

Full Length Research Paper

Simple and multiple linear regressions for harvest prediction of Prata type bananas

Bruno Vinícius Castro Guimarães^{1*}, Sérgio Luiz Rodrigues Donato², Victor Martins Maia³, Ignacio Aspiazú³, Maria Geralda Vilela Rodrigues⁴ and Pedro Ricardo Rocha Marques²

¹Federal Institute of Education, Science and Technology of the Amazon, IFAM, Campus São Gabriel da Cachoeira, BR 307, km 03, Estrada do Aeroporto, Cachoeirinha, ZIP Code 69750-000, São Gabriel da Cachoeira, AM, Brazil.

²Federal Institute of Education, Science and Technology of Bahia – Campus Guanambi, P. O. Box 09, Ceraima District, ZIP Code 46.430-000, Guanambi, BA, Brazil.

³Montes Claros State University– Campus Janaúba, Center for Exact Sciences and Technology, Department of Agricultural Sciences, 2.630 Reinaldo Viana Av., PO Box 91, Bico da Pedra, ZIP Code 39.440-000, Janaúba, MG, Brazil.

⁴Agricultural Research Company of Minas Gerais / Regional Unit Epamig Northern Minas Gerais, ZIP Code 39525-000, Nova Porteirinha, MG, Brazil.

Accepted 21 November, 2013

This study aimed to fit regression models for harvest prediction in Prata type bananas. The experiment consisted of plants and bunches of bananas carried out in Guanambi, BA, with genotypes Dwarf Prata (AAB) and BRS Platina (AAAB), planted at a spacing of 3.0 × 2.5 m, with irrigation. Measurements of vegetative and yield characteristics were sampled at random. Models of simple and multiple linear regression were estimated, considering as independent variables with highest correlation coefficients with the masses of the bunch (MB) and hands (MH). The simple linear regression models allow prediction of the masses of the bunch and hands, according to the number of hands (NB), with better precision, for both genotypes and at least 120 days prior to harvest. For the 'Dwarf Prata', the equations were: $MCA = -10.05 + 3.08NH$; $r^2 = 0.99$, e, $MPE = -12.38 + 3.55NH$; $r^2 = 0.99$. For the 'BRS Platina' equations were: $MCA = -1.37 + 3.02NH$; $r^2 = 0.97$ e $MPE = -2.74 + 2.88NP$; $r^2 = 0.97$. The determination coefficients for the adjusted models ensure consistency of the regressions for the estimation of Prata type banana production.

Key words: Estimated harvest, Dwarf Prata (AAB), BRS Platina (AAAB) genotypes, the regression models.

INTRODUCTION

The banana (*Musa* spp.) is the largest herbaceous monocot grown in the world. Over the centuries, the crop has expanded and is now grown in over 120 countries, highlighting bananas as the most consumed fruit in the world (Cordeiro and Moreira, 2006). It is considered an important food, because of its chemical composition and

vitamins and minerals content, especially potassium. It also constitutes an important element in the diet, not only by the high nutritional value, but also by the low cost (IBRAF, 2005).

The fruit industry is among the main generators of income, employment and rural development of all world

* Corresponding author. E-mail: bvinicius20@yahoo.com.br.

agribusiness (BRASIL/MAPA, 2007). Brazil is the fifth largest producer of bananas, with 7,329 million tons produced in 503,354 ha, resulting in average yield of 14.5 t ha⁻¹ (FAO, 2011a), and the per capita consumption is 29.10 kg yr⁻¹ (FAO, 2011b). Production and consumption in Brazil show very peculiar characteristics, prevailing in most Brazilian regions the dessert varieties, such as AAB bananas, Prata type, 'Dwarf Prata', 'Pacovan' and 'Common Prata', which represent about 80%. Because of its economic importance, the banana crop is a matter of growing interest of researchers worldwide (Dantas and Soares Filho, 2006).

In these studies, typically, the researcher is interested in the identification and selection of superior genotypes that meet desirable characteristics such as appropriate size, major pests and diseases resistance and adaptation to different ecosystems. Thus, the experimental analysis addresses the biometric characteristics of plants (plant height, leaf number, pseudostem perimeter) and bunch (bunch weight, number of hands, number of fruits per bunch, fruits length and diameter) (Silva et al., 2000). These variables are quantitative, easy to measure, and may be under polygenic control, being under environmental influence, having direct and indirect economic importance (Ortiz, 1997), and most of them show significant correlation between themselves (Donato et al., 2006). This way, the analysis of plant behavior and expression is very interesting to those engaged in research in crop production, being of great application and essential to proper planning of agricultural activities. Thus, several statistical studies, linear and nonlinear models, have been developed with the purpose of obtaining future information and describe plant growth over time (Hernández et al., 2007; Maia et al., 2009).

In this context, several authors make use of predictive models to reduce or circumvent the interference of biotic or abiotic environment in the expression of the variable of economic interest (Savin et al., 2007; Zhang et al., 2007). In addition, simulation models are strategic tools for estimating the duration of stages of plant development, choosing the time of planting, predicting an abnormal production and thus use these data in genetic improvement programs (Roberto et al., 2005; Stenzel et al., 2006; Bísaro, 2007), and are, therefore, important for breeders and producers. From this perspective, there is the use of mathematical models also to forecast the harvest in several crops (Streck et al., 2007; Scarpari and Beauclair, 2009; Wyzykowski, 2009). However, mathematical modeling in the estimation of banana production is limited and has low expression in the literature. In this respect, Jaramillo (1982) and Meyer (1975) describe studies of this nature in Cavendish type cultivars in Costa Rica and Soares et al., (2013), in Brazil, adjusted models for harvest prediction of the cultivar Tropical, type Maçã. Hence, it becomes clear the importance of identifying the variables that explain part of the variation in productivity through mathematical modeling to estimate harvest, based on agronomic traits measured throughout

the crop cycle. In this case, the mathematical model, if adjusted in a functional form, allows those involved with the banana crop, whether researcher and/or producer, planning and organization of operations of pre and post-harvest. In this context, this study has the objective of adjusting simple and multiple linear regression models to predict harvest in bananas type Prata, 'Dwarf Prata' and 'BRS Platina'.

MATERIALS AND METHODS

The experiment was established in a Red-Yellow Latosol (Hapludox), medium texture, hypoxerophilous coating phase, flat to moderate topography. The experimental area is located at the Federal Institute of Bahia, Campus Guanambi, BA, 14° 13' 30" S, 42° 46' 53" W, at an altitude of 545 m, average annual precipitation of 660 mm and average temperature of 26°C. The local climate is type Aw, according to Köppen's classification. Micropropagated plantlets were used, planted in a spacing of 3.0 x 2.5 m and submitted to the system of fixed conventional sprinkler irrigation with micro-sprinklers. Installation and cultivation followed the recommendations for the crop, and fertilizers were applied based on analysis of soil and leaves. The two evaluated genotypes were: Dwarf Prata, triploid (AAB), susceptible to yellow and black Sigatoka and Panama disease, and the hybrid BRS Platina, tetraploid (AAAB), resistant to yellow Sigatoka and Panama disease, derived from the cross between 'Dwarf Prata' (AAB) and M53 diploid (AA), formerly known in pre-release as PA42-44.

Measurements were made at the time of harvest. Each plant, the basic unit, was considered as a replicate. Therefore, to assess vegetative and yield characteristics, the plants were sampled at random into the two genotypes with different numbers of replicates, 98 for 'Dwarf Prata' and 96 for 'BRS Platina'. Those measurements consisted of phenotypic vegetative descriptors plant height, pseudostem perimeter at ground level, 30 and 100 cm in height, number of live leaves at harvest. Measurements of bunch yield were also carried out: Bunch weight, number of hands and fruits per bunch, hands mass, stalk mass, length and diameter. Hands yield was also assessed: number of total fruits and fruits per bunch, and mass, internal and external length and diameter of the central fruit on the external and internal rows. The measured values were obtained according to the methodological proposal contained in catalogs of standard morphological descriptors for banana (IPGRI, 1996). For each evaluated genotype, 'Dwarf Prata' and 'BRS Platina', phenotypic correlations were estimated regarding the associations between the masses of bunch and hands with the evaluated yield and vegetative characteristics, based on the Pearson correlation (Pimentel-Gomes, 2000). The correlations were tested by Student's t test at 1% probability. Data from observations of individual replicates of each genotype were used for establishing associations between characteristics. From the estimates of the correlations between all measured variables with the masses of the bunch and hands, the significant associations and with highest values were considered to proceed the regression analysis with the subsequent choice of the best-fitted model.

To evaluate the importance of the variables related to yield and vegetative characteristics and their influence in the masses of the bunch and hands, estimations were made of multiple linear regression equations using the variable selection procedure called backward elimination (Ribeiro Júnior, 2001), in SAEG software (Statistical Analysis System), version 9.1, Federal University of Viçosa (SAEG, 2007). This way, the statistical model with k independent variables can be determined with the following multiple regression equation: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i$. In the model, Y_i refers to the response variable: mass of bunch or

hands, as a function of the regressive variables of bunch yield (number and average mass of hands; mass, length and diameter of stalk and number of fruits). It also refers to hands yield (mass and number of fruits of the hand; mass and external length of the external and internal rows fruit; internal length of the external and internal rows fruit; diameter of the external and internal rows fruit ($X_i; \dots; X_k$). The error associated to the i -esime observation is ϵ_i , assumed as normal and independently distributed; constant β_0 is inherent to the model and $\beta_1 \dots \beta_k$ model coefficients. To elucidate the relationship with each variable, given that it was significant in the regression analysis with the final bunch of the mass, Pearson parametric correlations were estimated (Pimentel-Gomes, 2000).

In the present work, the statistical procedures made to estimate the prediction equations of the values of masses of bunch and hands for Prata type bananas, 'Dwarf Prata' and 'BRS Platina', only the variables with significant correlation coefficients and with highest values were selected. Significance of the regression coefficients by the "t" test, at a 1% probability level, behavior of the biological phenomenon, determination coefficient (r^2) and the significance of the F test for the regression analysis of variance were considered for the prediction equations adjusted, in each particular case. For that, the backward elimination procedure in the software SAEG (SAEG 2007) was used, according to Ribeiro Jr. (2001). In these cases, the data from individual observations of the replicates for each genotype, 'Dwarf Prata' and 'BRS Platina', were used separately.

Simple regression models also were fitted between masses of bunch and hands for each evaluated genotype, 'Dwarf Prata' and 'BRS Platina', with the yield and vegetative characteristics. For that, regressions were fitted from means of the replicates, considering as independent variable, for each genotype, the pseudostem perimeter measured at ground level and the number of hands, because these showed highest correlation coefficients with the masses of bunch and hands. Another reason for using those variables in the models is because they are easy to determinate, by direct counting (number of hands) and by simple measure with a measuring tape (pseudostem perimeter). They are also non-destructive and can be obtained at the stage of flowering, well before the harvest of the bunch of 'Dwarf Prata' and 'BRS Platina', about 120 to 150 days (Donato et al., 2009). This enables an efficient planning associated with the technique of marking the bunch by age as a criterion for the harvest time (Lichtemberg et al., 2008; Soto Ballesteros, 2008), allows to predict the timing and amount of harvest. The number of repetitions for each case was variable, with the stratification of the number of hands on classes of observations, ranging from 9 to 13 bunches for the 'Dwarf Prata' and seven to 12 bunches for hybrid 'BRS Platina'. Similar were the proceedings for the pseudostem perimeter measured at ground level, with stratification according to the number of replicates of each class and using the average of the repetitions in the procedure for simple and multiple linear regression in the software SAEG (2007).

RESULTS AND DISCUSSION

The association between agronomic traits in banana is crucial for estimating the production of the bunch, and can be evaluated by means of phenotypic, genetic and environmental correlations (Rocha, 2010). Thus, it should be added to the study of harvest prediction of the correlation analysis, in order to determine which variables influence, to a greater or lesser degree, the production. Correlations between the masses of bunch and hands and the vegetative characteristics plant height, pseudostem perimeter at ground level, 30 and 100 cm in

height and number of live leaves at harvest, for the 'Dwarf Prata' and 'BRS Platina' bananas, were significant and positive for all variables. This directly indicates the variation of the variables yield and masses of bunch and hands, with the vegetative variables analyzed in this study (Table 1). However, the association between the characters of yield, masses of the bunch and hands, and pseudostem perimeter measured at ground level expressed more strongly in relation to other variables (Table 1). Different studies relating vegetative and reproductive characteristics in banana reported a significant correlation between the perimeter of the pseudostem and the production of the bunch (Lima Neto et al., 2003; Arantes et al., 2010). In addition, Siqueira (1984) found from clones of banana 'Prata', that among the characters related to vegetative development, the pseudostem perimeter was the most positively correlated with the characters of production, which suggests it as an effective variable to compose the harvest prediction model.

Measurements on vegetative descriptors are widely used in practice, because they are easy to determine and require simple tools. Additionally, they can be measured at flowering, about 120 to 150 days before harvesting Prata type bananas (Donato et al., 2009), which subsidizes a efficient harvest programming. Thus, given the high correlation usually found between the yield factors and the perimeter of the pseudostem at ground level, combined with ease of measurement and its non-destructive character, this variable was tested as a component of the equation to predict the masses of bunch and hands for 'Dwarf Prata' and 'BRS Platina' bananas. In this way, simple linear regression models were fitted between these variables from the means of the repetitions, taking as the independent variable the pseudostem perimeter measured at ground level (Table 2). The linear regression models were significant and positive for the variable pseudostem perimeter at ground level, both for the bunch mass and for the mass of hands, and with appropriate determination coefficient (r^2) for both genotypes (Table 2). This result suggests the possibility of planning the harvest and its implications, such as the logistics of harvest and postharvest, as well as marketing, transportation and climatization, with considerable reliability and with three to four months in advance (at flowering), by means of an easily measured indicator.

The model estimates that for every centimeter of increase in the pseudostem perimeter of 'Dwarf Prata' banana, the masses of bunch and hands increase, respectively, 327 and 289 g. In this order, the determination coefficients were adjusted in 0.76 and 0.77. Thus, the r^2 shown in Table 2 represents the fit of the data to predict the harvest, as a function of the pseudostem perimeter at ground level for both genotypes. Still, according to Table 2, it can be suggested for hybrid BRS Platina that the coefficients of the models estimate that, for each centimeter of increase

Table 1. Correlation coefficients between the masses of the bunch and hands, in association with the vegetative characteristics in type Prata bananas 'Dwarf Prata' and 'BRS Platina', Guanambi, BA, 2009.

Vegetative characteristics	Genotypes		Genotypes	
	Dwarf Prata	BRS Platina	Dwarf Prata	BRS Platina
	<i>Mass of bunch</i> **		<i>Mass of hands</i> **	
Plant height	0.63	0.49	0.63	0.47
Pseudostem perimeter at ground level	0.75	0.75	0.75	0.73
Pseudostem perimeter at 30 cm from ground	0.69	0.64	0.68	0.63
Pseudostem perimeter at 100 cm from ground	0.68	0.74	0.67	0.72
Number of leaves on the harvest	0.35	0.52	0.35	0.52

** P<0.01.

Table 2. Prediction models for the masses of bunch and hands on type Prata bananas, 'Dwarf Prata' and 'BRS Platina', as a function of the pseudostem perimeter at ground level, Guanambi, BA, 2009.

Genotype	^a Estimate (\hat{y})	Simple linear regression equation**	(r^2)
Dwarf Prata	<i>MB</i>	$\hat{y} = -10.12 + 0.33 PPGL$	0.76
Dwarf Prata	<i>MH</i>	$\hat{y} = -8.74 + 0.29 PPGL$	0.77
BRS Platina	<i>MB</i>	$\hat{y} = -14.31 + 0.36 PPGL$	0.77
BRS Platina	<i>MH</i>	$\hat{y} = -13.45 + 0.32 PPGL$	0.76

^aYield estimate (\hat{y}): *MB* = mass of bunch; *MH* = mass of hands; ^bIndependent variable (X): *PPGL* = Pseudostem perimeter at ground level. ** P<0.01.

in pseudostem perimeter, of the bunch mass and the mass of hands increase, respectively, 355 and 324 g. Subsequently, the determination coefficients were adjusted in 0.77 and 0.76. In this sense, harvest estimates for the masses of bunch and hands indicated 77 and 76%, respectively, of reliability on the final production determination.

Although the pseudostem perimeter in this study was measured at harvest time, this would present the same dimensions if it were measured at flowering, as the characteristic remains constant after that. After flowering, the banana ceases emission of roots and leaves, beginning senescence of these organs and culminating with the maturing of the bunch (Soto Ballester, 2008; Robinson and Galán, 2010). Additionally, the prediction equations were composed by variables of easy measurement and simple application in practice. Therefore, these data allow the use of non-destructive methods to determine production and productivity. However, Soares et al. (2012) in a study of banana cv. Tropical for harvest prediction by means of multiple linear regression observed that the vegetative variables plant height, pseudostem perimeter and number of live leaves at flowering had low response in the equation fit with r^2 of 0.13. For the author, this value may have been due to the little influence that each variable has on the final mass or due to the reduced number of variables that compose this model, which denotes that a number of other factors not considered in the study may influence the mass of the

bunch. However, the r^2 value found was low, although the coefficient of variation was also relatively low (CV = 16%). This is justifiable, because the associations between the bunch mass and the other characters in banana may vary between genotypes and cycles and even between hybrids and their respective parents (Donato et al., 2006; Arantes et al., 2010).

The simple linear regression equations involving masses of bunch and hands were estimated as a function of the means of the repetitions of the independent variable pseudostem perimeter measured at ground level (Table 2). The developed models express magnitudes consistent with the correlation studies and with determination coefficients values which may be used. Regarding the correlation of the plant descriptors, similar results were found by Donato et al. (2006) with respect to the character pseudostem perimeter and its relationship with the bunch mass to other genotypes, ST12-31, Grand Naine, PV42-85 and Nanicão. However, that study did not aim to establish equations for predicting the mass of the bunch, but only the relationship between the characters.

Statistical procedures for multiple linear regression were used to estimate the equation for predicting the values of the masses of bunch and hands for the 'Dwarf Prata' and 'BRS Platina' bananas. To this end, only variables that showed significant correlation coefficients and with highest values in association with the masses of bunch and hands for each genotype were selected. The

Table 3. Correlation coefficients between the masses of the bunch and hands, in association with the yield characteristics in type Prata bananas 'Dwarf Prata' and 'BRS Platina', Guanambi, BA, 2009.

Yield characteristics	Genotypes			
	Dwarf Prata	BRS Platina	Dwarf Prata	BRS Platina
	<i>Mass of bunch**</i>		<i>Mass of hands**</i>	
Number of hands	0.70	-	0.68	-
Number of fruits	0.74	-	0.72	-
Number of fruits on the fourth hand	0.51	-	0.50	-
Mass of the fruit of the external row on the fourth hand	0.73	-	0.75	-
External length of the fruit of the external row on the fourth hand	0.64	-	0.66	-
Diameter of the fruit of the external row on the fourth hand	0.41	-	0.42	-
Number of hands	-	0.59	-	0.59
Number of fruits	-	0.73	-	0.73
Number of fruits on the fifth hand	-	0.65	-	0.65
Mass of the fruit of the external row on the fifth hand	-	0.66	-	0.66
External length of the fruit of the external row on the fifth hand	-	0.57	-	0.57
Diameter of the fruit of the external row on the fifth hand	-	0.34	-	0.34

** P<0.01.

procedure called backward elimination was used for choosing the best fit of the predicting equations.

Correlations between masses of bunch and hands with the number of bunches and number of fruits in the bunch, and the characters of the fourth and fifth bunch, respectively, (number of fruits; mass, external length and diameter of the fruit in the external row), for 'Dwarf Prata' and 'BRS Platina', were significant and positive for predicting the yield of bunch and hands (Table 3). The number of hands and fruits per bunch are the components that most influence on the mass of the bunch. However, the amount of fruits and fruit diameter of the external row of the fourth bunch had the lowest correlation with the masses of the bunch and hands. Soto Ballesteros (2008) reports on the relationship between the diameter of the central fruit of the external row of the second hand and the age of bunches for harvesting, and therefore, this fruit is used as a reference for indicating the harvest time of the bunch.

The associations between the masses of bunch and hands with the mass of the fruit of the external row of the fourth hand were statistically significant, positive and with highest value in relation to other variables for 'Dwarf Prata' bananas, being respectively, 0.73 and 0.75 (Table 3), indicating that the fourth hand of the bunch is strongly related to production (Meyer, 1975). However, still with the analysis of Table 3, it appears that the yield descriptors, fruit number and mass of the fruit of the external row of the fifth hand, showed the highest correlation coefficients in relation to other variables for hybrid 'BRS Platina'. This fact accredits these variables to compose the harvest prediction model.

In summary, the variable with the highest correlation for the bunch mass and mass of the hands was the number

of fruits, for both genotypes. Despite the mass of the fruit of the external row have presented interesting correlation coefficient values for the Prata type, there was divergence for the genotypes according to the position of the hand. The fifth hand showed the second highest expression for the hybrid 'BRS Platina', while for the 'Dwarf Prata', the highest correlation was obtained for the mass of the fruit of the external row of the fourth hand (Table 3). The second hand has been considered a reference for studies and the classical procedure for determination of the harvest time for some banana genotypes (Jaramillo, 1982). Based on data from Meyer (1975) and in the present work, it can be inferred that the fourth and fifth hands, respectively, can be just as suitable for the development of postharvest studies and programming the harvest time as the second hand, which suggests the need of specific research to prove this hypothesis.

According to the correlation study, it can be suggested that the yield descriptors have the potential to compose the prediction model. In this context, the highest determination coefficients obtained for the harvest estimates for both mass of bunch and hands were observed with yield variables fruit number and mass of the fruit of the external row of the fourth hand for the progenitor 'Dwarf Prata'. For the 'BRS Platina', the number of fruits and fruit mass of the external row of the fifth hand showed the highest determination coefficients. By the methodology used, the harvest prediction equation which obtained the best fit for 'Dwarf Prata', both for the mass of the bunch, and for mass of hands, showed a determination coefficient (r^2) of 0.87 (Table 4). For the yield characteristics measured, considering the 'Dwarf Prata', the best fitted equations to determine the mass of

Table 4. Components of the equation for prediction of the masses of bunch and hands of 'Dwarf Prata' bananas, according to yield characteristics. Guanambi, BA, 2009.

Yield characteristic	Constant**	Coefficients						r ²	CV(%)
		^a NH	NFR**	NFR4	MFER4**	ELFE4	DFER4		
Mass of bunch	-18.77	-	0.14	-	0.14	-	-	0.87	22.87
Mass of hands	-16.63	-	0.12	-	0.13	-	-	0.87	23.15

^aNH, number of hands; NFR, number of fruits; NFR4, number of fruits on the fourth hand; MFER4, mass of the fruit on the external row of the fourth hand; ELFE4, external length of the fruit on the external row of the fourth hand; DFER4, diameter of the fruit on the external row of the fourth hand; -, non-significant variables by the backward elimination procedure. ** P<0.01.

Table 5. Components of the equation for prediction of the masses of bunch and hands of 'BRS Platina' bananas, as a function of yield characteristics. Guanambi, BA, 2009.

Yield characteristic	Constant**	Coefficients						r ²	CV(%)
		^a NH	NFR**	NFR4	MFER4**	ELFE4	DFER4		
Mass of bunch	-14.25	-	0.16	-	0.12	-	-	0.80	22.87
Mass of hands	-13.87	-	0.14	-	0.11	-	-	0.79	23.15

^aNH, Number of hands; NFR, number of fruits; NFR4, number of fruits on the fourth hand; MFER4, mass of the fruit on the external row of the fourth hand; ELFE4, external length of the fruit on the external row of the fourth hand; DFER4, diameter of the fruit on the external row of the fourth hand; -, non-significant variables by the backward elimination procedure. ** P<0.01.

bunch and hands at the harvest time were, respectively, $MB = -18.78 + 0.14NFR + 0.14MFER4$; and $MH = -16.63 + 0.12NFR + 0.13MFER4$. In this model, *MB* and *MH* indicate, respectively, the bunch mass and the mass of hands; *NFR* = number of fruits and *MFER4* = mass of the fruit of the external row of the fourth hand. The other variables measured at the harvest time were not significant by the adopted procedure (Table 4). The fact that the determination coefficient of this model is highest is an indication that the descriptors measured at the time of harvest can represent production more accurately. This event takes on greater significance when associated with the occurrence of appropriate coefficient of variation (Table 4).

Meyer (1975) estimated a harvest prediction equation for Cavendish type bananas. The author investigated the relations between liquid bunch mass (mass of the bunches, without the stalk) and some parameters easily measured on the harvest day. The best fit found by the author was for the mass of the middle finger of the fourth hand, which is similar to this work. The equations were: $Y = 15.30X1 + 9.84X2 + 13.55$ and $Y = 8.41Z + 4.31$, in which *Y* = net mass of the bunch, expressed in kg; *X1* = number of fingers per hand x 10⁻²; *X2* = mass of the medium finger on the fourth hand, in hectograms; *Z* = *X1* x *X2*.

Soares et al. (2012), studying the 'Tropical' cultivar banana, for harvest prediction, worked with the stepwise statistical procedure for multiple linear regression, using yield and vegetative characteristics, and determined the best fit for the prediction equation of bunch mass on harvest time, adding the variables that composed the

prediction model. The equation was $MB = -5.25 + 0.11NLH + 0.07NFB + 0.05FW + 0.18LF + 2.04RW - 0.01LS$, in which: *MB* = Bunch weight, expressed in kg; *NLH* = Number of leaves at harvest; *NFB* = Number of fruits per bunch; *AFW* = Average fruit weight; *LF*: Length fruit; *RW* = Rachis weight. In this case, the determination coefficient obtained was 0.71. However, the variables are hard to be measured or are destructive to the samples, for the average mass of the fruit and mass of the rachis, respectively, a fact that complicates or little contributes to the efficiency and practicality of the prediction process, as reason Walpole et al. (2009). When analyzing the determination coefficient, for the same harvest estimates, the hybrid BRS Platina expressed inferior values to the presented by the genitor (Table 4), however, considered adequate, being $r^2 = 0.80$ and 0.79 , respectively, for the masses of bunch and hands (Table 5). That is justifiable, because the associations between bunch mass and the other characters in banana can vary between genotypes and cycles and also between hybrids and their respective genitors (Arantes et al., 2010; Donato et al., 2006).

The coefficients of variation showed similarity between genotypes. The best equations to determine the masses of bunch and hands on the harvest time, for the 'BRS Platina', were, respectively, $MB = -14.25 + 0.16NFR + 0.12MFER5$, and $MH = -13.87 + 0.14NFR + 0.11MFER5$, in which: *BM* and *BH* indicate, respectively, the bunch mass and the mass of the hands; *NFR* = number of fruits and *MFER5* = mass of the fruit of the external row of the fifth hand. The other variables measured on harvest time were not significant by the procedure used to choose the models (Table 5). Jaramillo (1982) states that the bunch

Table 6. Prediction models of the masses of bunch and hands on type Prata bananas, 'Dwarf Prata' and 'BRS Platina', as a function of the number of hands. Guanambi, BA, 2009.

Genotype	¹ Estimate (\hat{Y})	Simple linear regression equation**	(r^2)
<i>Dwarf Prata</i>	<i>MB</i>	$\hat{Y} = -12.3804 + 3.55NH$	0.99
<i>Dwarf Prata</i>	<i>MH</i>	$\hat{Y} = -10.0485 + 3.08NH$	0.99
<i>BRS Platina</i>	<i>MB</i>	$\hat{Y} = -1.36638 + 3.02NH$	0.97
<i>BRS Platina</i>	<i>MH</i>	$\hat{Y} = -2.73380 + 2.88NH$	0.97

¹Yield estimate (\hat{Y}): *MB* = mass of bunch; *MH* = mass of hands; ²Independent variable (X): *NH* = number of hands. ** $P < 0.01$.

mass of a genotype is closely related to the number of fruits, which, in turn, is directly proportional to the number of hands. Still, this author estimated regression equations and found that the number of hands per bunch is strongly related to the bunch mass. Other authors (Fernandez-Caldas et al., 1977; Holder and Cumbs, 1982; Donato et al., 2006; Arantes et al., 2010) confirm the relationship of the components mass of hands and bunch with the number of hands.

Thus, it can be inferred that the characters of production and number of hands are directly correlated. Donato et al. (2006) found correlation coefficients of 0.94 and 0.92, respectively, for 'Dwarf Prata' and 'BRS Platina', between bunch mass and number of hands, in the first production cycle. In addition, the mass of hands correlated significantly with the number of hands, as might be expected, since the hands mass is the main component of the bunch, only without the rachis. Arantes et al. (2010) found a correlation of 0.97 and 0.98 between bunch and hands mass, respectively, with number of hands in plantains. Therefore, given the high correlation usually found between the production and the number of hands, combined with ease of measurement and its non-destructive character, this variable was tested as a component of the equation fit to predict the masses of bunch and hands for the 'Dwarf Prata' and 'BRS Platina'. In this sense, simple linear regression models were adjusted between these variables from the average of the replicates, taking as independent variable the number of hands (Table 6).

The number of hands is an easy-assessment characteristic, by simple counting, and can be obtained at the stage of flowering, well before the harvest of the bunch of 'Dwarf Prata' and 'BRS Platina', about 120 to 150 days (Donato et al., 2009). When associated with the technique of marking of the bunch by age as a criterion for harvest time (Lichtemberg et al., 2008; Soto Ballesteros, 2008), it allows to predict the timing and amount of harvest with more accuracy. The simple linear regression models that estimate the masses of the bunch and hands for the 'Dwarf Prata' and 'BRS Platina' showed significant values for the yield variable number of hands. Thus, the number of hands showed linear dependence on the harvest prediction model for the Prata

type genotypes. Thus, for the progenitor 'Dwarf Prata', both for the mass of hands as for the mass of the bunch, with the variable number of hands, a r^2 of 0.99 was fitted (Table 6). For the hybrid 'BRS Platina', the determination coefficient was 0.97 for the masses of bunch and hands too, with the variable number of hands (Table 6). In this sense, Jaramillo (1982) observed, by means of linear regression, a fit of high linear dependence ($r^2 = 0.99$) between the number of hands and the bunch mass for Cavendish type bananas. Therefore, by the simple linear regression equations it was possible to estimate with high accuracy the yield in masses of bunch and hands, for 'Dwarf Prata'. In this order, the increase rate of mass of hands was estimated at 3.55 and 3.08 kg, and the yield in mass of bunch and mass of hands, for hybrid and 'BRS Platina', presented by hand, respectively, an increase estimated at 3.2 and 2.88 kg (Table 6).

To compose the equation, it was sought the ease in obtaining the values of yield variable coupled with a high correlation coefficient with production. In this set, the fit of the data provided r^2 0.99 and 0.97, respectively, for 'Dwarf Prata' and 'BRS Platina'. This result suggests that the prediction models found in Table 6 are significant and of great scientific and practical application. Moreover, as these models use the variable number of hands of the bunch, which is defined in flowering, that is, three to four months before harvest, they allow the producer a more efficient planning.

Conclusions

The simple regression linear models estimate with a relative accuracy the masses of bunch and hands based on the pseudostem perimeter measured at ground level, for 'Dwarf Prata' and 'BRS Platina'.

The multiple linear regression models estimated with adequate accuracy the masses of the bunch and hands, as a function of the characteristics number of fruits and mass of fruit of the fourth hand for 'Dwarf Prata', and fruit number and mass of fruit of the fifth hand for 'BRS Platina'.

The simple linear regression models allow prediction of the masses of the bunch and hands, according to the

number of hands in advance of at least 120 days to harvest and high precision, for 'Dwarf Prata' ($r^2 = 0.99$) and 'BRS Platina' ($r^2 = 0.97$), which ensures the consistency of the regressions to estimate the production Prata type bananas.

ACKNOWLEDGMENT

The authors express thanks to EPAMIG - Agricultural Complex of Minas Gerais and Brazil, for the support and technological solutions provided during this research.

Abbreviations: **AAB**, Dwarf Prata, triploid; **AAAB**, BRS Platina, tetraploid; **AA**, AAB and M53 diploid; **MB**, bunch weight; **MH**, mass of hands; **NFR**, number of fruits; **MFER4**, mass of the fruit of the external row of the fourth hand; **NLH**, number of leaves at harvest; **NFB**, number of fruits per bunch; **AFW**, average fruit weight; **LF**, Length fruit; **RW**, Rachis weight.

REFERENCES

- Arantes AM, Donato SLR, Silva SO (2010). Relação entre características morfológicas e componentes de produção em plátanos. *Pesq. Agropec. Bras.* 45(2):224-227.
- Bíscaro GA (2007). Meteorologia agrícola básica. Cassilândia: UNIGRAF. P. 86.
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Política Agrícola, Instituto Interamericano de Cooperação para a Agricultura (2007). Cadeia produtiva de frutas. Série agronegócios, v. 7. Antônio Márcio Buainain e Mário Otávio Batalha (Coord.). Brasília: IICA: MAPA/SPA, 2007.
- Cordeiro ZJM, Moreira RS (2006). A bananicultura brasileira. Bananicultura: um negócio sustentável. XVII REUNIÃO INTERNACIONAL ACORBAT 2006. 15 a 20 de outubro de 2006, Joinville – Santa Catarina – Brasil. Anais... XVII Reunião Internacional da Associação para a Cooperação nas Pesquisas sobre Banana no Caribe e da América Tropical, pp. 36-47.
- Dantas JLL, Soares Filho WS (2006). Classificação, origem e evolução. Frutas do Brasil, março 2006. Available in: <www.ceinfo.cnpq.br/arquivos/artigo_2317.pdf>. Access in: may 15 2011.
- Donato SLR, Silva SO, Lucca Filho AO, Lima MB, Domingues H, Alves JS (2006). Correlações entre caracteres da planta e do cacho em bananeira (*Musa spp.*). *Ciênc. agropec.* 30:21-30.
- Donato SLR, Arantes AM, Silva SO, Cordeiro ZJM (2009). Comportamento fitotécnico da bananeira 'Prata-Anã' e de seus híbridos. *Pesq. Agropec. Bras.* 44(12):1608-1615.
- FAO-Food and Agriculture Organization (2011a). Banana. Available in: <<http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor>>. Access in: Apr. 08, 2013a.
- FAO –Food and Agriculture Organization (2011b). Consumo. Available in: <<http://faostat.fao.org/site/609/DesktopDefault.aspx?PageID=609#ancor>>. Access in: Apr. 08, 2013b.
- Fernandez-Caldas E, Garcia V, Perez-Garcia V, Diaz A (1977). Análisis foliar del plátano en dos fases de sudesarrollo: floración y corte. *Fruits* 32(11):665-671.
- Holder GD, Cumbs FA (1982). Effects of water supply during floral initiation and differentiation on female flower production by robusta banana. *Exp. Agric.* 18(2):183-193.
- Hernández MS, Martínez O, Fernández-Trujillo JP (2007). Behavior of arazá (*Eugenia stipitata* Mc Vaugh) fruit quality traits during growth, development and ripening. *Sci. Hortic.* 111:220-227.
- IBRAF - Instituto Brasileiro de Frutas (2005). Estudo da cadeia produtiva de fruticultura do estado da Bahia: Análises. São Paulo.
- IPGRI - International Plant Genetic Resources Institute (1996). Descriptors for banana (*Musa spp.*). Roma: IPGRI, P. 55.
- Jaramillo RC (1982). Las principales características morfológicas del fruto de banano, variedad Cavendish Gigante (*Musa AAA*) em Costa Rica. Upeb-Impretext, P. 42.
- Lichtemberg LA, Vilas Boas EVB, Dias MSC (2008). Colheita e pós-colheita da banana. *Inf. Agropec.* 29(245):92-110.
- Lima Neto FP, Silva SO, Flores JCO, Jesus ON, Paiva LE (2003). Relações entre caracteres de rendimento e de desenvolvimento em genótipos de bananeira. *Magistra* 15(2):275-281.
- Maia E, Siqueira DL, Silva FF, Petermelli LA, Salomão LCC (2009). Método de comparação de modelos de regressão não-lineares em bananeiras. *Cienc. Rural.* 39(5):1380-1386.
- Meyer JP (1975). Estimation de productivité: calcul dupoids des régimes de bananier em fonction Du nombre de doigts et dupoids"undoigt. *Fruits* 30(12):739-744.
- Ortiz R (1997). Morphological variation in *Musa* germplasm. *Genet. Resour. Copevol.* 44:393-404.
- Pimentel-Gomes F (2000). Curso de estatística experimental. 14.ed. Piracicaba: Nobel. P. 477.
- Ribeiro Júnior JI (2001). Análises estatísticas no SAEG. Viçosa: UFV. P. 301.
- Robinson JC, Galán SV (2010). Bananas and plantains. 2nd ed. Oxford: CAB International (Crop production science in horticulture series, 19:311).
- Roberto SR, Sato AJ, Brenner EA, Jubilei BS, Santos CE, Genta W (2005). Caracterização da fenologia e exigência térmica (graus-dia) para a uva 'Cabernet Sauvignon' em zona subtropical. *Acta. Sci. Agron.* 27(1):183-187.
- SAEG (2007). Sistema para análises estatísticas. Versão 9.1. CD-ROM. Viçosa: FUNARBE, UFV, 2007. [CD-ROM].
- Rocha J (2010). Avaliação do coeficiente de variação e relações entre caracteres de rendimento e desenvolvimento na cultura da bananeira (Magister Scientiae Dissertation). Cruz das Almas, BA: Universidade Federal do Recôncavo da Bahia. P. 46.
- Savin IY, Stathakis D, Negre T, Isaev VA (2007). Prediction of crop yields with the use of neural networks. *Rus. Agric. Sci.* 33(9):361-363.
- Silva SO, Rocha AS, Alves EJ, Credico MDI, Passos AR (2000). Caracterização morfológica e avaliação de cultivares e híbridos de bananeira. *Rev. Bras. Frutic.* 22(2):161-169.
- Stenzel NMC, Neves CSVJ, Marur CJ, Scholz MBS, Gomes JC (2006). Maturation curves and degree-days accumulation for fruits of 'Folha Murcha' orange trees. *Sci. Agric.* 63(3):219-225.
- Scarpari MS, Beaclair EGF (2009). Physiological model to estimate the maturity of sugarcane. *Sci. Agric.* 66(5): 622-628.
- Siqueira DL (1984). Variabilidade e correlações de caracteres em clones da bananeira 'Prata'. Lavras, P. 68. Magister Scientiae Dissertation – Escola Superior de Agricultura de Lavras, 1984.
- Soares JDR, Pasqual M, Lacerda WS, Silva SO, Donato SLR (2013). Utilization of artificial neural networks in the prediction of the bunches weight in banana plants. *Sci. Hort.* 155:24-29.
- Soares JDR, Pasqual M, Rodrigues FA, Lacerda WS, Donato SLR, Silva SO, Paixão CA (2012). Correlation between morphological characters and estimated bunch weight of the Tropical banana cultivar. *Afr. J. Biotechnol.* 11(47):10682-10687.
- Soto Ballester MS (2008). Bananos: Técnicas de Producción, Poscosecha y Comercialización. 3a.ed. San José, Costa Rica: Lil., 1 CD – ROM.
- Streck NA, Michelon S, Bosco LC, Lago I, Walter LC, Tellesrosa H, Paula G (2007). Soma térmica de algumas fases do ciclo de desenvolvimento da escala de counce para cultivares sul-brasileiras de arroz irrigado. *Bragantia* 66(2):357-364.
- Walpole RE, Myers RH, Myers SL, Ye K (2009). Probabilidade e estatística para engenharia e ciências. 8. Ed. Americana. ISBN 978-85-7605-199-2. São Paulo: Pearson Prentice Hall.
- Wyzykowski J (2009). Modelos de regressão para a descrição do crescimento do cafeeiro irrigado e não irrigado após recepa. 2009. P. 80. Dissertação (Master of Science in Statistics and Agricultural

Experimentation), Universidade Federal de Lavras, Lavras.
Zhang W, Bai XC, Liu G (2007). Neural network modeling of ecosystems: a case study on cabbage growth system. *Ecol. Model.* 201(3):317-325.