Short Communication

Indian Journal of Agricultural Sciences **85** (6): 859–60, June 2015/Short communication

Yield prediction in banana (*Musa* × *paradisiaca*) (cv Grand Naine) by ANN models

R VENUGOPALAN[1]

Indian Institute of Horticultural Research, Hesseraghatta Lake PO, Bangalore, Karnataka 560 089

Received: 21 March 2014; Accepted 23 March 2015

Key words: ANN, Banana, Biometrical factors, SAS

Biometrical studies are gaining paramount importance in understanding perennial crop growth cycle in view of recent climate changes and thereby to identify potential traits based on which advance selection could be attempted for a desirable trait at early crop growth stage. Also, identified significant biometrical traits across crop growth stages may be an effective component of any effective precision farming management package for increasing crop productivity. Role of biometrical factors in crop modeling is gaining importance in horticultural studies (Venugopalan 2010). To achieve this, development of statistical models based on recorded biometrical traits across crop growth stages is highly essential.

Multiple linear regression (MLR) modeling is a very powerful statistical technique and is widely used to estimate linear relationship between response variable and predictors. Its main limitation is that it is useful only when the underlying relation between response and predictor variables is assumed to be "linear". However, in a realistic situation, this assumption is rarely satisfied. Also, if there are several predictors, it is well- nigh impossible to have an idea of underlying non-linear functional relationship between response and predictor variables. To this end, in the present study, models based on the theory of "Artificial neural networks" (ANN), which can handle non-linear pattern among the independent factors (Singh and Prajneshu 2008) based on prior knowledge of the system (Hochachka *et al.* 2007), are developed and stage-wise significant yield influencing characters in banana (*Musa* × *paradisiaca* L.) (cv Grand Naine) are identified.

Primary data on yield and yield attributing traits [Plant height (cm), girth (cm), leaf length (cm), leaf breadth (cm), number of functional leaves, number of suckers] across seven important growth stages of banana (Cv Grand Naine) were collected from farmer's field (n=120) during the period 2010-12, at Dodballapur, Bengaluru. ANN models were developed to capture the inherent non-linearity among the biometrical factors vis-à-vis to identify their importance in yield forecasting at early crop growth stage.

A typical ANN consists of one input layer, one output layer and hidden layers. Each layer can have several units whose output is a function of weighted sum of their inputs. Input in to a node is a weighted sum of outputs from nodes connected to it. Thus, net input into a node is given by equation (1):

$$\text{Net input}_i = \Sigma (w_{ij} * \text{output}_j) + u_i \qquad (1)$$

The weights in an ANN, similar to coefficients in a regression model, are adjusted to solve the problem presented to ANN. Learning or training is used to describe the process of finding values of these weights. Cheng and Titterington (1994) have reviewed the ANN methodology from a statistical perspective, while Warner and Misra (1996) have laid emphasis on understanding of ANN as a statistical tool.

SAS JMP package was used for developing ANN models by dividing the data set into three parts: 70% for training (to learn pattern present in the data); 20% for validation and rest for testing (to assess the performance of trained network) by using Multi Layer Perceptron (MLP) ANN architecture. Results of models developed for three important crop growth stages showed that during 3rd month of crop growth: after sucker emergence and before inflorescence stage, plant girth (optimum value as 27 cm), and leaf breadth (optimum values as 35.5 cm); during pulp development stage(7th month), plant height (optimum value as 82 cm) and 13 leaves; during starch accumulation stage (9th month) plant height (optimum value as 127 cm); leaf breadth (optimum value as 54 cm) were the significant crop logging biometrical traits having 82.3 to 92.1% power of predicting crop yield. Optimum values were computed based on the first order derivative of the response function involving the significant variables at each stage and also the one-sigma upper limit that corresponds to 99% probability of the observations expected to fall. Further, validation of results of ANN models (Table 1) for three important growth stages with all biometrical traits and with only significant traits

[1] Principal Scientist (Agricultural Statistics) (e mail: venur@icar.org.in), Section of Economics and Statistics

Table 1 Validation of results of ANN models in banana (cv Grand Naine)

Stage	Input neurons: All biometrical traits		Input neurons: significant biometrical traits	
	RMSE/R^2 (Training set)	RMSE/R^2 (Validation set)	RMSE/R^2 (Training set)	RMSE/R^2 (Validation set)
I	1.39/83.5	1.33/81.9	1.5/72.0	1.42/77.2
II	1.5/84.6	1.36/85.5	1.6/79.9	1.56/80.2
III	1.8/85.5	1.65/88.0	1.95/81.2	1.85/84.2

I: After sucker emergence stage, II: pulp development stage, III: starch accumulation stage

substantiated the importance of the identified factors, as the R^2 values exceeded 80% under most of the cases, with least RMSE values, for both training and validation sets. Identified significant traits at a given crop stage may serve as selection indices in crop improvement research and could be used as basic inputs for formulating crop precision farming protocols with a view to achieve higher productivity.

SUMMARY

Biological organism, in general, exhibit non-linear growth in contrary to linear growth as perceived in most of the data analysis procedure. Based on the primary data collected from farmer's field during the period 2010-12, ANN models contrary to MLR models for banana (*Musa* × *paradisiaca* L.) are developed to capture effectively the inherent nonlinearity with higher R^2 values (8.7 to 11.3%). Results showed that during 3rd month of crop growth: after sucker emergence and before inflorescence stage, plant girth (optimum value as 27 cm), and leaf breadth (optimum values as 35.5 cm); during pulp development stage (7th month), plant height (optimum value as 82 cm) and 13 leaves; during starch accumulation stage (9th month) plant height (optimum value as 127 cm); leaf breadth (optimum value as 54 cm) were the significant crop logging biometrical traits having 82.3 to 92.1% power of predicting crop yield. To conclude, ANN approaches which could handle existing nonlinear relation among the biometrical traits could be utilized in crop modeling research, for realistic representation of the system of nonlinearity.

REFERENCES

Cheng B and Titterington D M. 1994. Neural networks: A review from a statistical perspective. *Statistical Science* **9**: 42–54.

Hochachka W M, Caruana, R, Fink D Munson, A Riedewald, M Sorokina D and Kelling S. 2007. Data-mining discovery of pattern and process in ecological systems. *Journal of Wildlife Management* **71**: 2 427–37.

Singh R K and Prajneshu. 2008. Artificial neural network methodology for modelling and forecasting maize crop yield. *Agricultural Economics Research Review* **21**: 5–10.

Venugopalan R. 2010. Application of statistical principles for evolving crop-logging models in banana (cv *Ney Poovan*). *Tropical Agriculture* **87**(1): pp 29–32.

Warner B and Misra M. 1996. Understanding neural networks as statistical tools. *American Statistician* **50**: 284–93.