weather Effects over Flower Stability and Pollination

This study presents a new approach of SIMPLE model calibration for cocoa as tropical crop in South America for five environments. The analysis of the weather conditions (Fig. 4) did not confirmed clearly that flower are affected by the wind and rain by mechanical damage. However, farmers stated that the number of flowers pollinated decrease by months where wind and rain are high. Moreover, wind can affect the availability of tiny flies pollinators from Diptera order and from the families of of the biting midges *Ceratopogonidae*, genus *Forcipomyia* [39–41] to reach the cocoa flowers. However, the stability of cocoa flowers is influenced by seasonal wheather conditions (abiotic) and pollination (biotic) [42]. Therefore, pollinator population shoud be coincidence with the phenology of the flowering cocoa trees [43,44]. Flower opening is very well synchronised between the cohorts of mature flowers opening each night [7]. The flowers open at almost exactly the same time and rate, irrespective of their position on the trunk. Thus, unfertilised flowers abscise from the trunk approximately 1 day after flower opening [7]. Hence more than 90% of unpollinated flowers fall or abscised within 32 hours after anthesis [45]. Abscission processed of flowers are mainly controlled by three hormones: auxin, ethylene, and abscisic acid (ABA) [45]. Ethylene generally promotes abscission because it may inhibit the transport of auxin from the leaf blade, which allow the action of ABA to promove the fall of flowers [46]. In general, environmental conditions can also stimulate to a decrease the auxin/ethylene relationship [45].

In our preliminary analysis, Santanter data, showed that the number of successful flowers pollinated to produce final yield can be affected negatively mostly by the rain, TMAX and wind. Nevertheless, a better field data tracing flowers development is essential to understand if it is a mechanical or physiological effect. In contrast, [47,48] indicated that the numbers of cocoa pollinators were reduced during the dry season, but increased in the wet season. This, could be due to midges need a moist environment to develop [42], which is difficult during the dry season as cocoa leaves create a dried ground mat [42,47]. Moreover, the lack of water during dry seasons may reduce the nutrients uptake provoking the massive flower drops [49]. For future studies, wind could be included in the SIMPLEcocoa model as an input to simulated this mechanical effects over number of flowers. In general, field data regarding counting flowers pollinated by month should be better reported for the region here studied.

4.2. Thermal Time for Harvest day predictions

We define the thermal time required to harvest cocoa pods for five Colombian regions as maturation of the fruit is related with temperature during the growth cycle [18]. Previous studies have calculated thermal time for different cocoa cultivar in Brazil and Ghana [30]. They also confirm that the fruit maturation time decrease in with an increase in temperature as was presented on others researches [30,50,51]. the effects of temperature and solar radiation on fruit growth and development was previously studied by [30], showing that crop under higher temperatures thought the crop cycle induce greater fruit losses because of physiological maturation (cherelle wilt). When the fruit is mature, seeds are able to germinate inside the pod before the harvest day [18].

Our results showed that the hotter regions such as Apartado and Arauca presented higher thermal time values (Fig. 5 and 3). Even though, Arauca had very low values of SAR but very high temperatures, this may be caused by clouds cover. The opposite can be seen for Caldas and Santander. These, extreme relations T/SAR can compensate the crop efficiency, for example in Santander (Fig. 6). The thermal time calculation was defined base on 180 DAF because farmers cut the pods by calendar days. Previous studies stated that evaluating the level of knowledge of growers regarding cocoa crop management, showed that the harvest was in the group of activities that presented the lowest level of information by the farmers [10]. That is why, these results present important temperature boundaries to predict fruit maturation day. Therefore, may be other environmental factors that should be studied for further research.

4.3. Cocoa Crop Model Simulations

Although cocoa is a relevant crop and there is an extensive agronomic literature, there is only one physiological crop model specific for cocoa so far. The (SUCROS-Cocoa) developed by [1]. However, the code was not easy available for adaptations. In contrast, the SIMPLE model has an open code in R, which we could adapt such a model would be very useful to compare yields and predict harvest date in different climates. As the harvest day was predicted form the flowering date (Fig. 7, consequently, biomass production and fSolar presented a crop cycle shorter than 180 DAF (Fig. 6, a and b). These simulations are coincident with results presented by [18], where physiological maturity of coca pod varies from 140 to 162 DAF. Our results showed how the harvest day can vary depending on the accumulate temperature during each especific crop cycle simulated (Fig. 7).

Biomass simulations use SIMPLEcocoa model presented similar predicted values (10000 Kg ha^{-1}) for coca drops in Costa Rica using SUCROScocoa [1]. Biomass simulations are a common evaluation in crop modells such as Sirius [33], SUBSTOR-potato [52] and DSSAT, CropSyst, STICS and WOFOST [53]. The approach of this research was focus on the harvest date prediction, hence the leaves crop cycle was evaluated indirectly this study. The fraction of intercepted photosynthesis active radiation (fSolar) decreased (Fig. 6, b) when the senescence of the canopy [1]. In cocoa canopy senecence refereed to a group of leaves responsible at the moment of the fruit formation to produce carbohydrates. These leaves eventually drop becoming on litter over the soil. Leaves life cycle has been simulated using crop models [1,54], which can be the reference to improve our SIMPLEcocoa crop model in future studies.

Yield prediction presented a RRMSE values 7.2%, which were significant lower than those presented for other crops using the SIMPLE model which reported and RRMSE of 24.4% [6]. Resulting in reliable approach for cocoa yield prediction. In general, these results may help to improve the quality of cocoa seed considering the moment to harvest can be variable depending on weather changes.

4.4. App development and future challenges

The original code of SIMPLE model of [6] was modified to make easy the implementation of this cocoa model as and app to be used in smartphones and desktops by farmers in Colombia. Therefore, the new version for cocoa crops simulation will be used to predict yield, date of harvest and biomass production, inserting only the date of flowering and region. The app development is on charge of Grupo BIOS to be deliver to farmers in Caldas initially at the end of 2021.

5. Conclusions

This research presented and initial crop calibration that can be improved with further studies, including effects over the pod production by diseases, nutritional diffidences and abiotic stresses. The future challenge will be that traditional farmers start to harvest more aware of the environmental effects over their crops. It will be necessary that they engage growers with adapt founding from scientific studies. Moreover, It will be required the help of entrepreneurs, researchers, academics and non-specialized communities to transfers the knowledge to cocoa growers.

Cocoa fruit development for harvest in the right time depend on whether conditions and principles of crop physiology and flower phenology. This was common for the five regions. Thermal time characterisation range from 1200 to 3000 days °C, with a T_b of 10 °C for the fruit development. The SIMPLEcocoa model allowed to predict the harvest date with better precision than only considering days by calendar. Thus, the crop cycle of cocoa for harvest should be shorter than 180 days after flowering. These results confirm the potential of the Crop Simulation Model approaches for tropical crop in Latin America.

6. Author contributions

ARV and ARC model calibration and data analysis. PA and ODR field data collection. AMG, ARC, and ARV editing and results interpretations. ARV were primarily responsible for writing the manuscript.

Acknowledgments: The authors wish to thank to Grupo BIOS and Fedecacao for facilitated the data from cocoa fields.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zuidema, P.A.; Leffelaar, P.A.; Gerritsma, W.; Mommer, L.; Anten, N.P. A physiological production model for cocoa (Theobroma cacao): model presentation, validation and application. *Agricultural Systems* **2005**, *84*, 195–225. doi:10.1016/j.agsy.2004.06.015.
2. Motamayor, J.C.; Risterucci, A.M.; Lopez, P.A.; Ortiz, C.F.; Moreno, A.; Lanaud, C. Cacao domestication I: the origin of the cacao cultivated by the Mayas. *Heredity* **2002**, *89*, 380–386.
3. Argout, X.; Salse, J.; Aury, J.M.; Guiltinan, M.J.; Droc, G.; Gouzy, J.; Allegre, M.; Chaparro, C.; Legavre, T.; Maximova, S.N.; Abrouk, M.; Murat, F.; Fouet, O.; Poulain, J.; Ruiz, M.; Roguet, Y.; Rodier-Goud, M.; Barbosa-Neto, J.F.; Sabot, F.; Kudrna, D.; Ammiraju, J.S.S.; Schuster, S.C.; Carlson, J.E.; Sallet, E.; Schiex, T.; Dievart, A.; Kramer, M.; Gelley, L.; Shi, Z.; Bérard, A.; Viot, C.; Boccara, M.; Risterucci, A.M.; Guignon, V.; Sabau, X.; Axtell, M.J.; Ma, Z.; Zhang, Y.; Brown, S.; Bourge, M.; Golser, W.; Song, X.; Clement, D.; Rivallan, R.; Tahi, M.; Akaza, J.M.; Pitollat, B.; Gramacho, K.; D'Hont, A.; Brunel, D.; Infante, D.; Kebe, I.; Costet, P.; Wing, R.; McCombie, W.R.; Guiderdoni, E.; Quetier, F.; Panaud, O.; Wincker, P.; Bocs, S.; Lanaud, C. The genome of Theobroma cacao. *Nature genetics* **2011**, *43*, 101–108. doi:https://doi.org/10.1038/ng.736.
4. Rodriguez-Medina, C.; Arana, A.C.; Sounigo, O.; Argout, X.; Alvarado, G.A.; Yockteng, R. Cacao breeding in Colombia, past, present and future. *Breeding Science* **2019**, *69*, 373–382. doi:10.1270/jsbbs.19011.
5. Nix, A.; Paull, C.; Colgrave, M. Flavonoid Profile of the Cotton Plant, Gossypium hirsutum: A Review. *Plants* **2017**, *6*. doi:10.3390/plants6040043.

6. Zhao, C.; Liu, B.; Xiao, L.; Hoogenboom, G.; Boote, K.J.; Kassie, B.T.; Pavan, W.; Shelia, V.; Kim, K.S.; Hernandez-Ochoa, I.M.; Wallach, D.; Porter, C.H.; Stockle, C.O.; Zhu, Y.; Asseng, S. A SIMPLE crop model. *European Journal of Agronomy* **2019**, *104*, 97–106. doi:https://doi.org/10.1016/j.eja.2019.01.009.
7. Niemenak, N.; Cilas, C.; Rohsius, C.; Bleiholder, H.; Meier, U.; Lieberei, R. Phenological growth stages of cacao plants (Theobroma sp.): codification and description according to the BBCH scale. *Annals of Applied Biology* **2010**, *156*, 13–24, [https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1744-7348.2009.00356.x]. doi:https://doi.org/10.1111/j.1744-7348.2009.00356.x.
8. Suárez, Y.Y.J.; Castañeda, G.A.A.; Daza, E.Y.B.; Estrada, G.A.R.; Molina, J.R. Modelo productivo para el cultivo de cacao (Theobroma cacao L.) en el departamento de Santander. *Corporación Colombiana de Investigación Agropecuaria - AGROSAVIA* **2021**. doi:https://doi.org/10.21930/agrosavia.model.74046 47.
9. Abbott, P Cand Benjamin, T.; Burniske, G.; Croft, M.; Fenton, M.; Kelly, C.R.; Lundy, M.; Rodriguez-Camayo, F.; Wilcox, M.J. An analysis of the supply chain of cacao in Colombia. *United States Agency for International Development: Washington, DC* **2019**.
10. Gutiérrez García, G.A.; Gutiérrez-Montes, I.; Hernández Núñez, H.E.; Suárez Salazar, J.C.; Casanoves, F. Relevance of local knowledge in decision-making and rural innovation: A methodological proposal for leveraging participation of Colombian cocoa producers. *Journal of Rural Studies* **2020**, *75*, 119–124. doi:https://doi.org/10.1016/j.jrurstud.2020.01.012.
11. García-Cáceres, R.G.; Perdomo, A.; Ortiz, O.; Beltrán, P.; Lopéz, K. Characterization of the supply and value chains of Colombian cocoa. *DYNA* **2014**, *81*, 30 – 40. doi:https://doi.org/10.15446/dyna.v81n186.39555.
12. Escobar, S.; Santander, M.; Useche, P.; Contreras, C.; Rodríguez, J. Aligning strategic objectives with research and development activities in a soft commodity sector: a technological plan for colombian cocoa producers. *Agriculture* **2020**, *10*. doi:10.3390/agriculture10050141.
13. Lamos-Díaz, H.; Puentes-Garzón, D.E.; Zarate-Caicedo, D.A. Comparison Between Machine Learning Models for Yield Forecast in Cocoa Crops in Santander, Colombia. *Revista Facultad de Ingeniería* **2020**, *29*. doi:https://doi.org/10.19053/01211129.v29.n54.2020.10853.
14. Escobar, S.; Santander, M.; Zuluaga, M.; Chacón, I.; Rodríguez, J.; Vaillant, F. Fine cocoa beans production: Tracking aroma precursors through a comprehensive analysis of flavor attributes formation. *Food Chemistry* **2021**, *365*, 130627. doi:https://doi.org/10.1016/j.foodchem.2021.130627.
15. Reynolds, M.; Kropff, M.; Crossa, J.; Koo, J.; Kruseman, G.; Molero Milan, A.; Rutkoski, J.; Schulthess, U.; Balwinder-Singh.; Sonder, K.; Tonnang, H.; Vadez, V. Role of modelling in international crop research: overview and some case studies. *Agronomy* **2018**, *8*. doi:doi:10.3390/agronomy8120291.
16. Bai, T.c.; Tao, W.; Zhang, N.n.; Chen, Y.q.; Mercatoris, B. Growth simulation and yield prediction for perennial jujube fruit tree by integrating age into the WOFOST model. *Journal of Integrative Agriculture* **2020**, *19*, 721–734. https://doi.org/10.1016/S2095-3119(19)62753-X, doi:https://doi.org/10.1016/S2095-3119(19)62753-X.
17. Romero-Vergel, A. Crop Modelling and Remote Sensing for Yield Predictions of Asparagus Cultivated in Peru. *Ongoing publication* **2021**.
18. López-Hernández, J.; López-Hernández, L.; Avendaño-Arrazate, C.; Aguirre-Medina, J.; Espinosa-Zagaroza, S.; Moreno-Martínez, J.; Mendoza-López, A.; Suárez-Venero, G. Biología floral de cacao (Theobroma cacao L.); criollo, trinitario y forastero en México. *AGROProductividad* **2018**, *11*, 129–136.
19. Swanson, J.D. Flower development in Theobroma cacao L.: an assessment of morphological and molecular conservation of floral development between Arabidopsis thaliana and Theobroma cacao. PhD thesis, The Pennsylvania State University, 2005.
20. Cheesman, E.E. Fertilization and Embryogeny in Theobroma Cacao, L. *Annals of Botany* **1927**, *41*, 107–126.
21. Berry, D.; Cilas, C. Étude génétique de la réaction à la pourriture brune des cabosses chez des cacaoyers (Theobroma cacao L) issus d'un plan de croisements diallèle. *Agronomie* **1994**, *14*, 599–609.
22. Lehrian, D.W.; Keeney, P.G. Changes in lipid components of seeds during growth and ripening of cacao fruit. *Journal of the American Oil Chemists' Society* **1980**, *57*, 61–65, [https://aocs.onlinelibrary.wiley.com/doi/pdf/10.1007/BF02674361]. doi:https://doi.org/10.1007/BF02674361.
23. Toledo-Hernández, M.; Tscharntke, T.; Tjoa, A.; Anshary, A.; Cyio, B.; Wanger, T.C. Landscape and farm-level management for conservation of potential pollinators in Indonesian cocoa agroforests. *Biological Conservation* **2021**, *257*, 109106. doi:https://doi.org/10.1016/j.biocon.2021.109106.
24. NASA. Nasa Power. *The National Aeronautics and Space Administration* **2021**. accessed on 10 July 2021.
25. RStudio Team. *RStudio 1.4: Integrated Development Environment for R*. RStudio, PBC., Boston, MA, 2020.
26. Meza-Sepúlveda, D.C.; Castro, A.M.; Zamora, A.; Arboleda, J.W.; Gallego, A.M.; Camargo-Rodríguez, A.V. Bio-Based Value Chains Potential in the Management of Cacao Pod Waste in Colombia, a Case Study. *Agronomy* **2021**, *11*.
27. Fedecacao. Producción nacional de cacao. *Federación Nacional de Cacaoteros* **2021**.
28. Ritchie, J.T.; Nesmith, D.S., Temperature and Crop Development. In *Modeling Plant and Soil Systems*; John Wiley and Sons, Ltd, 1991; chapter 2, pp. 5–29, [https://acsess.onlinelibrary.wiley.com/doi/pdf/10.2134/agronmonogr31.c2]. doi:https://doi.org/10.2134/agronmonogr31.c2.
29. Slafer, G.A.; Rawson, H.M. Base and optimum temperatures vary with genotype and stage of development in wheat. *Plant, Cell & Environment* **1995**, *18*, 671–679, [https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-3040.1995.tb00568.x]. doi:https://doi.org/10.1111/j.1365-3040.1995.tb00568.x.

30. Daymond, A.; Hadley, P. Differential effects of temperature on fruit development and bean quality of contrasting genotypes of cacao (Theobroma cacao). *Annals of Applied Biology* **2008**, *153*, 175–185.
31. Lahive, F.; Hadley, P.; Daymond, A.J. The physiological responses of cacao to the environment and the implications for climate change resilience. A review. *Agronomy for Sustainable Development* **2019**, *39*, 1–22.
32. Erneholm, I.; others. Cacao production of South America. Historical development and present geographical distribution. *Cacao production of South America. Historical development and present geographical distribution.* **1948**.
33. Crout, N.; Craigon, J.; Cox, G.; Jao, Y.; Tarsitano, D.; Wood, A.; Semenov, M. An objective approach to model reduction: Application to the Sirius wheat model. *Agricultural and Forest Meteorology* **2014**, *189-190*, 211–219. doi:https://doi.org/10.1016/j.agrformet.2014.01.010.
34. Agele, S.; Famuwagun, B.; Ogunleye, A. Effects of shade on microclimate, canopy characteristics and light integrals in dry season field-grown cocoa (Theobroma cacao L.) seedlings. *Journal of Horticultural Sciences* **2016**, *11*, 47–56.
35. Fuentes, L.F.Q.; Castelblanco, S.G.; Jerez, A.G.; Guerrero, N.M. Caracterización de tres índices de cosecha de cacao de los clones CCN51, ICS60 e ICS 95, en la montaña santandereana, Colombia. *Revista de Investigación Agraria y Ambiental* **2015**, *6*, 252–265.
36. Amir, J.; Sinclair, T. A model of water limitation on spring wheat growth and yield. *Field Crops Research* **1991**, *28*, 59–69. doi:https://doi.org/10.1016/0378-4290(91)90074-6.
37. Stöckle, C.O.; Donatelli, M.; Nelson, R. CropSyst, a cropping systems simulation model. *European Journal of Agronomy* **2003**, *18*, 289–307. Modelling Cropping Systems: Science, Software and Applications, doi:https://doi.org/10.1016/S1161-0301(02)00109-0.
38. Steduto, P.; Hsiao, T.C.; Raes, D.; Fereres, E. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. *Agronomy Journal* **2009**, *101*, 426–437. doi:https://doi.org/10.2134/agronj2008.0139s.
39. Saunders, L.G. Methods for studying forcipomyia midges, with special reference to cacao-pollinating species (Diptera, ceratopogonidae). *Canadian Journal of Zoology* **1959**, *37*, 33–51, [https://doi.org/10.1139/z59-005]. doi:10.1139/z59-005.
40. Kaufmann, T. Studies on the ecology and biology of a cocoa pollinator, *Forcipomyia squamipennis* L. (Diptera, Ceratopogonidae), in Ghana. *Bulletin of Entomological Research* **1975**, *65*, 263–268. doi:10.1017/S0007485300005940.
41. Sotomayor Chávez, A.C. Evaluación del riesgo ambiental de la aplicación de cuatro pesticidas sobre el polinizador Forcipomyia spp.(Díptera: Ceratopogonidae) del cultivo de cacao (Theobroma cacao L.), en San Martín, Perú. *Ingenieria Ambiental* **2020**. doi:https://hdl.handle.net/20.500.12805/1483.
42. Frimpong-Anin, K.; Adjaloo, M.K.; Kwapong, P.K.; Oduro, W. Structure and stability of cocoa flowers and their response to pollination. *Journal of Botany* **2014**.
43. Young, A.M. Seasonal differences in abundance and distribution of cocoa-pollinating midges in relation to flowering and fruit set between shaded and sunny habitats of the La Lola Cocoa Farm in Costa Rica. *Journal of Applied Ecology* **1983**, pp. 801–831.
44. Young, A.M. *Population biology of tropical insects*; Springer Science & Business Media, 2012.
45. Aneja, M.; Gianfagna, T.; Ng, E. The roles of abscisic acid and ethylene in the abscission and senescence of cocoa flowers. *Plant Growth Regulation* **1999**, *27*, 149–155.
46. Beyer Jr, E.M. Abscission: the initial effect of ethylene is in the leaf blade. *Plant Physiology* **1975**, *55*, 322–327.
47. Frimpong, E.; Gordon, I.; Kwapong, P.; Gemmill-Herren, B. Dynamics of cocoa pollination: tools and applications for surveying and monitoring cocoa pollinators. *International Journal of Tropical Insect Science* **2009**, *29*, 62–69.
48. Frimpong, E.A.; Gemmill-Herren, B.; Gordon, I.; Kwapong, P.K. Dynamics of insect pollinators as influenced by cocoa production systems in Ghana. *Journal of Pollination Ecology* **2011**, *5*.
49. Vaughton, G.; Ramsey, M. Pollinators and seed production. In *Seed development and germination*; Routledge, 2017; pp. 475–490.
50. Alvim, P.d.T.; Machado, A.; Vello, F. Physiological responses of cacao to environmental factors. *Rev. Theobroma* **1974**, *4*, 3–25.
51. End, M.J. A Study of the Effects of the Photo-Thermal Environment on Fruit and Seed Growth and Development in Theobroma cacao L. *The University of Reading, Reading,Uk.* **1991**.
52. Raymundo, R.; Asseng, S.; Prassad, R.; Kleinwechter, U.; Concha, J.; Condori, B.; Bowen, W.; Wolf, J.; Olesen, J.E.; Dong, Q.; Zotarelli, L.; Gastelo, M.; Alva, A.; Travasso, M.; Quiroz, R.; Arora, V.; Graham, W.; Porter, C. Performance of the SUBSTOR-potato model across contrasting growing conditions. *Field Crops Research* **2017**, *202*, 57–76. Modeling crops from genotype to phenotype in a changing climate, doi:https://doi.org/10.1016/j.fcr.2016.04.012.
53. Confalonieri, R.; Orlando, F.; Paleari, L.; Stella, T.; Gilardelli, C.; Movedi, E.; Pagani, V.; Cappelli, G.; Vertemara, A.; Alberti, L.; Alberti, P.; Atanassiu, S.; Bonaiti, M.; Cappelletti, G.; Ceruti, M.; Confalonieri, A.; Corgatelli, G.; Corti, P.; Dell'Oro, M.; Ghidoni, A.; Lamarta, A.; Maghini, A.; Mambretti, M.; Manchia, A.; Massoni, G.; Mutti, P.; Pariani, S.; Pasini, D.; Pesenti, A.; Pizzamiglio, G.; Ravasio, A.; Rea, A.; Santorsola, D.; Serafini, G.; Slavazza, M.; Acutis, M. Uncertainty in crop model predictions: What is the role of users? *Environmental Modelling and Software* **2016**, *81*, 165–173. doi:https://doi.org/10.1016/j.envsoft.2016.04.009.
54. Karunaratne, A.; Azam-Ali, S.; Al-Shareef, I.; Sesay, A.; Jørgensen, S.; Crout, N. Modelling the canopy development of bambara groundnut. *Agricultural and Forest Meteorology* **2010**, *150*, 1007–1015. doi:https://doi.org/10.1016/j.agrformet.2010.03.006.