

Mathematical modeling of climatological data to estimate passion fruit crop yield (*Passiflora edulis Sims L. f. Flavicarpa y purpurea*)

Leila Nayibe Ramírez Castañeda¹, Gina Paola González Angarita²,
José-Alejandro Cleves-Leguizamo³

Abstract - Passion fruit crop yield depends on the behavior of climatic variables, and modeling the dependence relationship of these variables regarding crop yield offers information aimed at facilitating agribusiness decision making. As main aim, passion fruit crop yield was estimated using mathematical models. A multivariate and univariate statistical analysis of meteorological variables was carried out during the observation period between 2007 and 2014 of selected weather stations, identified and located in the Colombian middle tropics (County of Huila). The relationship between yield with the following agroclimatic variables were analyzed: temperature, sunlight, relative humidity, rainfall and ENSO at monthly resolution with empirical and mechanistic models, recommended in scientific literature. Results showed that the multiple regression model requires the highest yield peaks; the adjustment of the multiple regression model is low, while univariate models such as the ARIMA model showed better adjustment in the time series analyzed. The Stewart's water-yield model has better performance to estimate yield as a function of evapotranspiration in the different phenological phases.

Index terms: yield, crop yield modeling, uncertainty, passion fruit.

Modelagem matemática de dados climatológicos para estimar a produção do maracujazeiro-amarelo (*Passiflora edulis Sims L. F. Flavicarpa*)

Resumo - A produtividade da cultura do maracujá depende do comportamento das variáveis climáticas, e a modelagem da relação de dependência destas variáveis com a produtividade da cultura oferece informações que visam a facilitar a tomada de decisão do agronegócio. Como objetivo principal, a produtividade da cultura do maracujá foi estimada por meio de modelos matemáticos. Uma análise estatística multivariada e univariada das variáveis meteorológicas foi realizada durante o período de observação entre 2007 e 2014 das estações climatológicas selecionadas, identificadas e localizadas no trópico médio colombiano (Região administrativa de Huila). Foram analisadas as relações de produtividade com as seguintes variáveis agroclimáticas: temperatura, brilho solar, umidade relativa, precipitação e ENOS em resolução mensal com modelos empíricos e mecânicos, recomendados na literatura científica. Os resultados mostraram que o modelo de regressão múltipla requer os maiores picos de rendimento; o ajuste do modelo é baixo, enquanto modelos univariados, como o modelo ARIMA, apresentaram melhor ajuste nas séries temporais analisadas. O modelo de rendimento de água de Stewart tem melhor desempenho para estimar rendimento em função da evapotranspiração nas diferentes fases fenológicas.

Termos para indexação: produtividade, modelagem de rendimento de cultura, incerteza, maracujá.

Corresponding author:

leylan.ramirez@unilibre.edu.co

Received: October 06, 2020

Accepted: February 22, 2021

Copyright: All the contents of this journal, except where otherwise noted, is licensed under a Creative Commons Attribution License.



¹Doctorate student in Engineering, Full-Time Research Professor, Universidad Libre, Bogotá, Colombia. E-mail: leylan.ramirez@unilibre.edu.co (ORCID 0000-0002-0651-0971)

²PhD in Geography, Director of the Environmental Engineering Program, Universidad Libre, Bogotá, Colombia. E-mail: ginap.gonzalez@unilibre.edu.co (ORCID 0000-0001-9210-3653)

³PhD in Agroecology, Full-Time Research Professor, Universidad Pedagógica y Tecnológica de Colombia, Escuela de Administración de Empresas Agropecuarias, Duitama, Boyacá, Colombia. E-mail: jose.cleves@uptc.edu.co (ORCID 0000-0001-9717-9753)

Introduction

Climate plays an important role in the development of crops, and the three most important climate components are light, temperature and rainfall (CAMPOS ARANDA., 2005). In addition, relative humidity and the El Niño-Niña Southern Oscillation (ENSO) climatic pattern measured from the ONI index are also considered, because these extreme phenomena have an impact on crop yields, causing threats to food security (IDEAM, 2013).

Additionally, it has been shown that the microclimate has an impact on plant growth and development, which are linked to climate behavior on a larger scale (macroclimate), and under these conditions, it is possible to generalize the results of studies to describe the behavior of the relationships between vegetation and climate (CAMPOS ARANDA., 2005).

In the case of the production of fruit species such as *Passiflora edulis Sims L. f. Flavicarpa y purpurea*, it depends on multiple agroclimatic variables, such as altitude, latitude, temperature, relative humidity, radiation, rainfall, wind speed and sunlight, among others, which are conditions that humans cannot control, directly affecting crop performance. Management decisions depend on the farmer's knowledge (cultural or learned) regarding the climate behavior and its impact on crop yield (MORLEY-BUNKER, 1999).

The scientific-technical background of production forecasts is based on the knowledge of the relationship between ecophysiological requirements of the species and the environmental supply, referring in this particular case to the supply and intensity of meteorological factors, with fundamental support of statistics (KANTANANTHA; STEWART, 2007), so that they accurately fulfill the prediction and are useful for other specialists (DELGADILLO-RUIZ et al., 2016; RUÍZ-RAMÍREZ; HERNÁNDEZ-RODRÍGUEZ; ZULETA RODRÍGUEZ, 2011; MARTÍNEZ VENTURA, 2006).

In this regard, forecasting models that estimate production volumes in semi-permanent fruit species such as passion fruit contribute to its understanding. The models most widely used are simulation models (SOTO G.; COTES T.; RODRÍGUEZ C., 2017), statistical-environmental models, statistical-biometric models and statistical models by sampling (KANTANANTHA; STEWART, 2007).

For this reason, quantitative techniques are useful tools to minimize economic losses, strengthening the availability and supply of the product in the market. In addition, farmers reduce the uncertainty in their decisions and economic consequences due to the uncertain behavior of the climate system, which has strong impacts in the phytosanitary management, quality, production volume and finally prices (PANNELL; MALCOLM; KINGWELL, 2000; MOSCHINI; HENNESSY, 2001).

Other fields of application using forecasts to estimate crop yields are food security, crop insurance policy indemnities, import and export plans, as they are relevant decisions for planning state subsidies. These decisions are considered of national interest, therefore of strategic nature (KANTANANTHA; STEWART, 2007). At producer level, yield forecast before harvest time generates information to plan crops and considers strategies to minimize eventual losses (KANTANANTHA; STEWART, 2007).

For the purpose of this research, empirical and mechanistic modeling approaches widely reported in scientific literature were analyzed to estimate yield from agroclimatic variables, and the advantages of their application were analyzed through the selected case (KANTANANTHA; SERBAN; GRIFFIN, 2010; KANTANANTHA; STEWART, 2007; KUSUMASTUTI; DONK; TEUNTER, 2016; MUSSHOF; HIRSCHAUER, 2007). Forecasts aim to capture more closely the dynamics of each of cultural practices (transitory and permanent), taking into account the production cycle (MARTÍNEZ VENTURA, 2006).

Materials and methods

Selection of agroclimatic stations

Weather stations were selected from the catalog of stations of the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM, 2019), according to the following criteria: information completeness above 75%, greater number of series / year of climatic variables, which were main weather stations (CP) and finally the coverage of climatic stations taking as reference their minimum Euclidean distance to locations of observed production units.

The mathematical programming model is described below:

Sets

I : Producers

J : Weather stations

Parameters

d_{ij} : Euclidean distance from producer i to agroclimatic station j .

Decision variable

x_{ij} : Binary variable 1 if the weather station is assigned to producer i or 0 otherwise

$$\min z = \sum_{i=1}^n \sum_{j=1}^m d_{ij} x_{ij} \quad (1)$$

Subject to:

$$\sum_{i=1}^n x_{ij} \leq m \quad \forall j \in J \quad (2)$$

$$\sum_{j=1}^m x_{ij} = 1 \quad \forall i \in I \quad (3)$$

Where equation (1) is the objective function that minimizes the distance between two points, equation (2) allows assigning producers to the closest weather stations and equation (3) allows all producers to be assigned to one

weather station. This model was solved with the Branch and Bound method (LAWLER; WOOD, 1966) through the Open Solver software (OPENSOLVER.ORG, 2020). The selected stations are listed below (Table 1):

Table 1. Selected weather stations

CROP	CODE	ALTITUDE (msnm)	MUNICIPALITY	LATITUDE	LONGITUDE	BS	HR	PP	T
Passion fruit	21025030	1368	ALTAMIRA	2.079167	-75.736222	X	X	X	X
	21035040	1045	SUAZA	1.868389	-75.827778		X	X	X
	21055020	1070	LA PLATA	2.378278	-75.89125	X	X	X	X
	21105040	680	CAMPOALEGRE	2.696806	-75.298889		X	X	X

Source: IDEAM (2019) (BS: Sunlight, HR: Relative Humidity, PP: Rainfall, T: Temperature)

Within the location of the most important passion fruit producers, the municipality of La Plata (Huila) stands out, located in the central mountain range at coordinates 2°23'00" N and 75°56'00", average temperature of 23°C, and altitude of 1180 meters above sea level. Crop production has different technification degrees, and weather stations with the shortest distance to passion fruit producers are 21105040 and 21035040, selected in this research, based on the mathematical programming model.

Weather data

Historical data (average / month) of climatic variables temperature, relative humidity, sunlight and rainfall were collected. ONI indexes from 2000 to 2004 and from 2007 to 2014 were obtained with monthly resolution, and the database construction obeyed the time periods with greater availability of information.

Crop yield data

According to reports from the Ministry of Agriculture and Rural Development, yield data (ton * ha⁻¹) / month of passion fruit cultivation of municipalities where weather stations were located were collected, and the yield behavior according to the phenological stage of the crop was considered.

The first step consisted in the creation of the database in a time series structure (DALININA, 2017; DERRYBERRY, 2014; R, 2019; METCALFE; COWPERTWAIT, 2009), evaluating stationarity, seasonality and decomposition of crop yield series.

Descriptive exploratory analysis

The methodology proposed by different authors was followed (BOX et al., 2015; METCALFE; COWPERTWAIT, 2009; CLAYTON et al., 2002; SHUMWAY; STOFFER, 2017), since their techniques involve aspects of trend and seasonality, among others. Analyses were carried out using the version 5.0.0 of the

Python Programming language under the Jupyter interface (programming language for statistical processing and graph elaboration) and the Rstudio language version 1.2.1335 (available at <https://jupyter.org>, <https://www.rstudio.com/>).

Description of empirical mathematical models

To explain the yield behavior of *Passiflora edulis* Sims L. f. *Flavicarpa* and *purpurea*, empirical mathematical models based on performance forecasts were used, which are a statistical procedure with certain disadvantages since it must have significant volume of data during several periods of time to generate the model reliability. These simple and direct methods have been developed with variables such as rainfall, temperature, solar radiation and sometimes nutrients to estimate yield (VANDENDRIESSCHE; VAN ITTERSUM, 1995). The techniques reported with this approach are: Autoregressive Integrated Moving Average (ARIMA), Multiple Regression, robust regression models and Neural Networks. Each of the techniques that were developed are described below.

Multiple Regression

A regression model in which the dependent or response variable "Y" is related to more than one explanation variable is called multiple linear regression model (DERRYBERRY, 2014) (Eq. 4).

$$Y_i = \beta_0 + \sum \beta_p X_{pi} + \varepsilon_i \quad (4)$$

Where disturbance ε_i is the error associated with measurement i of the X_{pi} value, under the assumptions that $\varepsilon_i \sim N(0, \sigma^2)$. The analysis between two or more variables can be performed through equations as proposed with this statistical analysis method. Residues represent the variability that is not captured by the deterministic relationship represented by regression functions (HENGL, 2009; HENGL; HEUVELINK; ROSSITER, 2007).

Based on descriptive statistical analyses, the model of all climatic variables was proposed, applying the backward methodology of elimination of less significant variables proposed by Montgomery, Peck and Vining, (2002).

ARIMA model

The ARIMA model has been widely used in all scientific-technical fields to identify temporal stochastic processes and estimate models for their prediction and control (Box and Jenkins, 1976). In the ARIMA model (p, d, q), p is defined as the order of the autoregressive process, d is the number of differences that are necessary for a process to be stationary and q is the order of the moving average process, which can be represented as follows (Eq. 5).

$$Y_t = -(\Delta^d Y_t - Y_t) + \phi_0 + \sum_{i=1}^p \phi_i \Delta^d Y_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (5)$$

Where d is used to convert the original series into stationary, parameters $\phi_1, \dots, \phi_p, \dots$ are the autoregressive part, $\theta_1, \dots, \theta_q, \dots$ belong to the moving averages, the constant term and the stochastic disturbance. This methodology is based on the use of one-variable data to identify the characteristics of its underlying probabilistic structure, in contrast to traditional procedures used to identify models based on an explanatory theory of the phenomenon under study (DALININA, 2017).

Robust regression models

The TheilSen Regressor estimator uses the generalization of the median in multiple dimensions, being therefore a non-parametric statistical model robust to multivariate atypical values. This estimator is impartial for the real slope in simple linear regression (KUMAR SEN, 1968). For many distributions of the response error, the estimator has high asymptotic efficiency in relation to the least squares estimate (KUMAR SEN, 1968).

Gaussian process regression implements Gaussian processes for regression purposes, which is possible through optimization and random algorithms to improve the adjustment (KUMAR SEN, 1968).

Multilayer Perceptron (MLP)

The multilayer perceptron (MLP) trains iteratively since each time step is calculated through partial derivatives of the loss function with respect to parameters of the model to update it (KINGMA; BA, 2014; HE et al., 2015).

MLP is a supervised learning algorithm that learns, with the $f(\cdot): R^m \rightarrow R^o$ function, training a data set, where m is the number of input dimensions and o is the number of output dimensions. Given a set of $X = x_1, x_2, \dots, x_m$ characteristics and a y goal, one can learn through a non-linear regression approximation. It differs from

logistic regression since between the input layer and the output layer, there may be one or more non-linear layers, called hidden layers (SCIKIT-LEARN.ORG, 2019).

The input layer consists of a set of neurons that represent each input entity. Each neuron in the hidden layer transforms values from the previous layer into a weighted linear sum, followed by a nonlinear function, such as the hyperbolic tangent function. The output layer receives values from the last hidden layer and converts them into output values (SCIKIT-LEARN.ORG, 2019). The regression class used with the multilayer perceptron trains with backward propagation without activation function, the square error uses it as a loss function, and the output is a set of continuous values. The mathematical formulation is given by a set of training data $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, where $x_i \in R^n$ and $y_i \in \{0, 1\}$, where the hidden layer learns with the $f(x) = W_2 g(w_1^T x + b_1) + b_2$ function, where $w_1 \in R^m$ and $W_2, b_1, b_2 \in R$ are parameters of the model W_1, W_2 are the weights of input and hidden layers, respectively, and represent the trend added to the hidden layer and the output layer (SCIKIT-LEARN.ORG, 2019).

Description of mechanistic mathematical models

Mechanistic mathematical models assume that the crop growth system has known structure and can be mathematically described. The aims of these models allow integrating knowledge and testing hypotheses; however, the models have high predictive level since there are easily measurable variables (VANDENDRIESSCHE; VAN ITTERSUM, 1995). Within this category, recommended models are the water-yield models (ALLEN et al., 1998). The main advantage of the latter is the access to knowledge for non-experts, non-researchers and farmers.

Mechanistic mathematical models describe crop physiology but still require further development to improve prediction accuracy (VANDENDRIESSCHE; VAN ITTERSUM, 1995). The formal description of estimation models considered in this research will be given below.

Water-yield models

The water-yield response is a function defined by FAO (FAO, 2012), which model addresses this relationship through a simple Eq. 6 expressed as:

$$\left(1 - \frac{Y_a}{Y_x}\right) = K_y \left(1 - \frac{ET_a}{ET_x}\right) \quad (6)$$

Where Y_a and Y_x are actual and maximum yields, ET_a and ET_x are actual and maximum evapotranspiration and K_y is the yield response factor that represents the decrease in evapotranspiration over yield losses.

Other yield-water models are described as production functions that depend on the crop water requirements. The best known and most accepted are the Jensen and Stewart's model. Models are used by simulating the relationships of yield with the evapotranspiration deficit of various crop growth stages.

The Jensen's model was developed by Jensen (1968) (JENSEN, 1968), and is given by the following Eq. 7:

$$\frac{Y_a}{Y_{ck}} = \prod_{i=1}^n \left(\frac{ET_{ai}}{ET_{cki}} \right)^{\lambda_t} \quad (7)$$

Where Y_a is the yield of irrigation deficit treatments; Y_{ck} is the crop yield with full irrigation treatments; ET_{ai} is the current crop evapotranspiration of growth stage of irrigation deficit treatment; ET_{cki} is the maximum evapotranspiration in the growth stage i of total irrigation; λ_t is the Jensen index of crop yield sensitivity under water deficit; i is the growth stage; n is the number of stages.

The above Eq. 8 can be solved by obtaining

$$\lambda_t = \frac{\log \frac{Y_a}{Y_{ck}}}{\log \frac{ET_{ai}}{ET_{cki}}} \quad (8)$$

The Stewart's model developed by Stewart (1976) (STEWART; HAGAN; PRUITT, 1976) is given by (Eq. 9)

$$\frac{Y_a}{Y_{ck}} = 1 - \sum_{i=1}^n K_{yi} \left(1 - \frac{ET_{ai}}{ET_{cki}} \right) \quad (9)$$

Where K_{yi} is the water deficit sensitivity index of the crop yield, the other parameters have been previously defined.

Determination of the water deficit sensitivity indexes of models

Given the complexity and availability of information to establish the effects of irrigation in relation to yield response, regression models are used to estimate K_{yi} and λ_t constants and the Stewart and Jensen's models, respectively. For the evaluation of models, the weather station with the greatest amount of information and measured variables was used, which in this case was station 21055020. To calculate evapotranspiration, the Penman Monteith equation recommended by FAO (FAO, 2012) and accepted by the international scientific community in various research studies for its adjustment with respect to observed values was used. Evapotranspiration of the passion fruit crop was calculated from crop coefficients (K_c) for each phenological stage reported by Torrente (2009) (TORRENTE T., 2009) and De Lima (2004) (DE LIMA CORREA, 2004). The proposed methodology was based on recommendations by FAO (FAO, 2012) and by studies (CHEN et al., 2014; FAO, 2012) that facilitate the use of empirical techniques to achieve the approximation of constants.

To determine maximum yield, the study on the effect of irrigation and fertilization on passion fruit yield and quality was considered (DORADO G.; TAFUR H.; RIOS R., 2013). From the treatment of maximum irrigation and its response to yield without fertilization, the following polynomial regression model is obtained, which interprets this behavior, with adjusted R of 0.96 (Fig. 1).

From these studies, the following sensitivity indexes proposed by Stewart and Jensen are obtained for each of the stages and proposed models for each phenological stage analyzed for the passion fruit crop and the climatic conditions of the station identified with code 21055020.

Evaluation of the fit of models

To evaluate the goodness-of-fit of models, regression (b) and determination coefficients (R^2) were analyzed. In addition, estimation errors and modeling quality were also calculated through the Mean Absolute Deviation calculated as Eq. 10:

$$DAM = \frac{1}{n} \sum_{i=1}^n |S_i - M_i| \quad (10)$$

Where S_i is the relative value obtained from the yield model and M_i is the observed crop yield value.

Regarding ARIMA, Robust Regression and multilayer perceptron models, it is important to note that for the selection of the best model to predict the series under study, the Akaike information criteria (AIC) (CRYER; CHAN, 2008) was used, for which the lower the values of these measures, the better the model is in terms of relative quality in relation to the loss of information of the statistical model estimated for the crop yield series.

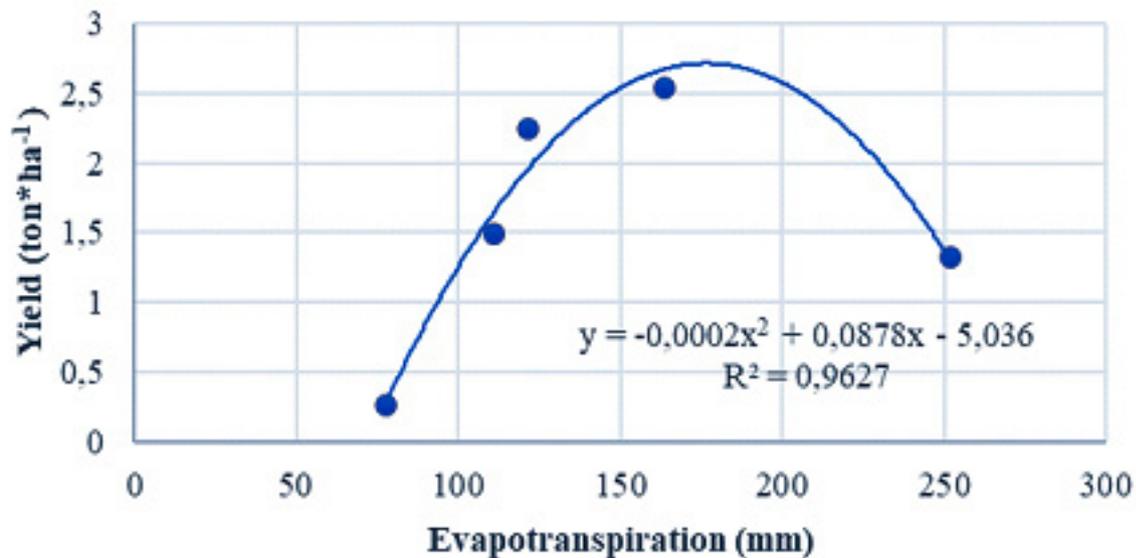
Results and discussion

Dispersion statistics for climatic and yield variables

The central tendency measures suggest normal distribution behavior (Table 2).

Table 2. Dispersion statistics of weather stations analyzed.

Station 21035040									
Crop	Variable	Minimum	Q1	Mean	Q3	Maximum	Standard deviation	Atypical Data	
								Lower limit	Upper limit
Passion fruit	Yield (ton*ha ⁻¹)	0.8058	1.14	1.37	1.59	2.238	0.3291	0.46	2.28
	Temperature (°C)	21.3	22.37	22.72	23	24.8	0.602	21.43	23.93
	Relative Humidity (%)	67	74	76.14	78	94	3.6646	68	84
	Rainfall (mm/month)	9.6	74.47	108.97	140.5	260.5	53.69	0	239.53
	ENSO (index)	-1.7	-0.8	-0.3093	0.125	1.6	0.7293	-2.18	1.51
Station 21105040									
Crop	Variable	Minimum	Q1	Mean	Q3	Maximum	Standard deviation	Atypical Data	
								Lower limit	Upper limit
Passion fruit	Yield (ton*ha ⁻¹)	0.46	0.79	0.99	1.14	1.66	0.27	0.61	2.32
	Temperature (°C)	24.4	25	26.12	26.1	24.4	3.6	23.35	27.75
	Relative Humidity (%)	54	67	71.16	75	83	6.76	55	87
	Rainfall (mm/month)	6.6	50.1	119.23	201.2	347.5	89.16	0	427.85
	ENSO (index)	-1.7	-1.1	-0.359	0.4	1.6	0.92	-3.35	2.65

**Figure 1.** Yield behavior regarding evapotranspiration of the passion fruit crop (*Passiflora edulis Sims L. f. Flavicarpa and purpurea*).

Among agroclimatic variables, rainfall has high variability, since its records range from 6.6 mm / month to 364.7 mm / month, a similar case was found for sunlight, which ranges from 54.2 hours / month to 218.9 hours / month, on average, and relative humidity does not exceed 94%. The analysis of atypical data under the methodology of interquartile ranges as a measure of statistical dispersion allows locating the extreme values of variables analyzed in the case of station 21035040, variable Relative Humidity had extreme values that correspond to 87%, 90, 5% and 94%, which occurred in the period from June to August 2008 respectively, based on historical information reported by IDEAM. The anomaly of the La Niña phenomenon was reported in the Andean region with strong intensity in periods between June 2007 and February 2008, and according to the source consulted, the occurrence of this phenomenon generated excessive rains in the middle and southern part of the Andean region, which significantly affected air temperature and rainfall (IDEAM, 2014). Another extreme value detected was rainfall of 260.5 mm / month in November 2013.

In station 21105040, relative humidity had atypical value of 54% in September 2009; in this period, El Niño phenomenon with weak intensity was reported by IDEAM, where there was a rainfall deficit condition (IDEAM, 2014), and according to observed data, rainfall in that period corresponded to 19.5 mm / month.

Passion fruit crop yield

In the seven years considered in station 21035040 near passion fruit producers, the average passion fruit crop yield was 1.38 ton * ha⁻¹ month, 75% of this variable reached maximum value of 1.59 ton * ha⁻¹ month, with maximum total value of 2.23 ton * ha⁻¹ month. Cyclical pattern of annual variation was observed, with higher production volume in the months of February and March and lower production volume in the months of July and August (Fig.2).

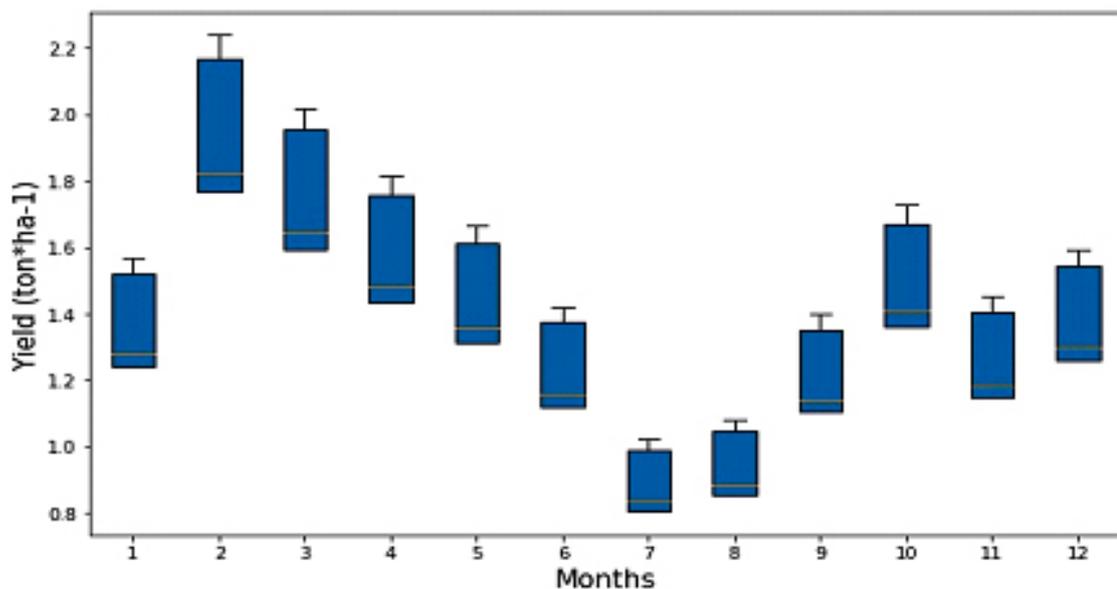


Figure 2. Multiple boxes representing passion fruit crop yield / month. Station 21035020.

In the cross-display diagrams for climatic variables, marked variations were observed in relation to crop yield. Less marked annual cyclical variations were detected for temperature, and some atypical values in some months for variables sunlight and rainfall.

On the other hand, the bivariate correlation matrix for station 21035040 shows less strong and positive association between yield and temperature (0.3426), as well as between ENSO and temperature (0.389) (Fig. 3).

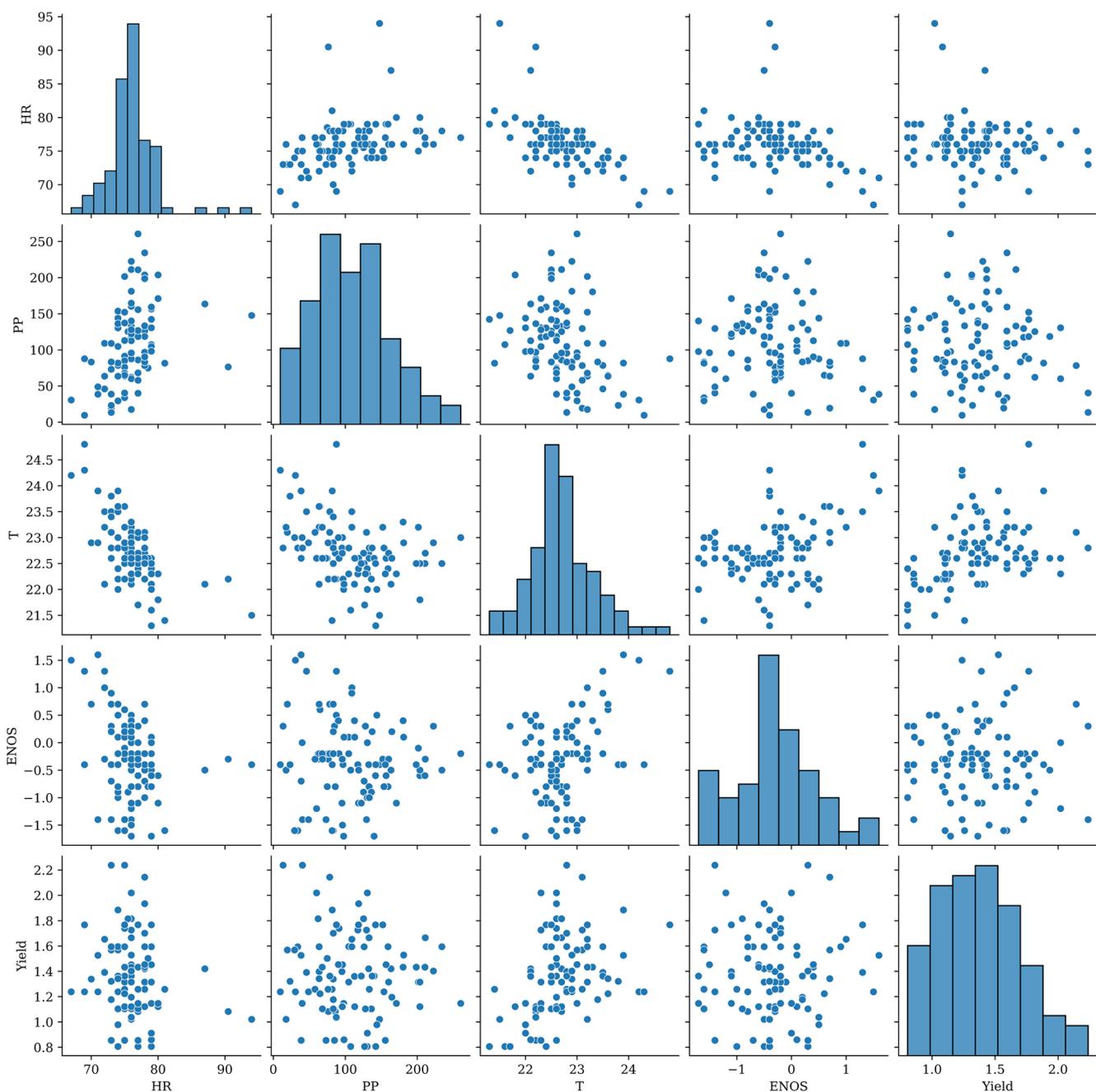


Figure 3. Cross correlation matrix between variables. Station 21035040

In station 21105040, the strongest linear association was found between variables rainfall and relative humidity (0.636), and strong negative association between relative humidity and temperature (-0.80), the other variables present less strong associations (0.36). The linear relationship is positive and less strong between yield and temperature for station 21035040 (0.34), while for station 21105040, yield has association with relative humidity (0.3) and rainfall (0.36) (Fig. 04).

Mathematical models to predict low yield with empirical approach

Multiple linear regression model

In this methodology, five models were tested to find the agroclimatic variables that best explained passion fruit crop yield. As previously explained, the statistical analysis did not show significant linear association between variables. In this sense, it was decided to analyze multiple regression models to explain yield for the selected stations (Table 3).

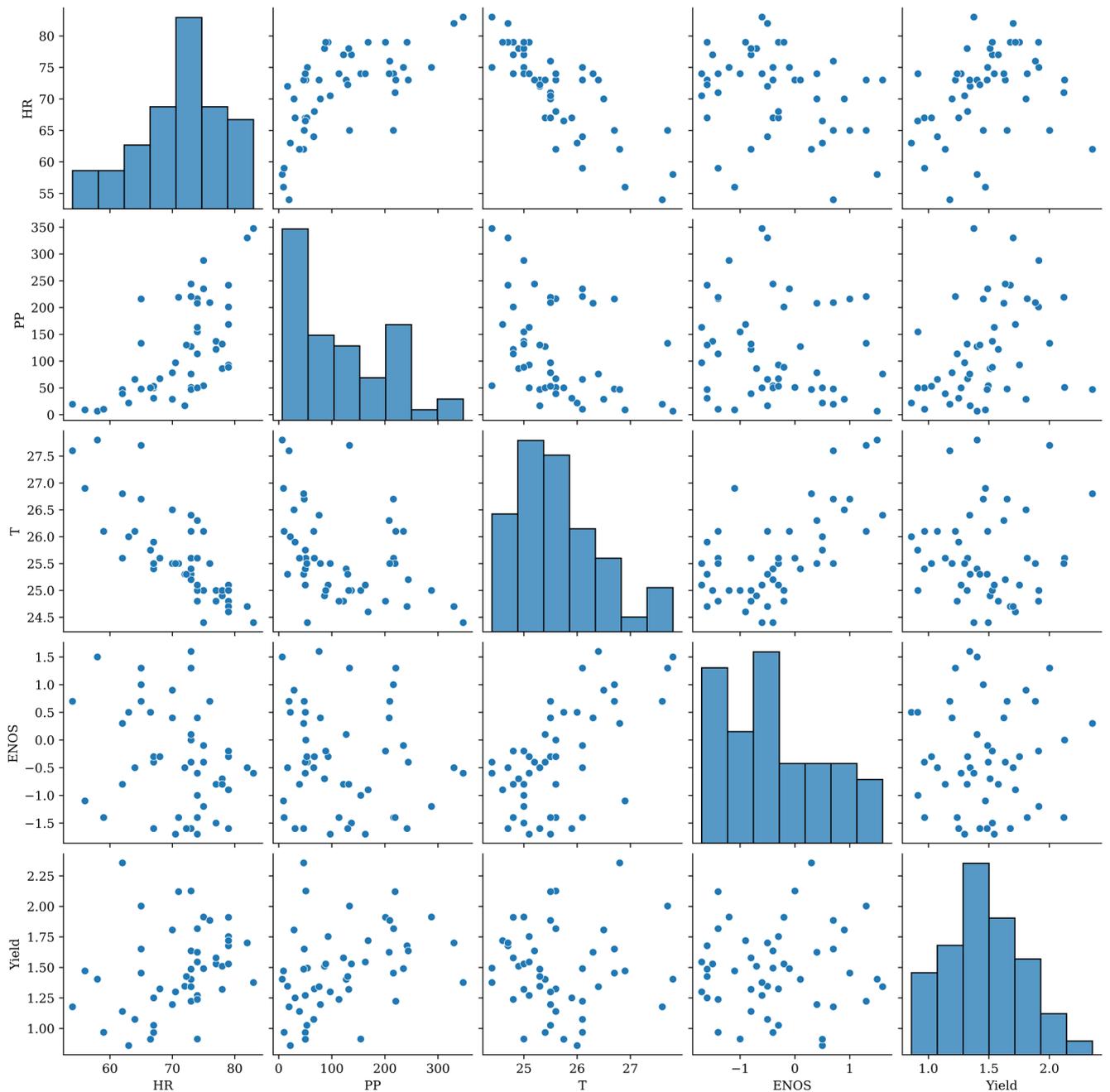


Figure 4. Cross correlation matrix between variables. Station 21105040

Table 3. Estimated multiple regression coefficients for the different weather stations analyzed.

Station	Variables	Coefficients	Pr(> t)	Indicators
21035040	Intercept	-5.04	0.026	R ² 0.141
	HR	0.0079	0.493	Adj R ² 0.103
	PP	0.0002	0.71	AIC 53.47
	T	0.2544	0.001	BIC 66.29
	ENSO	-0.0624	0.196	
21105040	Intercept	-12.34	0.003	R ² 0.327
	HR	0.0471	0.002	Adj R ² 0.266
	PP	0.0004	0.501	AIC 22.63
	T	0.4052	0.001	BIC 32.09
	ENSO	-0.1055	0.108	

In the first multivariate regression model, variable temperature, p-value significantly lower than 0.05 was observed (0.001); additionally, the probability value of the F statistic indicates that the model is significant at 5 %, with p-value (0.007) lower than 0.05. Low yield variability explained through the model was also observed, evidenced by R^2 (0.141) and adjusted R^2 (0.103) (CRYER; CHAN, 2008). In the case of the multivariate regression model of station 21105040, it was observed that p-value lower than 0.05 (0.002) corresponds to relative humidity, (0.001), temperature and low adjustment value with R^2 (0.327) and adjusted R^2 (0.266).

Although the adjustment of models is low, the possibility of estimating descriptive models is suggested; for practical purposes of the research, it was decided to continue with the analysis of residuals as validation of assumptions that give statistical support to estimates and predictions of the model (Fig. 5).

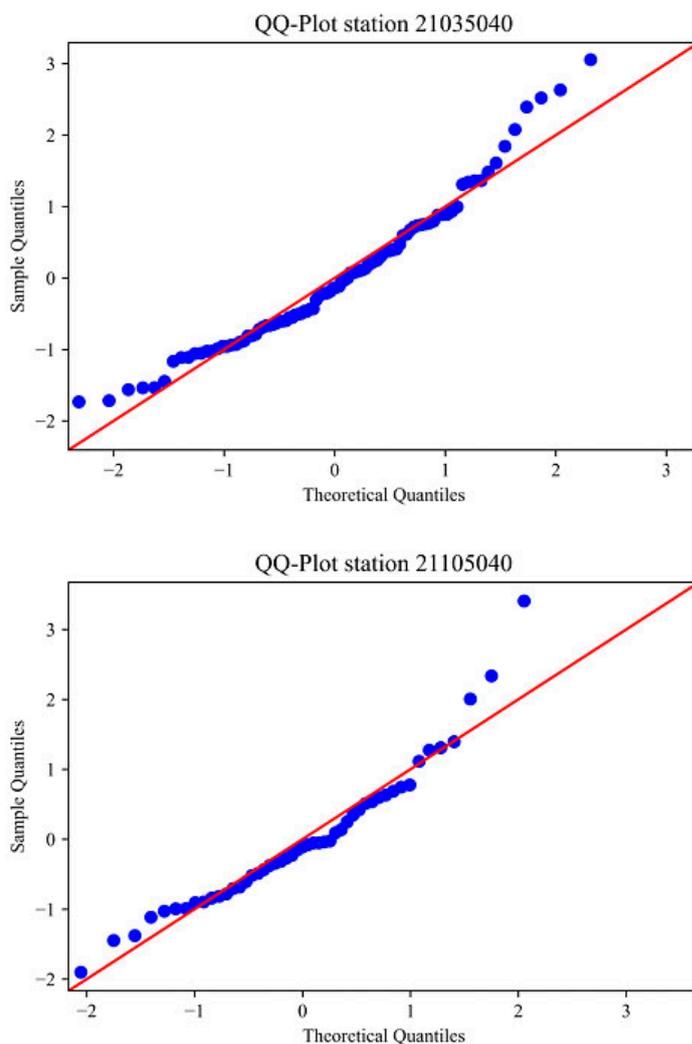


Figure 5. Analysis of residuals

In Figure 5, the normal probability graph for each of stations 21035040 and 21105040, a line is observed in the studentized residuals, demonstrating the non-normality of errors.

The Shapiro Wilks test is applied in this first analysis, as a normality test and to test the hypotheses H_0 : Residuals have normal behavior, H_a : Residuals do not present normality for $\alpha = 0.005$ with confidence level of 95 %, the p-value of the test for calculated residuals of station 21035040 is 0.003632, which is lower than α , confirming the rejection of the null hypothesis and the non-normality of residuals.

According to the above, to the multiple regression model for station 21035040, the significant variable that best explains yield is temperature. To improve the estimation of the model, parameters β are once again determined, where the independent variable is temperature and the dependent variable is yield. The expression obtained was:

$$\text{Yield} = 0.1872 \text{ Temperature} - 2.8767 \quad (11)$$

To verify the good specification, the Shapiro Wilks test is applied, and the p-value result of 0.007516 once again confirms the violation of the normality assumption of residuals.

To determine the good specification of the model, ANOVA was applied and F test was implemented, the hypotheses to be tested are H_0 : Poor model specification and H_a : Good model specification. In this case, p-value $0.0006318 < \alpha = 0.005$ was obtained, which indicates rejecting the H_0 hypothesis, which suggests that the model is biased and inconsistent. Under these criteria, it is necessary to apply data transformation to improve normality assumptions and model specification.

The suggested method for the transformation is Box-Cox (KUTNER et al., 2005), as it allows correcting biases in the distribution of errors and unequal variances, and to carry out the transformation, it is necessary to select a range of λ values, searching for the transformation to approximate data. When estimating the probability function for lambda, it is obtained through the Box-Cox transformation function, implemented in R.

In this case, the maximum probability function of lambda of 0.3838 suggests the transformation of natural logarithms for data. When applying the transformation and estimating the Shapiro Wilks normality test, p-value 0.5354 was obtained, with reliability of 95%, and the hypothesis of normality of residuals is accepted and the homoscedasticity assumption is verified.

To confirm the homoscedasticity assumption, the Breusch Pagan test was applied to confirm the equality of variances. P-value of 0.228, which accepts the null hypothesis, was obtained. Finally, low correlations were found between the explanatory variables of proposed models, which is why the assumption of non-multicollinearity between regression variables was validated.

For the independence assumption, the Durbin Watson test was applied to test the hypothesis of autocorrelation equal to zero and for the alternative non-zero, the p-value is $6.306e-13$, which determines that the alternative hypothesis was accepted. This result is due to the nature of data, where serial autocorrelation is shown.

The multiple linear regression models estimated for station 21105040 take previously indicated expressions, considering only variables that better explain yield (Eq. 12).

$$\text{Yield} = 0.0438 \text{ Relative Humidity} + 0.2932 \text{ Temperature} - 9.1566 \quad (12).$$

The adjustment of models is low, which reveals that climatic variables alone do not fully explain yield ($\text{ton} * \text{ha}^{-1} \text{ month}$), and other variables such as agrological soil class, fertility level, soil management, management and availability of water resources, etc., should also be considered. Additionally, each of the tests explained above were applied to each of the multiple regression models to verify assumptions, finding results and conclusions similar to those obtained with station 21035040.

ARIMA model

Figure 6 shows yield associated with climatic variables for stations 21035040, 21105040, stationarity, seasonality and decomposition of series of the yield variable ($\text{ton} * \text{ha}^{-1} \text{ month}$) in the analyzed time horizon. A pattern with annual periodicity was observed, with higher peaks in the first quarter of the year and lower peaks in the third quarter of the year with slight trend from 2011, which continues until 2014, where maximum yield of the series exceeds $2.23 \text{ ton} * \text{ha}^{-1} \text{ month}$ of passion fruit. The series was then decomposed into its seasonal, trend and residual parts.

The passion fruit yield series is stationary, a result corroborated by the Dickey Fuller test, with which the hypothesis is rejected at 5% significance (Table 04).

Table 4. Dickey Fuller statistic to test the seasonality of series.

Dickey-Fuller Statistic = - 6.463
Lag order = 4
P-Value = 0.01

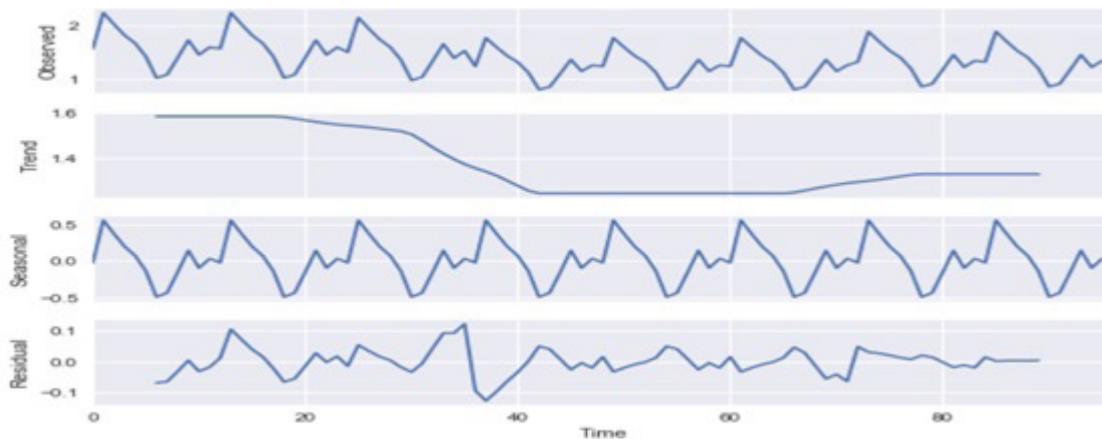


Figure 6. Decomposition of the passion fruit yield time series (*Passiflora edulis Sim L. f. Flavicarpa and purpurea*)

Using the auto.arima function of the forecast library, the best ARIMA model for the analyzed performance variable is (Table 05):

Table 5. ARIMA Models for Passion Fruit

ARIMA(0,1,0)
AIC=-341.78

When obtaining the residual graphs of the passion fruit yield autoregressive model ARIMA (0,1,0) and ARIMA (12,1,0), there is a clear pattern present in ACF / PACF and the residual graphs of the model repeats at lag 12. This suggests that the models may be better with a different specification, such as $p = 12$ or $q = 12$.

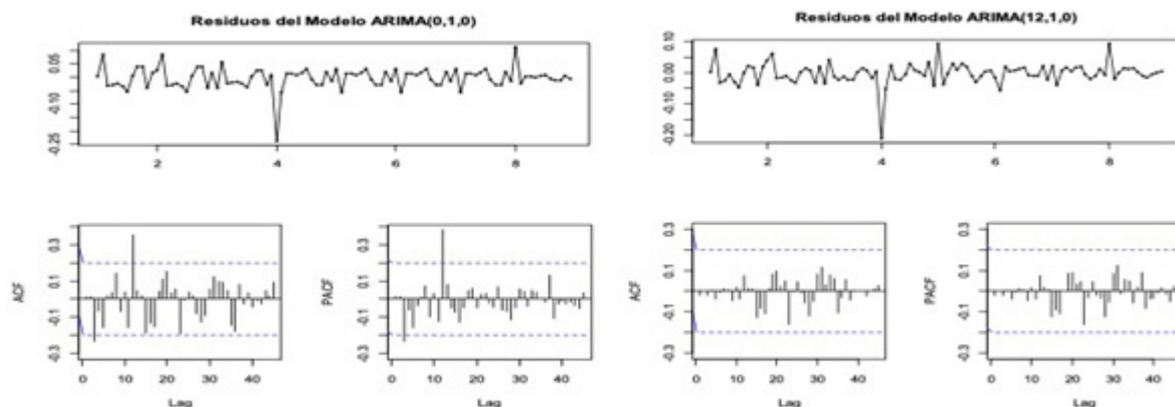


Figure 7. Residuals for the ARIMA (0, 1, 0) and ARIMA (12, 1, 0) model of passion fruit yield (*Passiflora edulis Sims L. f. Flavicarpa and purpurea*).

The graph of residuals of the second model (Fig. 7) correct distortions of the original model, and the AIC is lower with respect to ARIMA (0, 1, 0), which is why this model was selected as the best model.

Finally, the ARIMA (12, 1, 0) model that obtained lower AIC = -347.8 is expressed with the following Eq. 13:

$$\begin{aligned}
 Y_t = & 0.0884Y_{t-1} - 0.02934Y_{t-2} - 0.1922Y_{t-3} - 0.0724Y_{t-4} - 0.1348Y_{t-5} - \\
 & 0.0420Y_{t-6} + 0.0711Y_{t-7} + 0.0891Y_{t-8} + 0.0225Y_{t-9} + 0.0359Y_{t-10} - \\
 & 0.1284Y_{t-11} + 0.4522Y_{t-12} \quad (13)
 \end{aligned}$$

Robust regression models

Robust regression models were defined considering yield as a function of explanatory variable temperature.

Using the Anaconda jupyter package, the following results were obtained (Table 6).

Table 6. Adjustment values for robust regression models

Technique	21035040				21105040			
	VD	VI	R ²	MSE	VD	VI	R ²	MSE
Theil Sen regression	Yield	T	-112.97	12.22	Yield	T	-5.69	0.754
Regression Gaussian Processes	Yield	T	-17.68	2.00	Yield	T	-19.29	2.280

From these results, it was observed that the adjustment of models is minimal, which shows that these techniques are not adequate for data under analysis.

Multilayer Perceptron (MLP)

Given the low adjustment of the multiple regression and the results of the correlation matrix between variables, it was suggested that these relationships may be nonlinear. This is why the multilayer perceptron is considered a nonlinear multiple regression method. Input neurons are defined through training data considered for the different agroclimatic stations analyzed, which variables are relative humidity, temperature, sunlight, rainfall and the ENSO index.

Using the sklearn.neural_network import MLPRegressor (SCIKIT-LEARN.ORG, 2019) python function, MLP with 3 hidden layers was estimated with neurons equal to input variables where the activation function is $f(x) = \max(0, x)$. The solution method is an optimizer of the quasi Newton family of methods, and the number of iterations for the MLP estimation is 1500. The yield results obtained for each of the selected stations are shown in Table 7.

Table 7. Comparison of forecasts for passion fruit crop yield (*Passiflora edulis Sims L. f. Flavicarpa and purpurea*).

Passion Fruit Yield Station 21035040	MLP Forecast	Passion Fruit Yield Station 21105040	MLP Forecast
1.43314023	0.79780071	1.6509231	1.14620891
1.31485592	0.84468479	2.35607436	1.31609017
1.12100004	1.2214109	2.12598271	1.16243562
0.80585048	0.75954478	1.91085364	2.48672649
0.85456356	1.06203469	1.75314122	1.69345305
1.10338232	0.52045073	1.49466672	1.2669856
1.36292007	0.79937955	1.0744673	1.27769041
1.14657253	0.85864893	1.13941807	1.37414176
1.25797974	1.08403344	1.47117643	1.22978414
1.23819233	0.54353454	1.81722676	1.4826291
1.76705577	1.09270819	1.52876337	1.62689978
1.59448703	0.95293264	1.67730632	2.02684065
1.43314023	1.37955492	1.48583079	1.5963763
1.31485592	1.01198073	2.12046692	1.40880192
1.12100004	1.19241073	1.91338444	1.13741794
MAD	0.36709809	MAD	0.42875147

According to results, multilayer perceptron models have average mean absolute deviation of 0.36, which indicates that the accuracy and adjustment are very low given the deviation of the forecast with respect to yield observed in each of the selected stations.

Mathematical Models to Predict Low Yield with Mechanistic Approach

K_{yi} and λ_t estimated for the Stewart and Jensen models are observed in the following figure in the pre-flowering stage, using K_c (DE LIMA CORREA, 2004; TORRENTE, 2009) reported in literature for passion fruit crop (Fig 8).

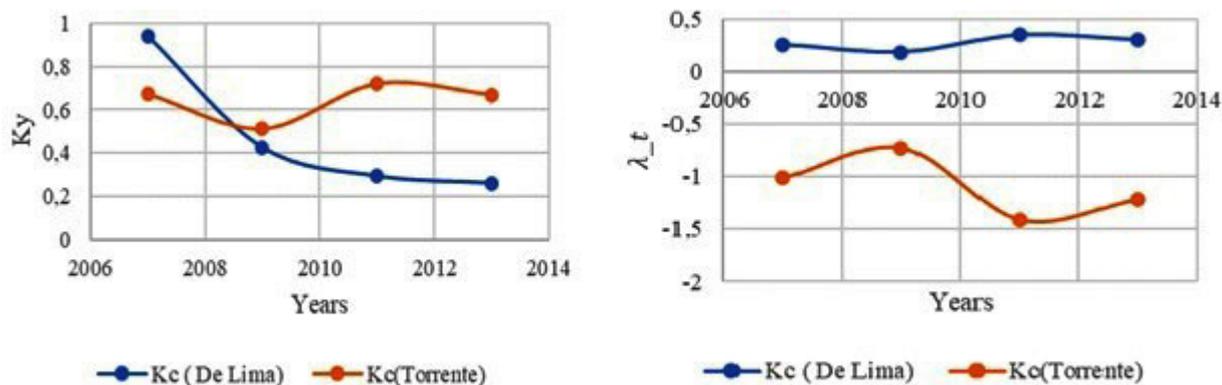


Figure 8. Sensitivity indexes for the Jensen and Stewart models in the pre-flowering stage (2007-2014)

It was observed that the sensitivity indexes associated with the effect of irrigation on yield are different for each of the years throughout the analysis period, according to K_c result obtained by De Lima, it shows that $K_{yi} < 1$, which means that the crop is more tolerant to water deficit and partially recovers from stress, showing proportionally smaller yield reductions due to the decrease in water use (FAO, 2012). In general, for passion fruit cultivation, the behavior is associated with $K_{yi} < 1$ for K_c proposed by De Lima and Torrente.

For the flowering stage, Figure 9 shows the calculated values of sensitivity indexes of water-yield models.

Regarding results obtained considering K_c proposed by De Lima, the sensitivity index K_{yi} is lower than 1, while in the case of K_c proposed by Torrente, K_{yi} is greater than 1, only for the year 2007. In general, as in the pre-flowering stage, passion fruit crop is tolerant to water deficit during the flowering stage.

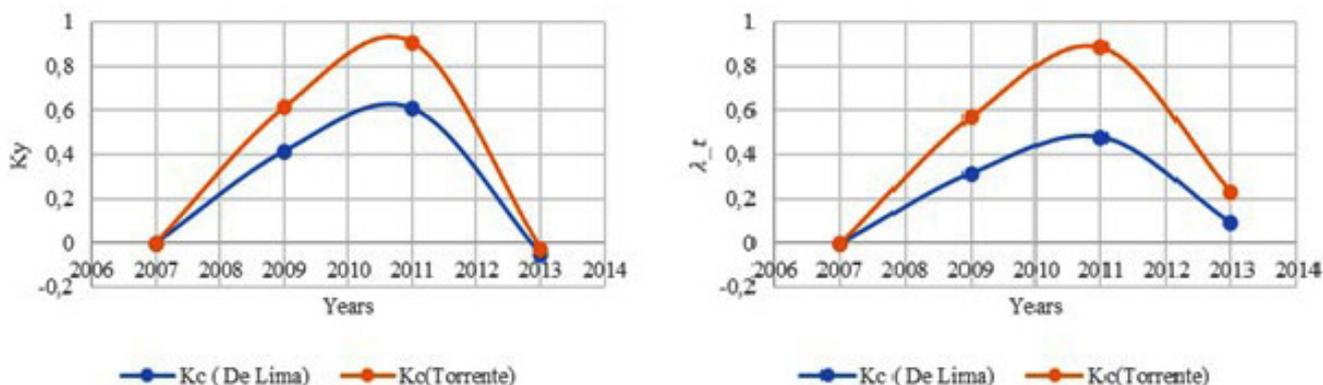


Figure 9. Sensitivity indexes for the Stewart and Jensen model in the flowering stage (2007 - 2014)

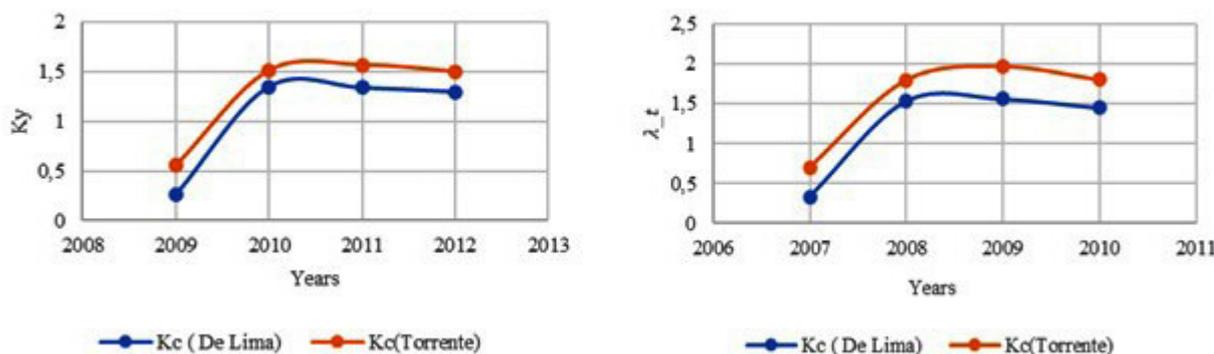


Figure 10. Water sensitivity index in the harvest stage (2009 -2012)

Finally, during the harvest stage (Fig. 10), $K_{y_i} > 1$, which indicates that they K_c interpret the crop yield conditions as a function of the water needs, confirming that the crop is tolerant to water stress during all phenological

stages. Sensitivity index constants for the phenological phases of passion fruit crop calculated for the Stewart and Jensen equations under the mechanistic approach are described below (Table 8):

Table 8. Sensitivity indexes for the Stewart and Jensen models.

Stage	K_c (De Lima)		K_c (Torrente)	
	K_y Stewart	λ_t Jensen	K_y Stewart	λ_t Jensen
Pre-flowering	0.48096728	0.27548971	0.64389582	-1.0901561
Flowering	0.24107874	0.22017158	0.37195351	0.42013473
Harvest	1.05949943	1.21841269	1.28406811	1.56592423

Comparison of empirical and mechanistic models

Empirical models such as multiple linear regression describe yield through explanatory variables, and ARIMA through the history of autoregressive components of series, moving averages and trend smoothing. According to results and due to the low adjustment of regression forecast models, their use for descriptive purposes is recommended. In the same comparative analysis, it could be observed that in the initial years, the model tends to underestimate peaks of higher yield, since it is not able to take into account temporal variations, which is one of the advantages of the ARIMA model over multiple linear regression (Fig. 11).

Regarding mechanistic models, the best yield was obtained for the Stewart model, using the K_c value of proposed by De Lima with Mean Absolute Deviation (MAD) of 0.39; deviations of the Jensen Model and the Stewart Model for K_c of Torrente, they have higher Mean Absolute Deviation compared to the other models. Furthermore, the multilayer perceptron as a technique to predict nonlinear associations obtained MAD of 0.367. Finally, the order according to the minimization of errors is recommended within the empirical models: ARIMA,

multilayer Perceptron, while within mechanistic models, the Stewart water-yield model using K_c recommended by De Lima.

The climatic conditions affect the crop ecophysiology, and by understanding and managing these variables, the environmental offer in association with the cultural management of farmers can be optimized (FISCHER; MELGAREJO, 2017).

In tropical conditions, the incidence of temperature more significantly explains crop yield in accordance with other investigations (CLEVES et al., 2016; MAYORGA et al., 2020). It was evidenced that crop evapotranspiration significantly influences aspects such as crop quality and yield (FISCHER; MELGAREJO; CUTLER, 2018).

This article contributes to technical scientific knowledge based on empirical models constructed from data, initially considering multivariate models, to determine the relationship between independent variables temperature, relative humidity, rainfall, sunlight, the ONI index and dependent variable yield. According to literature, new analysis variables are incorporated such as the ONI index, sunlight and relative humidity.

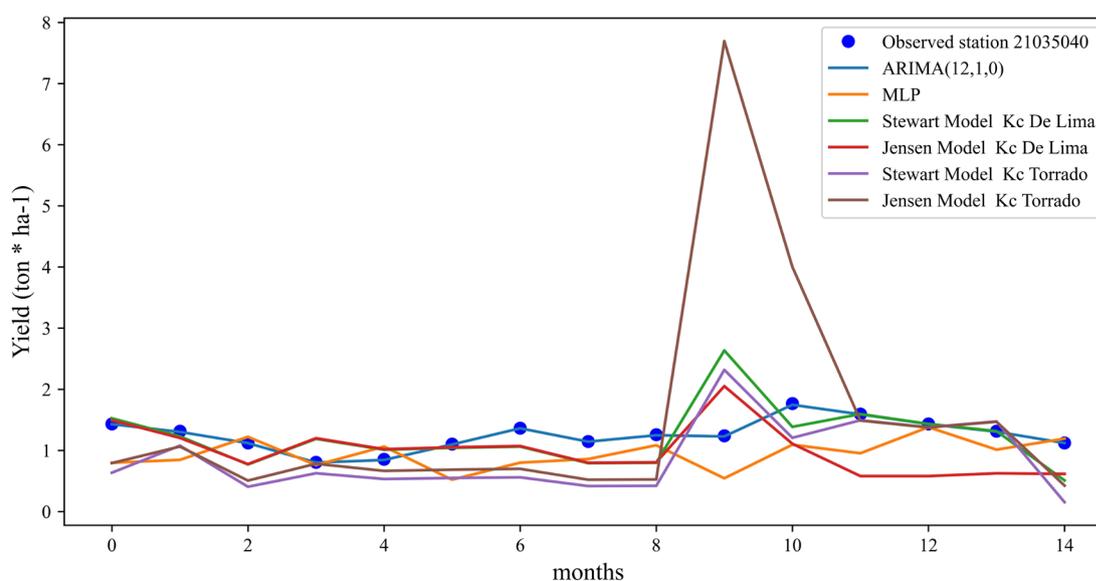


Figure 11. Original series and yield forecast of passion fruit crop (Ton * ha⁻¹) using empirical and mechanistic techniques.

The document makes significant contributions to the discrimination of the different methodologies reported in literature for yield forecasting, such as mechanistic and empirical mathematical models. Among mechanistic models, it was possible to determine the water sensitivity index and the recommended crop coefficient for passion fruit and regarding empirical ones, machine learning models such as robust regression and multilayer perceptron were tested, and this knowledge allows establishing the best methodology to predict crop yield.

In future studies, it is suggested to estimate specific models and develop all the necessary error and coefficient tests for each variable in order to establish that in these models, there is greater adjustment coefficients to improve their stability and provide a valid basis on which it can be predicted, so that the results of forecasts are considered acceptable and close to future reality.

Conclusions

Some alternatives to be considered in the future are the analysis of ARIMA models with the inclusion of explanatory variables, the use of Multiple Linear Regression models with lagged data and the modeling of Multiple Series, including in this case other series corresponding to some variables (evaporation, transpiration, evapotranspiration etc.).

Missing data of the main climatological variables limit the application of models, and the disarticulation among public entities that administer data is evident, which affects the quality of results.

The use of methodologies and models that evaluate the effect of agroclimatic factors on yield are a fundamental tool to be used in decision making by farmers, contributing to reduce the likelihood of uncertainties.

Inferential statistics for the construction of models allow the deduction and estimation of the phenomena from available information. The modeling of data patterns allows the researcher to draw inferences and confirms hypotheses about the characteristics obtained from future observations. The hypotheses established in this article such as the relationship between yield and climatic variables could not be confirmed through the analyzed information, given that correlations that coexist between variables can be nonlinear in nature and the methods applied to find this correlation are obviously based on linear relationships.

Univariate models such as ARIMA better explain yield behavior since MAD of 0.007 was obtained, while the water-yield model has better performance and explains the yield behavior as a function of evapotranspiration related to temperature, rainfall and wind speed. The absolute deviation of errors was 0.309 using Kc proposed by De Lima. The multivariate regression model is a tool that allows describing the behavior of the climatic variable system, and its predictive use is not recommended due to its low performance. The multilayer perceptron had performance measured through MAD of 0.367.

References

- ALLEN, R.G.; PEREIRA, L.S.; RAES, D.; SMITH, M. **Evapotranspiración del cultivo. Guías para la determinación de los requerimientos de agua de los cultivos**. Roma, 1998. Disponible em: <http://www.fao.org/3/a-x0490s.pdf>. Acceso em: 22 jul. 2019.
- BOX, G.E.P.; JENKINS, G.M.; REINSEL, G.C. **Time series analysis: forecasting and control**. 5.ed. New York: Wiley, 2015.
- CAMPOSARANDA, D.F. **Agroclimatología cuantitativa de cultivos**. México: Trillas, 2005.
- CHEN, J.; KANG, S.; DU, T.; GUO, P.; QIU, R.; CHEN, R.; GU, F. Modeling relations of tomato yield and fruit quality with water deficit at different growth stages under greenhouse condition. **Agricultural Water Management**, New York, v.146, p.131–148, 2014.
- CLAYTON, T.C.; ROBERSON, J.A.; ELGER, D.F.; MONTGOMERY, D.C.; PECK, E.A.; VINING, G. **Introducción al análisis de regresión lineal**. Ciudad de Mexico: Compañía Editorial Continental, 2002.
- CLEVES LEGUIZAMO, J.A.; TORO, J.C.; MARTÍNEZ BERNAL, L.F. Los balances hídricos agrícolas en modelos de simulación agroclimáticos. Una revisión analítica. **Revista Colombiana de Ciencias Hortícolas**, Tunja, v.10, n.1, p.149–163, 2016.
- CRYER, J.D.; CHAN, K.-S. **Time series analysis**. New York: Springer, 2008.
- DALININA, R. **Introduction to Forecasting with ARIMA in R**. 2017. Disponible em: <https://blogs.oracle.com/ai-and-datascience/post/introduction-to-forecasting-with-arima-in-r>
- DE LIMA CORREA, R.A. **Evapotranspiração e coeficiente de cultura em dois ciclos de produção do maracujazeiro amarelo**. 2004. Dissertação (Mestre em Agronomia) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2004.
- DELGADILLO-RUIZ, O.; RAMÍREZ-MORENO, P.P.; LEOS-RODRÍGUEZ, R.A.; GONZÁLEZ, J.M.S.; VALDEZ-CEPEDA, R.D. Pronósticos y series de tiempo de rendimientos de granos básicos en México. **Acta Universitaria**, Guanajuato v.26, n.3, p.23–32, 2016.
- DERRYBERRY, D. **Basic data analysis for time series with R**. Nova Jersey: Wiley, 2014.
- DORADO G., D.; TAFUR H., H.; RIOS R., L. Rendimiento y calidad de la fruta del Maracuyá Amarillo (*Passiflora edulis*) en respuesta a la combinación de riego y fertilización. **Ingeniería de Recursos Naturales y del Ambiente**, Bogotá, v.12, p.109–117, 2013.
- FAO. **Respuesta del rendimiento de los cultivos al agua**. Roma, 2012.
- FISCHER, G.; MELGAREJO, L.M.; CUTLER, J. Pre-harvest factors that influence the quality of passion fruit: A review. **Agronomía Colombiana**, Bogotá, v.36, n.3, p.217–226, 2018.
- FISCHER, G.; MELGAREJO, L.M. Factores que influyen en la calidad pre y poscosecha de las frutas pasifloráceas. *In: CONGRESO LATINOAMERICANO Y MUNDIAL DE PASIFLORAS*, 3., 2017, Neiva, Huila. **Anais [...]**. San Agustín-Huila: Corporación Cepass, 2017.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. *In: INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV)*, 2015, Washington. **Proceedings [...]**. Washington: Computer Vision Foundation, 2015. p.1026–1034.
- HENGL, T. **A practical guide to geostatistical mapping**. Luxembourg: Office for Official Publications of the European Communities, 2009. 270 p.
- HENGL, T.; HEUVELINK, G.B.M.; ROSSITER, D.G. About regression-kriging: From equations to case studies. **Computers & Geosciences**, Oxford, v.33, n.10, p.1301–1315, 2007.
- IDEAM - Instituto de Hidrología, Meteorología y Estudios Ambientales. **Efectos del cambio climático en la producción y rendimiento de cultivos por sectores**. Bogotá, 2019. Disponible em: <http://www.ideam.gov.co/documents/21021/21138/Efectos+del+Cambio+Climatico+en+la+agricultura.pdf/3b209fae-f078-4823-afa0-1679224a5e85>. Acceso em: 11 jul. 2019.
- IDEAM - Instituto de Hidrología, Meteorología y Estudios Ambientales. **Actualización del componente meteorológico del modelo institucional del IDEAM sobre el efecto climático de los fenómenos El Niño y La Niña en Colombia, como insumo para el Atlas Climatológico**. Bogotá, 2014.
- JENSEN, M. Water consumption by agricultural plants. *In: KOZLOWSKI, T.T. Water deficits and plant growth: development, control, and measurement*. New York: Academic Press, 1968. p.1–22.

- KANTANANTHA, N.; STEWART, H.M. **Crop decision planning under yield and price uncertainties**. Atlanta: Georgia Institute of Technology, 2007.
- KINGMA, D.P.; BA, J. Adam: a method for stochastic optimization. *In: INTERNATIONAL CONFERENCE FOR LEARNING REPRESENTATIONS*, 3., 2015. San Diego. **Conference paper** [...]. Disponível em: <http://arxiv.org/abs/1412.6980>. Acesso em: 10 jun. 2019.
- KUMAR SEN, P. Estimates of the regression coefficient based on Kendall's tau. **Journal of the American Statistical Association**, New York, v.63, n.324, p.1379–1389, 1968.
- KUSUMASTUTI, R.D.; DONK, D.P.VAN; TEUNTER, R. Crop-related harvesting and processing planning: a review. **International Journal of Production Economics**, Amsterdam, v.174, p.76–92, 2016.
- KUTNER, M.H.; NACHTSHEIM, C.J.; NETER, J.; LI, W. **Applied linear statistical models**. 5.ed. New York: Mc Graw Hill Irwin, 2005.
- LAWLER, E.L.; WOOD, D.E. Branch-and-bound methods: a survey. **Operations Research**, Würzburg, v.14, n.4, p.699–719, 1966.
- MARTÍNEZ VENTURA, C. Pronósticos de producción agrícola. **Archivos de Economía**, Bogotá, n.305, p.1–31, 2006. Disponível em: http://bibliotecadigital.agronet.gov.co/bitstream/11348/5014/1/200652212721_Pronósticos_de_producción_agrícola.pdf. Acesso em: 24 set. 2018.
- MAYORGA, M.; FISCHER, G.; MELGAREJO, LUZ M.; PARRA C, A. Growth, development and quality of *Passiflora tripartita* var. *mollissima* fruits under two environmental tropical conditions. **Journal of Applied Botany and Food Quality**, Quedlinburg, v.93, p.66–75, 2020.
- METCALFE, A.V.; COWPERTWAIT, P.S.P. **Introductory time series with R**. New York: Springer New York, 2009.
- MOSCHINI, G.; HENNESSY, D.A. Uncertainty, risk aversion, and risk management for agricultural producers. *In: PINGALI, P.K.; EVENSON, R. Handbook of agricultural economics*. Amsterdam: Elsevier, 2001. v.1.
- MUSSHOFF, O.; HIRSCHAUER, N. What benefits are to be derived from improved farm program planning approaches? – The role of time series models and stochastic optimization. **Agricultural Systems**, Oxford, v.95, n.1–3, p.11–27, 2007.
- OPENSOLVER. **OpenSolver for Google Sheets: OpenSolver for Excel**. 2019. Disponível em: <https://opensolver.org/opensolver-for-google-sheets/>. Acesso em: 9 jul. 2020.
- PANNELL, D.J.; MALCOLM, B.; KINGWELL, R.S. Are we risking too much? Perspectives on risk in farm modelling. **Agricultural Economics**, Reading, v.23, n.1, p.69–78, 2000.
- R Foundation. **R: the R Project for statistical computing**. 2019. Disponível em: <https://www.r-project.org/>
- RUÍZ RAMÍREZ, J.; HERNÁNDEZ RODRÍGUEZ, G.E.; ZULETA RODRÍGUEZ, R. Análisis de series de tiempo en el pronóstico de la producción de caña de azúcar. **Tierra Latinoamericana**, Chapingo, v.29, n.1, p.103–109, 2011.
- SCIKIT-LEARN. **Multi layer perceptron regressor**. 2019. Disponível em: <https://www.python-machinelearning.com/multi-layer-perceptron-regressor/>
- SHUMWAY, R.; STOFFER, D.S. **Time series regression and exploratory data analysis**. New York: Springer, 2017.
- SOTO G., A.M.; COTES T., J.M.; RODRÍGUEZ, D.C. Modelo de simulación del crecimiento y desarrollo de la papa criolla. **Ciencia en Desarrollo**, Tunja, v.9, n.1, 2017.
- STEWART, J.I.; HAGAN, R.M.; PRUITT, W.O. **Production functions and predicted irrigation programmes for principal crops as required for water resources planning and increased water use efficiency**. Washintong: US Department of Interior, 1976.
- TORRENTE T.A. **Potencial de agua en el suelo y su relación hídrica en los cultivos de Passifloras**. Neiva: Colciencias, 2009.
- VANDENDRIESSCHE, H.J.; VAN ITTERSUM, M.K. Crop models and decision support systems for yield forecasting and management of the sugar beet crop. **European Journal of Agronomy**, Amsterdam, v.4, n.3, p.269–279, 1995.