Decision trees for knowledge discovery on the yield decline of sugarcane ratoons

Thiago da Silva Siqueira*, Felipe Ferreira Bocca, Monique P. G. de Oliveira, Luiz Henrique A. Rodrigues.

Abstract

Due to the high costs associated with planting a new sugarcane field, sugarcane ratooning is explored to decrease production costs. However, ratoons have successively smaller yields, because of the effect known as sugarcane yield decline, which can impair the profits. The factors underpinning the ratoon yield decline are yet to be established. The objective of this work is to apply decision trees to sugarcane production mill data to evaluate factors related to the sugarcane ratoon yield decline. For this, meteorological and production data from four sugarcane mills were evaluated, comparing the yield obtained with the yield of the following year.

Key words:

Potential yield, data mining, analytics.

Introduction

Sugarcane (Saccharum spp.) ratooning capacity is one of the plant's characteristics that allows its commercial use. Beyond economical aspects, preserving ratoons is important due to the potential reduction of soil operations, reducing soil degradation. However, because of the successive decline in ratoon yield, there is an economical limit to how many ratoons can be harvested.

While there are no genetic reasons to explain this decline [1], soil nutrients depletion [2] and soil structure degradation and soil compaction [3,4] have been pointed as limiting conditions to crop development. Ramburan et al. [5] indicated that exploring non-linear relations and interactions may be necessary for the study of the ratoon yield decline. The objective of the present work was to model ratoon yield decline with decision trees using data from sugarcane production at sugarcane mills associated with meteorological data.

Results and Discussion

Production data were obtained from four different sugarcane mills (Odebrecht Agroindustrial) for three seasons (2012, 2013 and 2014). The datasets contained data about harvested sugarcane, soil characteristics and inputs applied. These datasets were enriched with meteorological data obtained from remote sensing for before and after the harvest (TRMM for precipitation and Land Surface Temperature from MODIS). Samples without soil characteristics and with more than four harvests were not used.

The yield decline was measured as the ratio of the actual yield and the potential yield for that cycle, compared with the same ratio for the following year. The potential yield was obtained by simulating the water limited yields with the growth models DSSAT [6] and APSIM [7]. Three different approaches for defining the target attribute were used: based only on the yield given by DSSAT, based only on the yield given by APSIM and based on an average of the results of both models. One limitation for running those simulations was the access to meteorological data.

The ratoon yield decline was then discretized for two labels “decline” and “non-decline”. The final dataset comprised 7334 samples with 133 variables (132 independent and the target). The dataset was split into training and test sets. Tuning of the decision tree hyperparameters was performed with spatially blocked cross-validation in the training set with grid search. The resulting model achieved 0.47 accuracy in the test set, which was considered unsatisfactory, as the simple prediction with the majority class could achieve 0.51 accuracy.

Beyond the unacceptable accuracy values, the best models were trees with only one node, which made it impossible to draw conclusions regarding the variables used for the splits.

Conclusions

Modeling sugarcane ratoon yield decline with decision trees was not effective, preventing further investigation of the models obtained. Possible reasons are the lack of meteorological data for stations close to the mill areas and the quality of the production data.

Acknowledgement

We thank PIBIC/CNPq for providing the scholarship for the first author and CAPES and CNPq for the scholarships of second and third authors. This work was also partially supported by Bioen/Fapesp and Odebrecht Agroindustrial.

1 Matsuoka, S.; Stolf, R. Chapter 5 - Sugarcane tillering and ratooning: Key factors for a profitable cropping. **Sugarcane: Production, Cultivation and Uses**, p. 137–157, 2012.

2 Henry, P.C.; Ellis, R.D. Soil as a Factor in Sugarcane Ratoon Yield Decline on an Irrigated Estate in Swaziland. **Proceedings: International Society of Sugar Cane Technologists**, v. 22, n. Murdoch 1972, p. 84–91, 1995.

3 McGarry, D.; Bristow, K. Sugarcane production and soil physical decline. **Proc. Int. Soc. Sugar Cane Technolo.**, v. 24, p. 3–7, 2001.

4 Braunack, M. V.; McGarry, D. Traffic control and tillage strategies for harvesting and planting of sugarcane (Saccharum officinarum) in Australia. **Soil and Tillage Research**, v. 89, n. 1, p. 86–102, 2006.

5 Ramburan, S.; Wettergreen, T.; et al. Effects of variety ,environment and management on sugarcane ratoon yield decline. **Proc. S. Afr. Sug. Technol. Ass**, p. 180–192, 2012.

6 Jones, J.W., Hoogenboom, G., et al. The DSSAT cropping system model. In: **European Journal of Agronomy**. v.18.2003, p. 235–265.

7 Keating, B.A.; Carberry, P.S.; et al. An overview of APSIM , a model designed for farming systems simulation. **European Journal of Agronomy**, v. 18, n. 3-4, p. 267–288, 2003.