Australian Journal of Crop Science

AJCS

AJCS 9(11):1049-1056 (2015) ISSN:1835-2707

Agrometeorological models for forecasting yield and quality of sugarcane

Marcos Antônio Marcari, Glauco de Souza Rolim, Lucas Eduardo de Oliveira Aparecido*

UNESP–São Paulo State University, Department of Exact Sciences, 14884-900, Jaboticabal, SP, Brazil

***Corresponding author: marcos.marcari@hotmail.com**

Abstract

Climate is an important factor in sugarcane production, and its study is fundamental for understanding the climatic requirements of the crop. We developed regional agro-meteorological models to forecast monthly yields in tonnes of sugarcane per hectare (TCH) and quality of the total recoverable sugar (ATR). We used monthly climatological data (air temperature, precipitation, water deficiency and surplus, potential and actual evapotranspiration, soil-water storage, and global solar irradiation) of the previous year to forecast TCH and ATR for the next year using multiple linear regression. The parameters of monthly climatological data were chosen for their small mean absolute percentage errors (MAPEs) and $p < 0.05$ and ability to model longer periods of prediction. Data for Jaboticabal, a major area of sugarcane production in the state of São Paulo, Brazil, from 2002-2009 were used for calibration, and data from 2010-2013 were used for validation. All calibrated models were significant ($p < 0.05$) and accurate, with a MAPE of 4.06% for the forecast of TCH in the ambient "C" for July. The model calibrated for November had variable water deficits in all environments, showing the importance of this variable to the crops. The monthly models tested performed well. For example, the forecast by TCH_{MAY} in the AB environment (MAPE = 1.89% and adjusted coefficient of determination = 0.90) overestimated the average yield of 90.6 t ha^{-1} by only 1.7 t ha^{-1}. The predictive period for forecasting TCH_{MAY} was eight months when the last climatological parameter used in the model was DEF_{SEP}.

Keywords: Estimate yield; Water balance; Climate.

Abbreviations: TCH_tonnes of sugarcane per hectare; ATR_total recoverable sugar; DEF_water deficiency; PET_potential evapotranspiration; WS_water storage in the soil; CWB_climatological water balance; T_air temperature; Qg_global irradiation; P_rainfall; EXC_water surplus; PE_production environments: AB, C, DE high, medium and low PE .

Introduction

The sugarcane agribusiness is a competitive activity that uses high technology. The success of the business depends on sales management associated with forecasting yields. Early information is thus necessary for planning activities, budgeting for the period, and forecasting the amount of raw material available. Crop models can forecast yield and quality. Estimates and forecasts made by models, however, should be clearly distinguished. Estimates use historical data and evaluate current conditions (Carriero et al., 2009), and forecasts attempt to predict the future, with current data simulating future conditions (Clements et al., 2012). This information linked to the complexity of the sugarcane crop is important due to the number of products that influence prices, so companies need agility and accurate information for following the market and deciding when and how much of a harvest can be marketed (Leite et al., 2008). Sugarcane is a semi-perennial plant of great importance to Brazil. The crop cycle is approximately five years (Milk et al., 2008). The diversity of local climates and soils and the selection of varieties best adapted to the environment are crucial aspects for obtaining the best economic return in the production cycle (UNICA, 2012). The estimation of yield and of the technological index that expresses the amount of total recoverable sugar (ATR) also plays an essential role in planning budget inputs, harvesting, loading, transport of raw materials, processing, storage, and marketing of products (Scarpari and Beauclair, 2009). Accurate climatic forecasting provides important information, especially for making decisions and for the use of soil and/or water on farms (Cabrera et al., 2006). Sugarcane yields are commonly forecasted one month before harvest based on experience. These forecasts are made without the use of statistical parameters (Schmidt et al., 2001). Some forecasters use agro-meteorological models to estimate regional yields and sugar levels. Crop modelling should address the most important aspects of the interaction between and management of climate, plants, and soil, but climate is the main factor determining agricultural yield (Hoogenboom, 2000; Aparecido et al. 2015). The development of computer models and their implementation in agro-climatic information systems are important for planning and for increasing agricultural yields (Marin et al., 2011). Models for estimating crop yield based on agro-meteorological principles simulate the stages of development and maturation of crops, the availability of soil moisture, and the effects of water stress on crop yield (Heinemann et al., 2002). The models can be dynamic/mechanistic or statistical. Dynamic/mechanistic models describe daily changes in crop variables and include main morphological and physiological processes. The physiological understanding of the processes of growth and plant development have sought to improve these models and thus the estimates of the models (Gouveia et al., 2009). The effect of water stress on different stages of development is an important aspect of these models (Doorenbos and Kassam, 1979). Various mechanistic models have been developed for simulating sugarcane growth, such as Auscane (Jones et al., 1989), DSSAT/Canegro (Inman-BambeR, 1991), QCane (Liu and Kingston, 1995), APSIM (Keating et al., 1999), and Mosicas (Martine, 2003). These models require parameter calibration to adjust to the characteristics of the soil, climate, and crop genotype. Statistical models seek a quantitative relationship between yield/quality and various climatic factors, soils, and associated management (Oliveira et al., 2013). Numerical forecasting techniques based on

Table 1. Monthly agrometeorological models to forecast total recoverable sugar (ATR). Legend: The independent variables are, T = air temperature (°C), EXC and DEF = surplus and deficit water (mm), WS = water storage (mm), P = rainfall (mm), PET = potential evapotranspiration (mm), Qg = global radiation (w m^{-2}) of the previous year. The subtitles indicate the months.

ATR Models	Calibration			Test	
	p-Value	MAPE	R^2 Adj	MAPE	R^2 Adj
May					
$ATR = 0{,}0566 \cdot STO_{MAY} + 0{,}0744 \cdot PET_{SEP} - 0{,}0345 \cdot PET_{DEC} + 0{,}0403 \cdot P_{OCT} + 120{,}0658$	0.033	0.228	0.951	1.606	0.667
June					
$ATR = -0{,}0632 \cdot EXC_{JAN} + 0{,}1762 \cdot WS_{MAY} + 0{,}0995 \cdot PET_{MAY} + 4{,}0855 \cdot T_{MAY} + 45{,}5615$	0.028	0.244	0.957	1.176	0.879
July					
$ATR = -0{,}1373 \cdot DEF_{SEP} + 0{,}0888 \cdot DEF_{OCT} + 0{,}2106 + PET_{APR} + 0{,}9202 \cdot Qg_{AUG} + 110{,}43$	0.003	0.124	0.996	0.473	0.976
August					
$ATR = -0{,}2455 \cdot DEF_{JUN} + 0{,}4211 \cdot PET_{JAN} - 0{,}7641 \cdot Qg_{SEP} + 4{,}4104 \cdot T_{MAY} + 27{,}7477$	0.017	0.410	0.970	0.660	0.950
September					
$ATR = 0{,}0593 \cdot DEF_{OCT} + 0{,}1871 \cdot P_{MAY} + 0{,}1220 \cdot P_{NOV} + 5{,}0200 \cdot Qg_{JUN} + 49{,}1647$	0.026	0.520	0.960	0.930	0.971
October					
$ATR = -1{,}2075 \cdot DEF_{FEB} + 0{,}2509 \cdot STO_{AUG} - 0{,}4817 \cdot PET_{AUG} + 3{,}5906 \cdot Qg_{OCT} + 92{,}4914$	0.005	0.199	0.992	1.032	0.871
November					
$ATR = 0{,}0268 \cdot EXC_{JAN} - 0{,}2898 \cdot WS_{MAY} - 0{,}0240 \cdot P_{FEB} + 1{,}5479 \cdot T_{MAC} + 106{,}5084$	0.031	0.419	0.953	0.435	0.964

MAPE = mean absolute percentage error.

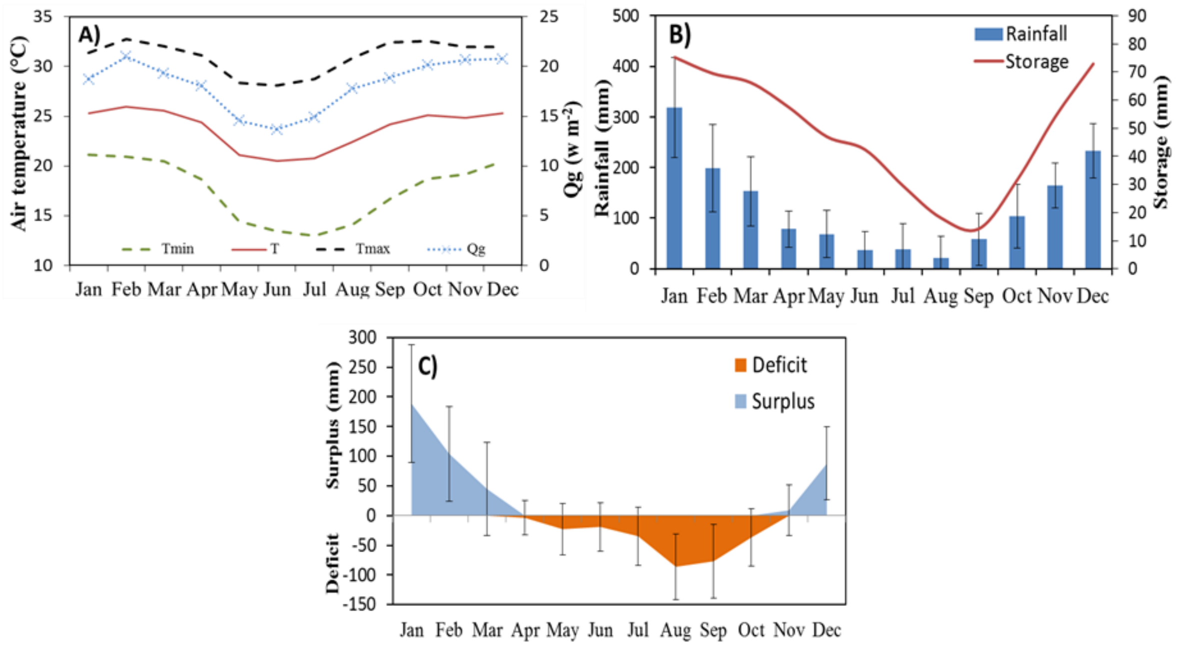

Fig 1. Characterization climate monthly to the location of Jaboticabal – SP (2001 to 2013), A) the average monthly air temperature (minimum, average and maximum) (°C) and the global radiation (W m^{-2}), average monthly rainfall, C) water storage in the soil (mm) and components of the climatic water balance (surplus and water deficit, mm). The vertical bar means the standard deviation monthly values.

agro-meteorology are normally statistical relationships between the dependent variables that need to be estimated (yield) and the independent agro-meteorological variables (weather) (Heinemann et al., 2002; Araújo et al., 2014). Published statistical models such as that by Moreto and Rolim (2015) have used statistical methods of regression for monthly water deficits (WDs) to predict the yield of "Valencia" oranges in São Paulo. Carvalho et al. (2004) also used regression models to predict the yield of coffee in the state of Minas Gerais. The model simulations depended on the homogeneity of climatic conditions, agricultural practices, and soil characteristics (Scarpari and Beauclair, 2009).

Alternative approaches also use neural networks and expert systems, both components of artificial intelligence (Carvalho et al., 1998). Few models for forecasting yield have been based on probabilities using various statistical models. Mkhabela et al. (2004) forecasted maize yields in South Africa two months before the harvest. Savin et al. (2007) used neural networks to predict wheat yield in Russia during flowering with 74% accuracy. Gauranga and Ashwani (2014) forecasted rice yield in India 30 days before the harvest. The use of the scientific method in the development of agrometeorological models and in decision-making, replacing intuition and/or practical experience, provides a powerful tool for the possible realisation of advantageous scenarios, operations, competitive differentials, and improvements in agricultural planning (Gouveia et al., 2009).

Our aim was thus to develop regional agro-meteorological models for forecasting sugarcane yield, in tonnes of sugarcane per hectare (TCH), and quality of ATR on a monthly scale.

Results and Discussion

Meteorological and water analyses

The average air temperature (T) at Jaboticabal ranged from 21 to 26 °C, with a maximum of 33 °C in summer and a minimum of 13 °C in winter. The global irradiation (Qg) ranged from 14 to 21 W m^{-2}. The average rainfall (P) was 180 mm from October to March and 40 mm from April to September. Less water was stored in the soil in winter, but storage was lowest in September, with a water capacity (WC) of 18 mm. Analysing the components of the climatological water balance (CWB), we observed a water

Table 2. Monthly agrometeorological models to forecast tons of sugarcane per hectare (TCH). Legend: The independent variables are, T = air temperature (°C), EXC and DEF = surplus and deficit water (mm), WS = water storage (mm), P = rainfall (mm), PET = potential evapotranspiration (mm), Qg = global radiation (w m^{-2}) of the previous year. The subtitles indicate the months.

Yield Environment	TCH Models	Calibration P Value	Calibration MAPE	Calibration R^2 Adj	Test MAPE	Test R^2 Adj
	May					
AB	$TCH = -0{,}8047 . DEF_{JAN} - 0{,}0370 . DEF_{AUG} + 0{,}7723 . DEF_{DEC} + 0{,}2196 . WS_{JUL} + 87{,}7535$	0.049	1.340	0.926	1.893	0.909
C	$TCH = 0{,}5854 . PET_{FEB} + 0{,}1565 .P ET_{MAC} + 0{,}2419 . P_{AUG} - 2{,}9197 . Qg_{JUN} + 42{,}3904$	0.002	0.331	0.997	5.739	0.973
DE	$TCH = -0{,}2608 . DEF_{SEP} - 2{,}1021 . Qg_{MAY} - 3{,}3693 . Qg_{OCT} + 3{,}6777 . T_{JAN} + 118{,}2424$	0.004	0.420	0.993	3.052	0.828
	June					
AB	$TCH = 0{,}0306 . EXC_{JAN} + 0{,}1846 . PET_{AUG} + 0{,}1187 . P_{APR} + 6{,}3202 . Qg_{JUL} - 19{,}0031$	0.090	1.118	0.862	2.124	0.867
C	$TCH = 0{,}1520 . WS_{APR} + 1{,}1584 . PET_{JAN} + 0{,}9407 . PET_{JUL} + 2{,}5358 . T_{JUL} -121{,}0817$	0.053	1.624	0.918	5.114	0.869
DE	$TCH = 0{,}7212 . DEF_{MAY} + 0{,}6325 . WS_{JUL}+ 0{,}5104 . PET_{MAC} - 1{,}7991 . Qg_{OCT} + 41{,}5761$	0.002	0.342	0.997	3.555	0.821
	July					
AB	$TCH = 0{,}2687 . WS_{JUN} + 0{,}2690 . PET_{JAN} + 0{,}0767 . P_{MAC} - 2{,}1432 . Qg_{JUN} + 73{,}0103$	0.084	1.284	0.872	3.295	0.840
C	$TCH = 1{,}2609. DEF_{MAY} + 0{,}3897 . EXC_{FEB}+ 1{,}9317 . WS_{JUN} - 11{,}3161 . Qg_{AUG} + 150{,}0125$	0.033	4.064	0.949	5.659	0.974
DE	$TCH = -0{,}0928 . DEF_{OCT} + 0{,}5611 .P ET_{MAY} + -0{,}0529 . P_{FEB}. -1{,}3961 . Qg_{OCT} + 103{,}7103$	0.000	0.146	1.000	2.015	0.982
	August					
AB	$TCH= 0{,}3013 . PET_{AUG} + 0{,}1838 . P_{APR} + 0{,}0865 . P_{OCT} + 3{,}0902 . Qg_{FEB} - 5{,}8093$	0.001	0.186	0.999	4.222	0.955
C	$TCH = 0{,}1598 . DEF_{JAN} + 0{,}7641 . WS_{MAY} -0{,}4127. PET_{APR} + 0{,}712 . Qg_{APR} + 67{,}7622$	0.082	1.736	0.875	4.621	0.969
DE	$TCH = 0{,}9012 . WS_{MAY} -0{,}6597. PET_{JUN} -3{,}3852 . Qg_{MAC} + 4{,}8836 . T_{FEB} + 7{,}5267$	0.002	0.412	0.996	7.523	0.863
	September					
AB	$TCH = -0{,}1386 . EXC_{JAN} - 2{,}1599 . WS_{DEC} + 0{,}3089 . P_{JUN} + 5{,}7292 . T_{OCT} + 118{,}3829$	0.051	0.849	0.922	5.583	0.675
C	$TCH = -2{,}5432 . WS_{JAN} + 1{,}1436. WS_{MAY} + 12{,}1593 . Qg_{APR} - 4{,}1671 . T_{MAC} + 122{,}3014$	0.051	1.586	0.923	6.850	0.991
DE	$TCH = 0{,}3245 . DEF_{APR} -0{,}3145 .PET_{JUN} - 7{,}7541 . Qg_{AUG} + 1{,}9522 . T_{SEP} + 170{,}9180$	0.001	0.455	0.999	3.829	0.891
	October					
AB	$TCH = 0{,}0488 . EXC_{OCT} + 0{,}3837 . WS_{MAY} + 0{,}4429 . PET_{AUG} + -0{,}2173 . P_{JUL} + 59{,}6127$	0.096	1.565	0.852	3.286	0.962
C	$TCH = 0{,}0609 . DEF_{MAY} - 0{,}0514 . EXC_{JAN} - 0{,}0847 . EXC_{FEB} + 0{,}2401 . PET_{MAY}+ 84{,}9556$	0.009	0.433	0.986	3.204	0.707
DE	$TCH = -0{,}0860 . EXC_{FEB} + 1{,}7785 . WS_{JAN} - 0{,}2086 . WS_{SEP} + 0{,}6877 . PET_{OCT} - 97{,}4830$	0.005	0.679	0.993	5.836	0.650
	November					
AB	$TCH= -0{,}6352 . DEF_{NOV} + 0{,}3010 . P_{MAY} + 6{,}3428 . Qg_{JAN} - 6{,}4148 . Qg_{AUG} + 68{,}0824$	0.008	0.570	0.987	2.177	0.945
C	$TCH = -0{,}7840 . DEF_{JUN} - 0{,}5282 . EXC_{OCT} + 6{,}5426 . Qg_{SEP} - 4{,}0980. T_{MAC} + 91{,}0047$	0.009	0.203	1.000	5.117	0.628
DE	$TCH = -1{,}9463 . DEF_{JAN} + 0{,}8956 . DEF_{SEP} + 1{,}0285 . WS_{AUG} - 10{,}1168 . T_{JAN} + 239{,}5707$	0.016	2.372	0.975	5.634	0.953

MAPE = mean absolute percentage error.

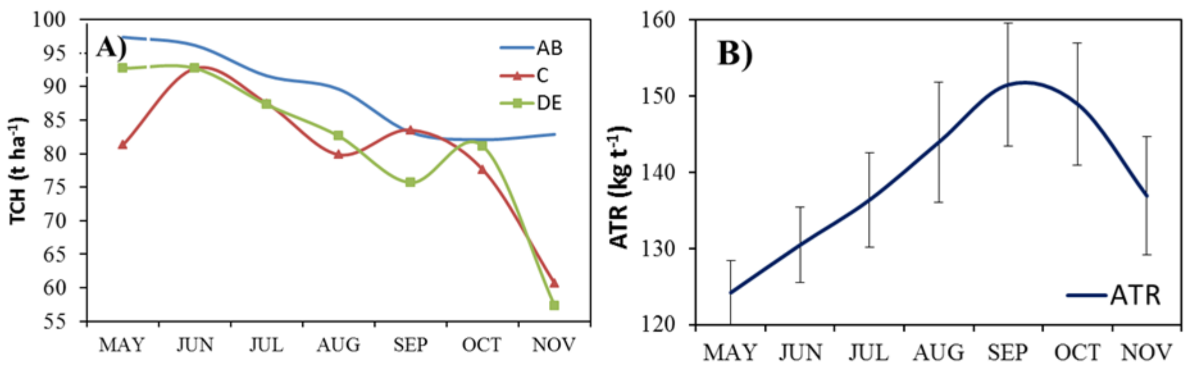

Fig 2. Monthly average of tons of sugarcane per hectare (TCH) (A) and total recoverable sugar (ATR) (B) in the sugarcane crop in the period of May to November. The vertical bars means the standard deviation observed values.

surplus (EXC) from November to April, reaching 200 mm in January. WDs occurred from May to October but were severest in August, reaching 75 mm (Figure 1).

Characterisation of yield and quality

Monthly TCH and ATR varied with production environment (PE). Environment AB had favourable characteristics of water storage and nutrition, which contributed to a greater longevity of the crop and a high mean yield. The crops did not last as long in environments C and DE, so the sugarcane fields needed to be renewed earlier than for AB. WDs decreased TCH. ATR increased after May and peaked in September (Figure 2).

Forecasting yield and quality

The agro-meteorological models developed to forecast sugarcane yield and quality using the APC (all possible combinations) method with up to four variables had a total of 1584740 possible combinations of independent variables (Figure 3) for each month, each dependent variable (TCH and ATR), and each environment (AB, C and DE). The removal of models with multicollinearity (849 317) left 735 423 viable models from which the best forecasting models for Jaboticabal were selected. We tested a total of 6338960 equations for each month. The APC method was effective; as accuracy increased (the adjusted coefficient of determination (R^2adj ≈ 1.00), and the mean absolute percentage error (MAPE) approached zero (Figure 4). The best models for forecasting ATR indicated that T_{JUN} and T_{NOV} were the most important variables in most environments, with positive coefficients indicating positive relationships (Figure 5). T_{AUG}, T_{NOV}, and T_{AUG}, likely affecting the end of tillering, were the most important variables for the models for predicting TCH in environments AB, C, and DE, respectively, with direct relationships to yield. All calibrated models were significant and accurate ($p < 0.05$, MAPE = 7.52%). ATR_{JUL} and ATR_{NOV} were the most accurate models with the lowest *p*-values and MAPEs (Table 1). Tests for the models of ATR forecasting for each month were accurate, with minimum and maximum MAPEs of 0.435 and 1.60% for the November and May forecasts,

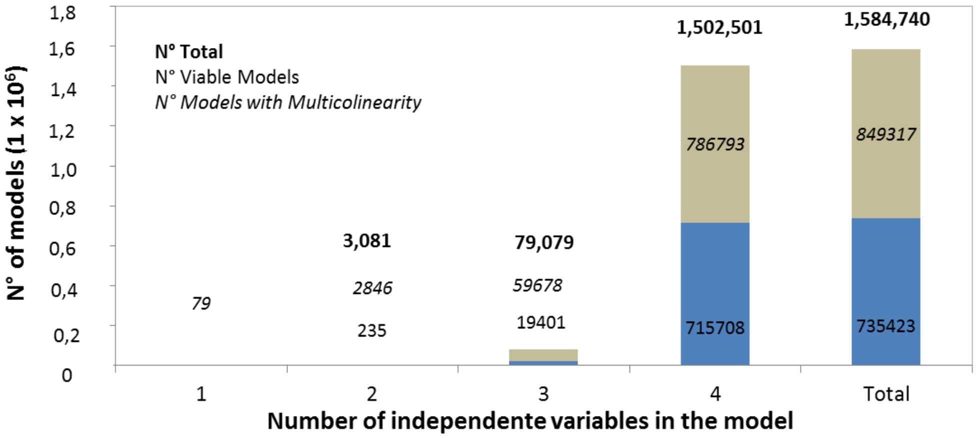

Fig 3. Example of total number of generated models, models with multicollinearity and viable models for one environment (i.e. 'A'), one dependent variable (i.e. 'TCH') and one month. Blue = viable models. Gray = models with multicollinearity.

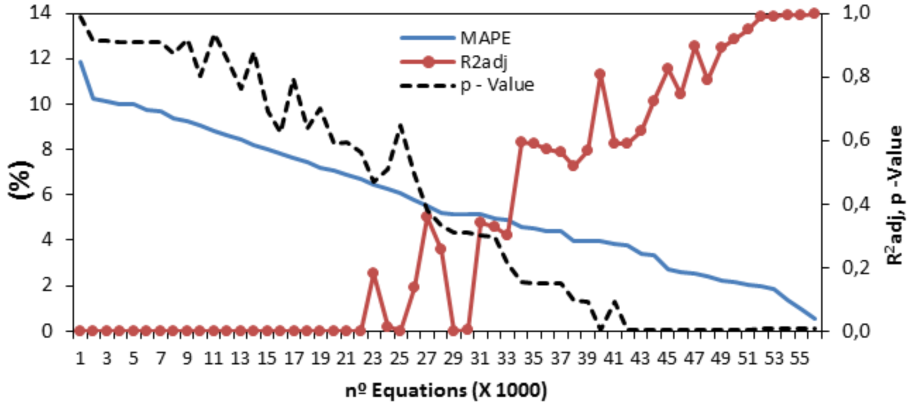

Fig 4. Example of a models classification according to criteria the accuracy (lowest MAPE "mean absolute percentage error"), precision (greater R^2 adj) and reliability (p-value).

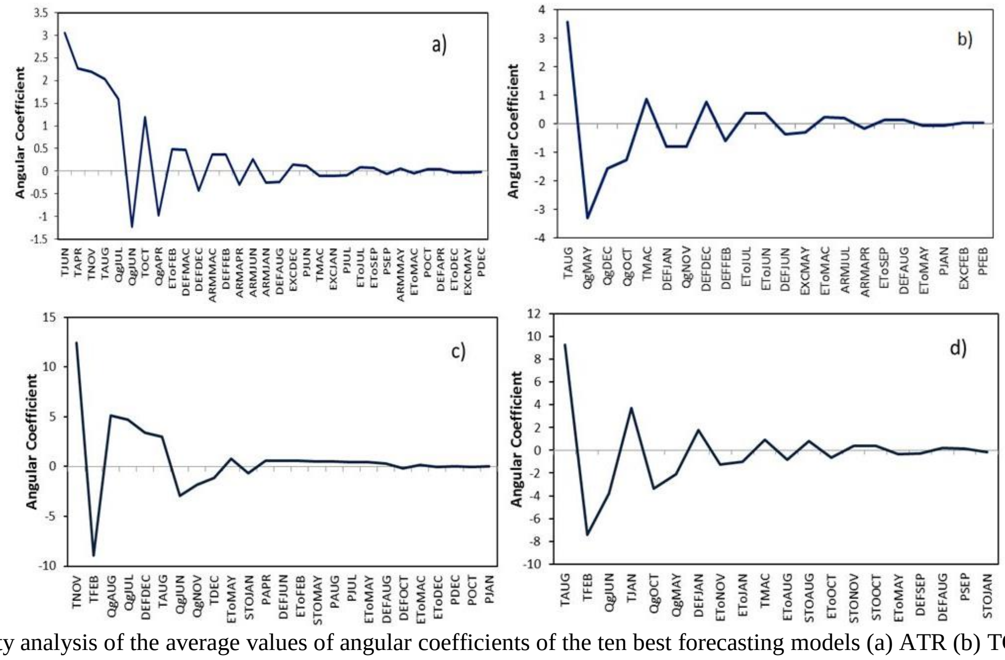

Fig 5. Sensitivity analysis of the average values of angular coefficients of the ten best forecasting models (a) ATR (b) TCH (environment AB), (c) TCH (environment C) and (d) TCH (environment DE).

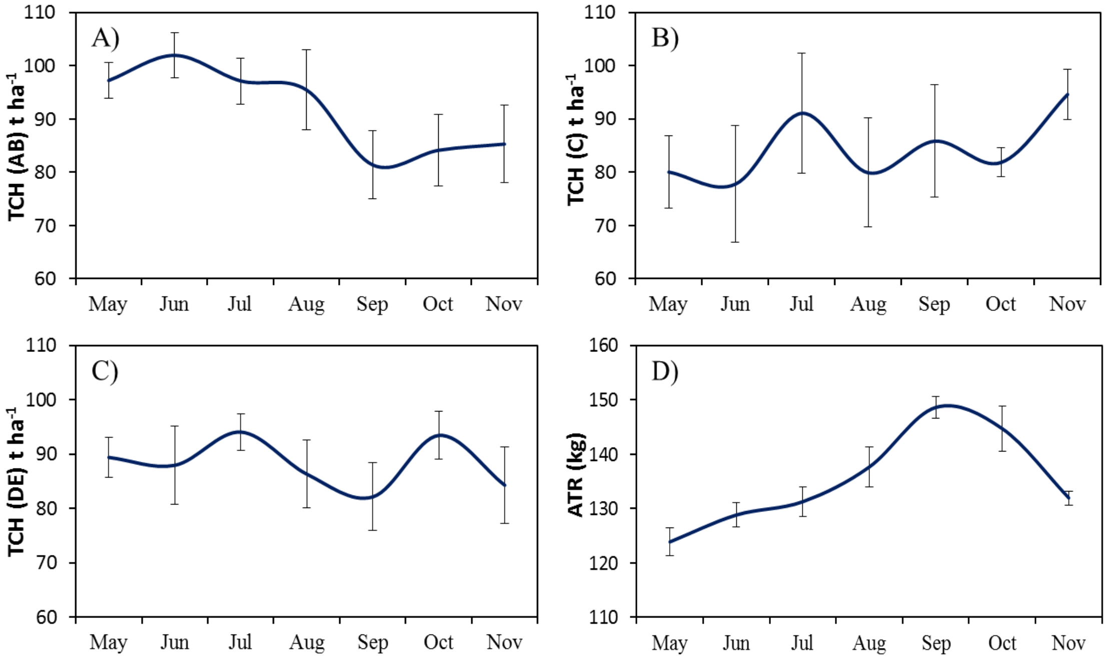

Fig 6. Estimated monthly data of TCH in the AB (A), C (B) and D (C) environments and ATR (D) with the confidence interval (95% of probability).

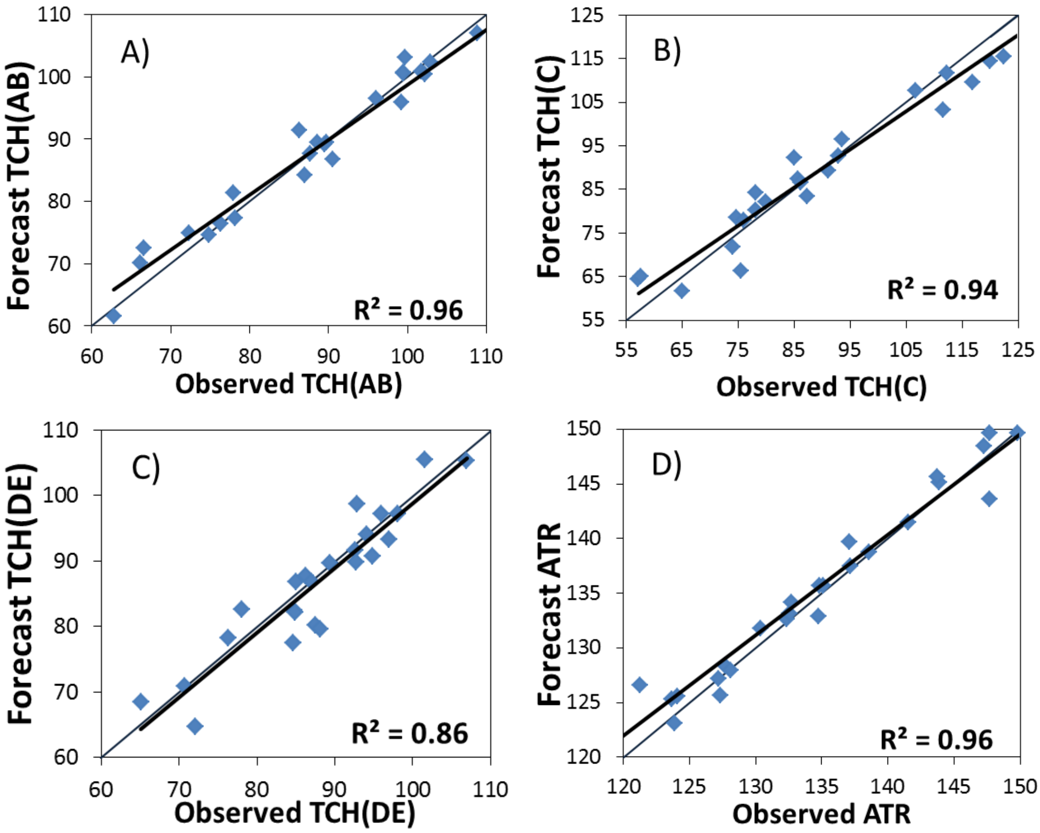

Fig 7. Independent (Test) Analysis of the precision of monthly forecasts of the models of yield in environments AB (A), C (B) e DE (C) and total recoverable sugar (D).

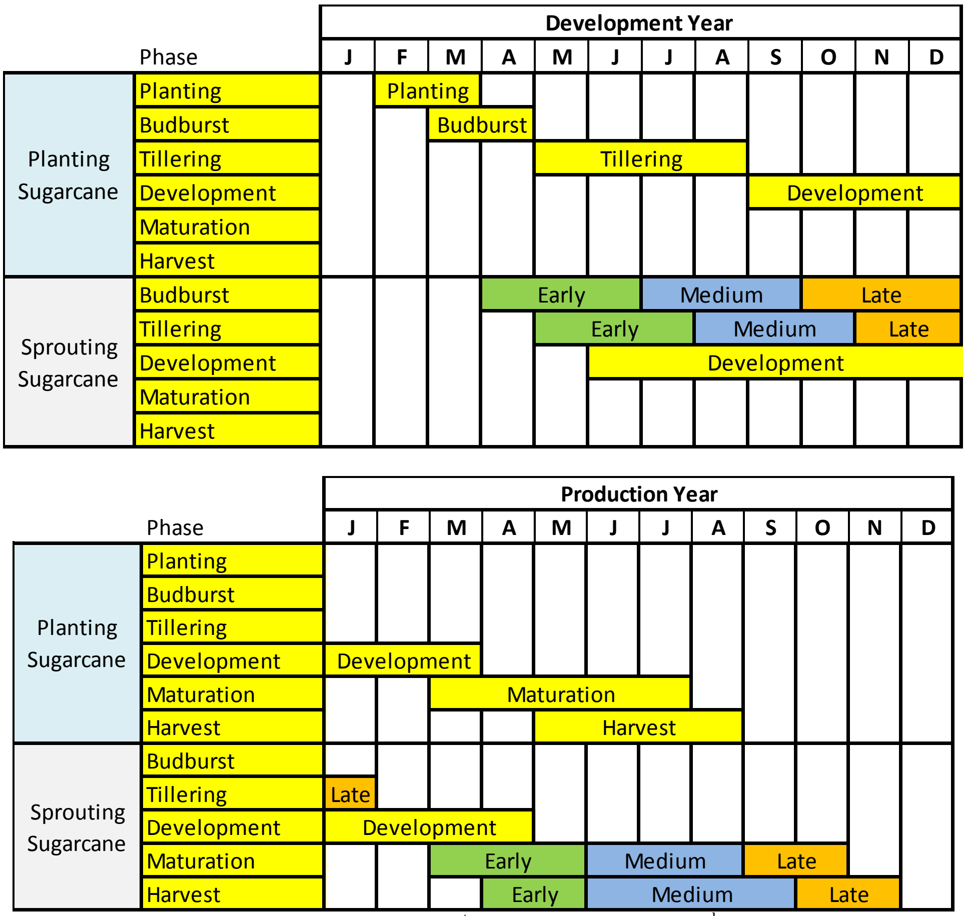

Fig 8. Phenology in the year of development (1st year) and production (2nd year) for Sugarcane crop.

respectively (Table 1). The most influential variables were WS and potential evapotranspiration (PET), with high angular coefficients. Other variables such as P and DEF were important in the model calibrated for September. The ATR_{NOV} model could forecast 17 months in advance, with May T as the most important variable. The ATR_{MAY} model had the shortest forecasting period of five months (Table 1). The forecast of yield is an excellent tool for planning new planting (Bocca et al., 2015). The models for forecasting TCH in the test period were accurate, with a minimum MAPE of 1.89% in the AB environment in May. Qg and DEF were the most influential variables in the agro-meteorological models for forecasting TCH (Table 2), with high angular coefficients. DEF is a major constraint that reduces the yield of crops (Khamssi et al., 2011). DEF occurred in every environment for the model calibrated for November, illustrating the importance of this variable to the crop. The TCH_{MAY} forecast was best in the AB environment, with a MAPE of 1.89% and R^2adj = 0.90 in the tests. Considering an average yield of 90.6 t ha^{-1} the model have a bias of 1.7 t ha^{-1}. DEF_{JAN} and DEF_{AUG} were the most and least influential variables, respectively (Table 2).

The TCH_{OCT} model in environment C had the longest yield forecasting period of 16 months, with PET_{MAY} as the most influential variable. The TCH_{MAY} model in environment DE had the shortest forecasting period of seven months (Table 2). General circulation models used for forecasting weather had forecasting periods for wheat yield in Ukraine of 2-3 months (Kogan et al., 2013). The results of the forecasts for the TCH variables in all environments and ATR during the year are shown in Figure 6. For example, TCH in environment AB in May was 97.2 t ha^{-1} but ranged from 94.2 to 100.1 t ha^{-1} (Figure 6A). The monthly models we tested performed well. The general analysis of the model performances compared to independent data for ATR and TCH with each month had highly precise forecasts, with a minimum R^2adj of 0.86 for TCH in environment DE (Figure 7C). Agro-meteorological models have shown great advantages, because forecasts using climatic variables reduce the risks of agricultural activities and increase the reliability of farming projects (Hammer et al., 2000; Araujo et al., 2014).

Materials and Methods

Climatic and crop characterisation

The study was conducted in Jaboticabal in the northeast of the state of São Paulo (Brazil) (21°19′18.15″S, 48°06′27.81″W; 530 m a.s.l.), an important region of sugarcane production in Brazil. The predominant climate of this region is classified in the Thornthwaite (1948) system as B1rA’a’, humid with low WDs, megathermic, with summer PETs <48% of the annual PET (Rolim and Aparecido, 2015). The region has a P of 1300 mm, concentrated from November to February. The natural vegetation consists of semideciduous tropical forest. The relief is predominantly gently rolling, with an average slope of 3.4%. The experimental area has been under sugarcane cultivation for over 50 years, 16 years with mechanical harvesting without burning.

Daily meteorological data were collected by an automated weather station (Model 21X, Campbell Scientific) for 2001-2013 organised monthly. The station had a data-transmission system (Wireless Vantage Pro Plus) and measured Qg with a Model 6450 sensor, T and relative humidity (RH) with a Model 7859 external sensor, and P with a Model 7852 Rain Collector rain gauge. P data for the experiment were collected from 21 pluviometric stations distributed in the study area. The data for TCH in the region were divided into three PEs corresponding to

those suggested by CTC (1995) and used in all areas of sugarcane production in Brazil. This classification takes into account the production potential due to the integration of three factors: climate, soil, and cultivar. The PEs were high, medium, and low production, corresponding to the CTC classifications A and B (AB), C, and D and E (DE), respectively. The AB, C, and DE PEs consisted of 30000, 21000, and 13000 ha, respectively. The data for ATR (kg) were analysed without PE categorisation.
All meteorological data during production were provided by the sugarcane companies in the Jaboticabal region.
The sugarcane crop in the region has four phenological phases: budburst, tillering, development, and maturation (Gascho and Shih, 1983). The period when T, Qg, and P are low coincides with the end of tillering and the start of development with vegetative dormancy, so the accumulation of fresh matter is small (Figure 8). The crop grows intensely from October to March when conditions of T, P, and Qg are favourable and when 75% of the fresh matter of the crop accumulates. The sugarcane is harvested from October, when weather conditions are good for budding and development, so the operation can damage the crop. This method decreases the exposure of the crop to optimal growing conditions and tends to produce less sugar. PET was estimated on a monthly scale from the meteorological data by the Thornthwaite method (1948). The PET and P data were used to estimate CWB as proposed by Thornthwaite and Mather (1955), with an available WC of 79 mm to determine the CWB components as actual evapotranspiration, water storage in the soil (WS), EXC, and DEF. WC was determined in the laboratory by retention curves, with a field moisture capacity of 22%, wilting-point moisture of 15%, soil bulk density of 1.25 g cm^{-3}, and root depth of 0.90 m.

Model analysis

We performed a correlation analysis with independent variables to identify the meteorological elements and CWBs in different months with higher correlations. The variables with correlation coefficients ≥|0.7| were removed to avoid problems of multicollinearity. The removal of the collinear variables allowed us to understand the weight (slope coefficient) of each monthly climatic variable in the models for forecasting TCH and ATR.
The most relevant variables were selected for the construction of the agro-meteorological models for forecasting sugarcane yield and quality. We used the multiple linear regression method (equation 1):

$$Y = a \times X_1 + b \times X_2 + c \times X_3 + \cdots + CL \quad (1)$$

where Y is the yield (kg ha^{-1}) or ATR (kg) for a specific month; a, b, c, ... are the adjusted coefficients (weights); X_1, X_2, X_3, ... are the monthly meteorological variables (Qg, T, RH, P, PET, WS, EXC, and DEF) and CWBs of the year preceding the harvest from January to December; and CL is the linear coefficient. The largest problem in multiple linear regression is the selection of the independent variables to be combined to generate good models. Any iterative numerical method such as the stepwise method has problems stabilising errors due to improper initial combinations. One option is to test all possible combinations when the number of independent variables is relatively small (Walpole et al., 2012). We used the APC method despite the large number of independent variables ($Qg_{JAN}\ldots Qg_{DEC}$, $T_{JAN}\ldots T_{DEC}$, …) by testing models with up to four independent variables of development year (1st year) on a monthly scale. We developed a Visual Basic for Application algorithm for these calculations. The criteria for selecting the variables were the significance of the coefficients ($t < 0.05$) and regressions ($F < 0.05$), minimal MAPEs, and maximal R^2adj.

Data analysis

The models were selected by evaluating the accuracy by the MAPE (Equation 2) and the precision by R^2adj (Equation 3) (Cornell and Berger, 1987):

$$MAPE = \frac{\sum_{i=1}^{N}\left(\left|\frac{Yest_i - Yobs_i}{Yobs_i}\right| * 100\right)}{n} \quad (2)$$

$$R^2adj = \left[1 - \frac{(1-R^2)\times(n-1)}{n-k-1}\right] \quad (3)$$

where $Yest_i$ is an estimated variable, $Yobs_i$ is an observed variable, $Y_{est\text{-}C}$ is a variable estimated by linear regression between $Yobs_i$ and $Yest_i$, n is the number of years, n is the number of datapoints, and k is the number of independent variables in the regression. The periods 2002-2009 and 2010-2013 were used for calibration and testing, respectively.

We identified the ten most accurate models for forecasting ATR and TCH for each of the productive environments and then ranked the most influential variables for the sensitivity analysis of the angular coefficients. We used the ten best models for each month for the sensitivity analysis to identify the most influential meteorological elements in each month and the confidence intervals for forecasting.

Conclusions

Variable selection and the multiple linear regressions were efficient in the construction of models for predicting sugarcane yield and quality in different environments of the Jaboticabal region of São Paulo. The development of accurate models for each month as functions of climatic variables was possible. The minimum period for the prediction of yield and quality was five months.

References

Araújo EC, Uribe-Opazo MA, Johann JA (2014) Modelos de regressão espacial para estimativa da produtividade da soja associada a variáveis agrometeorológicas na região do estado do Paraná. Eng Agric. 34:286-299.

Aparecido LEO, Rolim GS, Souza PS (2015) Sensitivity of newly transplanted coffee plants to climatic conditions at altitudes of Minas Gerais, Brazil. Aust J Crop Sci. 9(2):160-167.

Bocca FF, Rodrigues LHA, Arraes NAM (2015) When do i want to know and why? Different demands on sugarcane yield predictions. Agr Syst. 135:48–56.

Cabrera VE, Fraisse CW, Letson D, Podest´a G, Novak J (2006) Impact of climate information on reducing farm risk by optimizing crop insurance strategy. Trans ASABE. 49:1223–1233.

Carriero A, Kapetanios G, Marcellino M (2009) Forecasting exchange rates with a large bayesian var. Int J Forecasting. 25(2):400-417.

Carvalho LG, Sediyama GC, Cecon PR, Alves HMR (2004) Modelo de regressão para a previsão de produtividade de cafeeiros no estado de Minas Gerais. Rev Bras Eng Agric Ambient. 8:204-211.

Carvalho APLF, Braga AP, Ludemir TB (1998) Fundamentos de redes neurais artificiais. In: Escola de Computação, 11., 1998. Rio de Janeiro: UFRJ. 1998. 246p.

Clements MP, Galvão AB (2012) Forecasting with vector autoregressive models of data vintages: US output growth and inflation. Int J Forecasting. Available at: http://wrap.warwick.ac.uk/44603/. (Accessed 28 march 2015).

CTC (1995) Centro de tecnologia copersucar. Potencial de produção. Boletim agrícola, p.17.

Doorenbos J, Kassam AH (1979) Yield response to water. FAO: Irrigation and Drainage Paper, 33, Rome, FAO, 172 p.

Draper NR, Smith H (1966) Applied regression analysis. New York: John Wiley & Sons. 407 p.

Gouranga K, Ashwani K (2014) Forecasting rainfed rice yield with biomass of early phenophases, peak intercepted PAR and ground based remotely sensed vegetation indices. J Agron. 16:94-103.

Gouvêa JRF, Sentelhas PC, Gazzola ST, Santos MC (2009) Climate changes and technological advances: impacts on sugarcane productivity in tropical Southern Brazil. Sci Agric. 66(5):593-605.

Hammer GL, Carberry P, Stone R (2000) Comparing the value of seasonal climate forecasting systems in managing cropping systems. In Applications of seasonal climate forecasting in agricultural and natural ecosystems – The Australian Experience, Hammer Gl, Nicholls N Mitchell C (eds). Kluwer Academic Publishers: Dordrecht, Netherlands; 183–195.

Heinemann AB, Hoogenboom G, Faria RT (2002) Determination of spatial water requirements at country and regional levels using crop models and GIS. An example for the state of Parana, Brazil. Agr Water Manage. 52:177-196.

Hoogenboom G (2000) Contribution of agrometeorology to the simulation of crop production and its applications. Agr Forest Meteorol. 103(2):137–157.

Inman Bamber NG (1991) A growth model for sugarcane based on a simple carbon and the CERES – Maize water balance. S Afr J Pl. 8:93-99.

Jones CA, Wegener MK, Russell JS, Mcleod IM, Williams JR (1989) AUSCANE—Simulation of australian sugarcane with EPIC. CSIRO Division of tropical crops and pastures. Trop Agron Tech. 29:99.

Khamssi NN, Golezani KG, Najaphy A, Zehtab S (2011) Evaluation of grain filling rate, effective grain filling period and resistance indices under acclimation to gradual water deficit stress in chickpea cultivars. Aust J Crop Sci. 5(6):1044-1049.

Keating BA, Robertson MJ, Muchow RC, Huth NI (1999) Modeling sugarcane production systems. Develepment and performance of the sugarcane module. Field Crops Res. 271p.

Kogan F, Kussul N, Adamenko T, Skakun S, Kravchenko O, Kryvobok O, Shelestov A, Kolotii a, Kussul O, Lavrenyuk A (2013). Winter yield forecasting in ukraine based on earth observation, meteorological data and biophysical models. Int J Ap E Geo. 23:192-203.

Leite GMV, Andrade LAB, Garcia JC, Anjos IA (2008) Efeitos de fontes e doses de silicato de cálcio no rendimento agrícola e na qualidade tecnológica da cana-de-açúcar, cultivar SP80-1816. Ciênc Agrotec. 32(4):1120-1125.

Liu DL, Kingston G (1995) Qcane: a simulation model of sugarcane growth and sugar accumulation. In: ROBERTSON, M.J. (Ed.). Research and modelling approaches to assess sugarcane production opportunities and constrains. St. Lucia: University of Queensland. 25‑29.

Marin FR, Jones JW, Royce F, Suguitani C, Donzeli JL, Pallone Filho WJ, Nassif DSP (2011) Parameterization and evaluation of predictions of DSSAT/ CANEGRO for sugarcane brazilian production systems. Agron J. 103(2):304-315.

Mkhabela MS, Mkhabela MS, Mashinini NN (2005) Early maize yield forecasting in the four agro-ecologicalregions of swaziland using NDVI data derived from NOAA's-AVHRR. Agr Forest Meteorol. 129:1-9.

Martine JF (2003) Modélisation de la production potentielle de la canne à sucre em zone tropicale, sous conditions thermiques et hydriques contrastées. Applications du modéle. 2003. 130 p. Thése (Docteur) – Institut National Agronomique Paris-Gragnon, Paris.

Moreto VB, Rolim GS (2015) Estimation of annual yield and quality of "valencia" orange related to monthly water deficiencies. Afr J Agric Res. 10(1):543-553.

Oliveira HF, Mantovani EC, Sediyama GC (2013) Avaliação de modelos de estimativa de produtividade da cana-de-açúcar irrigada em Jaíba-MG. Rev Bras Agric Irrig. 7(2):112-127.

Paulino SEP, Volpe CA (2001) Relações entre a produção de laranjeira – "Pêra" e algumas variáveis meteorológicas, em Limeira, SP. Rev Bras Frutic. 23(1):130-1333.

Rolim GS, Aparecido LEO (2015) Camargo, Köppen and Thornthwaite climate classification systems in defining climatical regions of the state of São Paulo, Brazil. Int J Climatol. 35. DOI: 10.1002/joc.4372.

Savin IY, Stathakis D, Negre T, Isaev VA (2007) Prediction of crop yields with the use of neural networks. Rus Agric Sci. 33(6):361–363.

Scarpari MS, Beauclair EGF (2009) Physiological model to estimate the maturity of sugarcane. Sci Agric. 66(5):622-628.

Thornthwaite CW (1948) An approach towards a rational classification of climate. Geographical Review. 38(1):55-94.

Thornthwaite CW, Mather JR (1955) The water balance. Centerton: drexel institute of technology, laboratory of climatology. 104p. (Publications in Climatology, v.8, n.1).

Única (2012) União da indústria de cana-de-açúcar. Available at: www.unica.com.br. (Accessed 28 march 2015).

Walpole RE, Myers RH, Myers SL, Ye KE (2001) Probability and statistics for engineers and scientists (9th Edition). 816p.