Article

# A modelling strategy to improve cacao quality and productivity

**Angela RV [1,†,‡], Adriana [1,‡] and Anyela VCR[1,*]**

[1] The John Bingham Laboratory, NIAB, 93 Lawrence Weaver Road, Cambridge CB3 0LE, UK 1; e-mail@e-mail.com
[2] BIOS 2; e-mail@e-mail.com
* Correspondence:anyela.camargorodriguez@niab.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials) +xx-xxxx-xxx-xxxx (F.L.)
† Current address: Affiliation 3
‡ These authors contributed equally to this work.

**Abstract:** Crop modelling can support agronomical decisions of crop production under a range of scenarios improving competitiveness. Cocoa production systems in Latin America has a high importance over social and economic development, facing the fight against hunger and poverty. 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. We parametrized the SIMPLE crop model [1] to predict the best time for harvest cocoa fruits in Colombia. The aim of this study is develop a practical tool for farmers using flowering date and weather variables (Solar radiation, rain and temperature) easily available for crop modelling to understand and predict the phenology of cacao trees which is key to produce high quality cacao beans.

**Keywords:** ICS95; CCN51; thermal time, flowering date

**Citation:** Lastname, F.; Lastname, F.; Lastname, F. 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 [2–5]. Cacao plant member of the Malvaceae (formerly Sterculiaceae) botanical family such as cotton *Gossypium hirstium* [6] wich is modeled in SIMPLE model [1]. Cocoa is grown for its fruits, known as cacao pods. [7,8]. Only the 5% of the world cocoa yield is desalinated for Fine-cocoa production due to the low productivity of the traditional crop management [4]. 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 [5,9] as cocoa is grown by approximately 52.000 [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]. It is still not widely produced as the lack of adoptions of technologies by the traditional farmers. They empirically harvest after 5 or 6 months after flowering date, hence they ferment cocoa beans without considering the quality of the seeds at the harvest time. This produce heterogeneous characteristics between each fermentation batch diminishing the quality of cocoa final 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 [1,2,16,17]. 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 [2]. 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" [2]. 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 [2]. 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 [1].

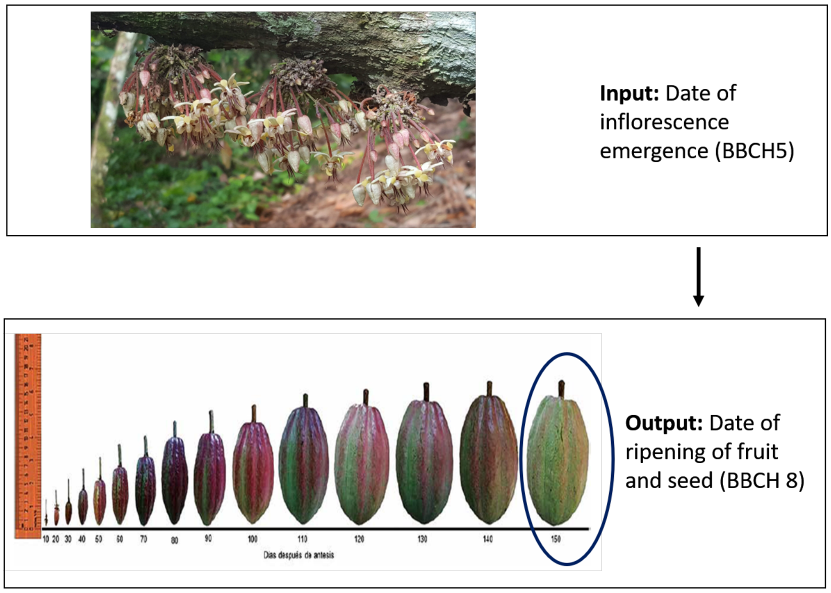
In this paper, we present a physiological parametrization of SIMPLE crop model for cocoa to predict the best harvest time and yield production. We used the SIMPLE crop model [1] for three reasons: 1: That it is very comprehensively described in the original paper. 2: That the code was available in R for initial trials and 3: That it had already been successfully fitted to perennial crops in south America. Overall, the model simulates crop development, growth and yield, and predict the maturation day when the fruit is ready to harvest. It includes 13 parameters ( daily weather data, irrigation, and soil and key dates) to specify a crop type, with four of these for cultivar characteristics easily available from farmer. Thus, this could be used as tool for small farmers with the aim to improve the quality of cocoa to become more competitive in Fine-cocoa market.

## 2. Materials and Methods

### *2.1. Floral Phenology of Cocoa*

In the SIMPLE model we simulate the floral phenology from the date of inflorescence emergence (BBCH scale 5) to predict the date of ripening of fruit and seed (BBCH 8) [7] (See figure 1). For modelling purposes the crop cycle of cocoa as perennial plant does not start at the plantation date such as annual crops where usually the planting date as start point of the crop cycle to model. The growth period of the fruit can varied from 110 to 150 daa (days after anthesis) [18] when cacao fruits reaches the physiological maturity, but if can be harvested at 170 days daa [7].

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




Credits: Taken from phys.org [19] and [18]

### 2.2. *Input data acquisition for SIMPLE crop model calibration*

Input variables required to run SIMPLE model for cocoa include the flowering date and daily weather of solar radiation (SRAD), maximum and minimum temperature (TMAX, TMIN) and rain. Weather data as csv file was downloaded from the POWER Data Access Viewer [20] from January 01 2018 to December 31 2020. The csv file had to be transformed to .WHT file using R 1.4 version [21].

### 2.3. *SIMPLE crop model calibration*

The procedure for the model calibration used was a sequential process of modifying or adding the appropriate files, changing the parameters in the simulation management file, running the program and inspecting the results. This cycle was then repeated until the simulated yields for cocoa were reasonably close to the observed yield.

A zip file of the R codes and input files for the SIMPLE model was kindly provided by Dr Chuang Zhao of the Agricultural and Biological Engineering Department, University of Florida. The main directory contains R codes, the *"Parameter definition"* Excel ® file, where the units required in each input file for the model are described, instructions and directories containing other input files. Examples of all input files are provided in the Zip files together with a range of data sets that can be used to check model operation after installation. Before running the SimpleB.R program initially line 72 was edited, with RStudio, to provide the full path for the installation. Also, the single year/single experiment mode was selected by setting the Grid-simulation switch in line 58 to "off" by entering [1] at the end of line 58. After checking the first column in the Input/Simulation Management.csv file is all zeros apart from 1 on line alongside cocoa experiment, then the model was run.

In the SIMPLE model dummy files are provided for adding new cocoa crop data and weather data, and files 2 to 6 in the list below can be edited to define new cultivars or experiments. Then modifying the simulation management file will cause the new files to be read when the program is run.

1. Input/Simulation Management.csv
2. Input/Species parameter.csv
3. Input/Cultivar.csv
4. Input/Treatment.csv

5. Input/Irrigation.csv
6. Input/Soil.csv
7. Observation/Obsdummy crop Exp name.csv
8. Weather/dummy weather.WTH

The first step was add the new experiment for cocoa in the dummy crop and experiment names were replaced with our crop and treatment names in the input files and flowering date was changed. Thermal time values required for Input/Cultivar.csv had to be defined for ICS 95 and CCN51 cultivars using field data from Fedecacao plot reports in Santander and Caldas, Colombia. Thus, the date for harvest is predicted when the model calculations reach the thermal time for each specific cultivar ICS 95 and CCN51. The new experiment name "KOKOlatl" for was used in the species parameter file with the respective values of parameters for cocoa including irrigation, soil features and radiation use efficiency [2] add table parameters values will change through the calibration.

Table 1: Cocoa crop parameter values used in Species-parameterfile.

| Tbase | Topt | RUE | I50maxH | I50maxW | MaxT | ExtremeT | CO2-RUE | S-water |
|---|---|---|---|---|---|---|---|---|
| 105 | 26 | 0.8 | 100 | 5 | 35 | 45 | 0.09 | 1.5 |

Table 2: Cocoa crop parameter values used in Cultivar file.

| Cultivar | Tsum | HI | I50A | I50B |
|---|---|---|---|---|
| ICS 95 | 3300 | 0.4 | 680 | 200 |
| CCN51 | 3600 | 0.4 | 680 | 200 |

Table 3: Cocoa crop parameter values used in Treatment and Observation file.

| Variable name | Value |
|---|---|
| Exp | KOKOlatl |
| Label | Colombia |
| Weather | KOKOS (.WTH file name) |
| CO2 | 400 |
| SowingDate | flowering date |
| SoilName | Loamy sand4 |
| Cultivar | ICS 95 or CCN51 |
| MaxIntercept | 0.95 |
| InitialBio | 1 |
| InitialTT | 0 |
| InitialFsolar | 0.01 |
| LAI | 2 |

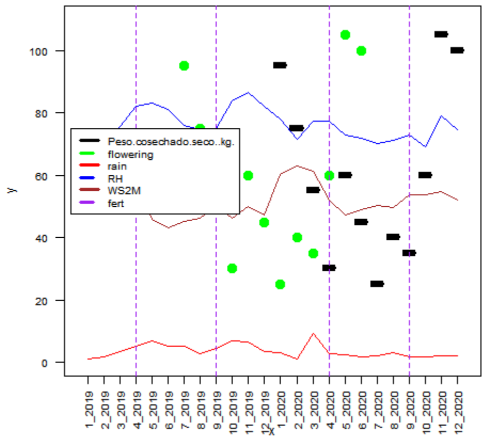
### *2.4. Evaluation of model performance*

The SIMPLE model performance was evaluated by comparing simulated values cocoa yield with those reported by Fedecacao from cocoa plantations, using statistical indices of relative-RMSE (RRMSE) [1] and coefficient of determination ($R^2$) [1,16,22]. The harvest date predicted was supported by Adriana results

## 3. Results

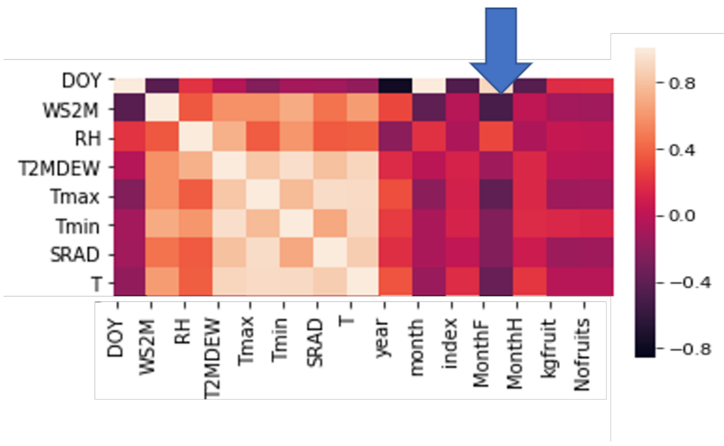
Studying the weather data effects over the flowering date and their final yield. We found that the number of successful flowers pollinated to produce final yield can be affected mostly by the rain, TMAX and wind (Figure 3). The later can affect the availability of tiny flies pollinators from Diptera order and from the families of of the biting midges *Ceratopogonidae*, genus *Forcipomyia* [23–25] to reach the cocoa flowers. However, wind is not included in the SIMPLE model as an input as cocoa crop has

**Figure 2.** SIMPLE model output for cocoa in Colombia




not been simulated with this model before. Flower opening is very well synchronised between the cohorts of mature flowers opening each night. The flowers open at almost exactly the same time and rate, irrespective of their position on the trunk. Unfertilised flowers abscise from the trunk approximately 1 day after flower opening [7]. .

**Figure 3.** Phenology of cocoa in Colombia for crop modelling




.

## 4. Discussion

Authors should discuss the results and how they can be interpreted from the perspective of previous studies and of the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted. this section on going to organize

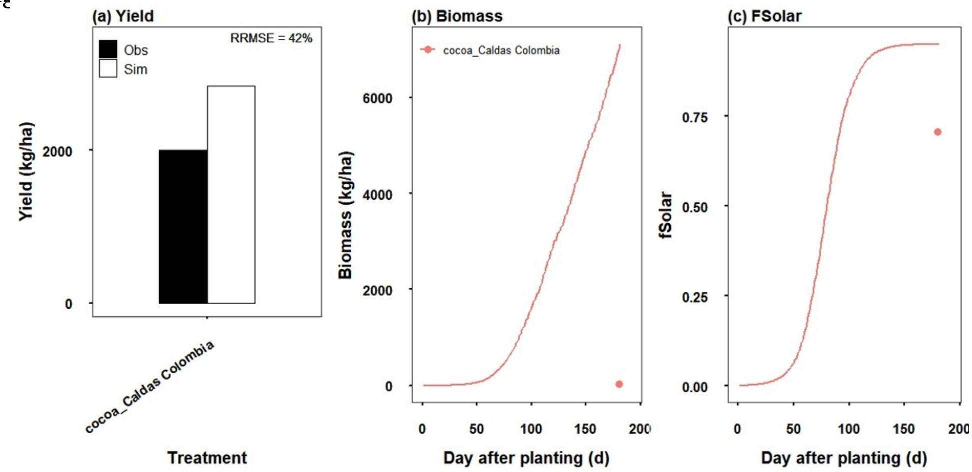
Evaluating the level of knowledge of producers regarding cocoa crop management, the harvest was in the group of activities that presented the lowest level of knowledge on the part of the producers according to the general averages [10].

## 5. Conclusions

This section is not mandatory, but can be added to the manuscript if the discussion is unusually long or complex.

**Author Contributions:** For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used

**Figure 4.** SIMPLE model output for cocoa in Colombia.




"Conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.", please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

**Funding:** Please add: "This research received no external funding" or "This research was funded by NAME OF FUNDER grant number XXX." and and "The APC was funded by XXX". Check carefully that the details given are accurate and use the standard spelling of funding agency names at https://search.crossref.org/funding, any errors may affect your future funding.

**Data Availability Statement:** In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Please refer to suggested Data Availability Statements in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics. You might choose to exclude this statement if the study did not report any data.

**Acknowledgments:** In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

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

## References

1. 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.
2. 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.
3. 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.
4. 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.
5. 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.

6. Nix, A.; Paull, C.; Colgrave, M. Flavonoid Profile of the Cotton Plant, Gossypium hirsutum: A Review. *Plants* **2017**, *6*. doi:10.3390/plants6040043.
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. 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.
20. NASA. Nasa Power. *The National Aeronautics and Space Administration* **2021**. accessed on 10 July 2021.
21. RStudio Team. *RStudio 1.4: Integrated Development Environment for R*. RStudio, PBC., Boston, MA, 2020.
22. Camargo Rodriguez, A.V.; Ober, E.S. AquaCropR: Crop Growth Model for R. *Agronomy* **2019**, *9*. doi:10.3390/agronomy9070378, doi:10.3390/agronomy9070378.
23. 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.
24. Kaufmann, T. Studies on the ecology and biology of a cocoa pollinator, Forcipomyia squamipennis I. + M. (Diptera, Ceratopogonidae), in Ghana. *Bulletin of Entomological Research* **1975**, *65*, 263–268. doi:10.1017/S0007485300005940.
25. 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.