THE CONTRIBUTION OF QUALITATIVE VARIABLES SUGARCANE YIELD MODEL BASED ON SPECTRAL VEGETATION INDEX

Bernardo Friedrich Theodor Rudorff
Yosio Edemir Shimabukuro
Getulio Teixeira Batista
David Chung Liang Lee

National Institute for Space Research - INPE
Remote Sensing Division
P.O. Box 515, 12201-970 - São José dos Campos, S.P., Brazil
E-mail: bernardo@ltid.inpe.br

ABSTRACT

Ratio vegetation index (RVI) obtained from Landsat-MSS images and qualitative variables (variety and growth stage) were used in regression models to predict yield of stalks of sugarcane (Saccharum officinarum L.) in tons ha⁻¹. The RVI over the sugarcane plantation sites were acquired prior to the beginning of the sugarcane harvesting in the study area, located in Lençóis Paulista, São Paulo State, Brazil. The sugarcane selected sites were planted with two varieties (either NA56-79 or SP70-1143) at two growth stages (either first cut or ratoon). First, yield data were regressed against RVI, varieties, and growth stages, for the crop year 1984/85. This model was then used to estimate the sugarcane yield for the next crop year, i.e., 1985/86. Similarly, data from these two first crop years (i.e. 84/85 and 85/86) were used to obtain the model to predict sugarcane yield estimation for the crop year 1986/87. The model for the crop year 1985/86 explained 40% of the yield variation; while the model for the crop year 1986/87 explained 41% of the yield variation, on a field by field basis. The equivalent figures by using only RVI data in a regression model were 20% in 1985/86 and 13% in 1986/87. The use of qualitative variables significantly improved the prediction of sugarcane yield estimation based solely on the spectral response, expressed in the RVI. However, more investigation should be pursued to derive a dependable model for prediction of sugarcane yield estimation which could benefit from the combined use of spectral, agronomic, and environmental variables.

1. INTRODUCTION

Sugarcane in Brazil is planted for sugar and alcohol production. São Paulo State represents almost half of the Brazilian sugarcane production which is around 130 million tons of stalks (IBGE, 1988). The development of crop yield models to provide accurate predictions of sugarcane harvest is important for management, economical, social and political reasons. We
have previously developed a crop yield model based on linear regression techniques using a vegetation index (ratio vegetation index; RVI; Pearson and Miller, 1972) derived from Landsat MSS data to predict sugarcane yield in stalks ha\(^{-1}\) (Rudorff and Batista, 1989; 1990). In the present study we attempt to improve the estimate including qualitative variables (variety and growth stage) in the regression model.

### 2. MATERIAL AND METHOD

The Barra Grande Plant, located in Lençóis Paulista, São Paulo State was chosen to develop and test the sugarcane crop yield model. Sugarcane fields planted with the two most cultivated varieties (NA56-79 and SP70-1143) were selected inside the several farms that provide sugarcane stalks for the Barra Grande Plant. A field planted to sugarcane will grow during 12 or 18 months before stalks are cut at their base and are named — first cut. After the first cut, the ratoons are harvested at 12-months intervals over a period of 5 to 8 years depending on genetic and environmental factors. Table 1 presents the number of sugarcane fields (approximately 100 ha each) for two varieties (NA56-79 and SP70-1143) and two growth stages (first cut and ratoon) selected in each crop year. The number of selected fields were proportional to the sugarcane area cultivated with these varieties/growth stages that represent approximately 70% of the total cultivated area designated to provide sugarcane stalks for the plant. In general, first cut sugarcane plantation ranges approximately from 15 to 25% of the total cultivated area while ratoons are cultivated in the remaining area.

#### Table 1. Number of selected sugarcane fields in each crop year.

<table>
<thead>
<tr>
<th>Variety</th>
<th>Growth Stage</th>
<th>Crop Year 1984/85</th>
<th>Crop Year 1985/86</th>
<th>Crop Year 1986/87</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA56-79</td>
<td>First Cut</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>NA56-79</td>
<td>Ratoon</td>
<td>75</td>
<td>75</td>
<td>56</td>
</tr>
<tr>
<td>SP70-1143</td>
<td>First Cut</td>
<td>6</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>SP70-1143</td>
<td>Ratoon</td>
<td>14</td>
<td>22</td>
<td>33</td>
</tr>
</tbody>
</table>

Spectral data were obtained from Landsat 4 and 5 from the multispectral scanner subsystem (MSS) during the months of February and March, i.e., 2 months prior to the beginning of the harvest period. Acquisition dates for each crop year were as follow: 28 February, 1984 (crop year 1984/85); 26 March, 1985 (crop year 1984/85); and 25 February, 1986 (crop year 1986/87). Digital counts from bands 2 (red) and 4 (near infrared) were extracted for each selected sugarcane field and transformed to reflectance values (Brian and Barker, 1986) in order to minimize differences in sensor calibrations and sensor geometry (different satellites and acquisition dates). Reflectance values were transformed to the ratio vegetation index (RVI; RVI = MSS 4 / MSS 2; Pearson and Miller, 1972) which is recommended by Jackson et al. (1983) for crops that cover more than 50% of the soil background. Statistical analysis were performed using...
RESULTS AND DISCUSSION

First, a regression of observed sugarcane yield (stalks ha⁻¹) on the vegetation index (RVI) was performed in order to establish a relationship between these two variables using data from the crop year 1984/85. This model was then used to predict sugarcane yield estimation for the crop year 1985/86 and was able to explain 20% of the field by field sugarcane yield variation. Combining the data from these two crop years (1984/85 and 1985/86) a regression was performed to establish a new model that was used to predict sugarcane yield for the crop year of 1986/87. This model explained 13% of the predicted sugarcane yield variation (crop year 1986/87). Using the coefficient of determination to analyze this result it seems that the spectral data transformed to the RVI were poor estimators of the predicted sugarcane yield variation in both crop years 1985/86 ($r^2 = 0.20$) and 1986/87 ($r^2 = 0.13$).

It is well known that a first cut sugarcane is more productive than a ratoon sugarcane and differences can change between 10% and 30% (Rudorff and Batista, 1989). Also, different varieties have different production potential which is largely defined by their genetic characteristics. However, in a sugarcane production system, different varieties have to be cultivated for many reasons (e.g., maximum sucrose content at different times during the harvesting period). Again, RVI data might not be sensitive to the yield production potential of different varieties and as a consequence they present limitations and do not provide accurate sugarcane yield prediction.

The inclusion of qualitative variables or dummy variables (variety and growth stage) in the regression model significantly improved the prediction of sugarcane yield estimation as illustrated in Table 2. The explained variation of sugarcane yield using only the spectral information increased from 20% to 40% when using the qualitative variables in the prediction model for the crop year 1985/1986 and from 13% to 41% for the crop year of 1986/87. The yield prediction model for crop year 1987/88 is also presented in Table 2 and was developed using RVI, variety and growth stage data from the previous three crop years (1984/85, 1985/86 and 1986/87).

Further improvement for sugarcane yield prediction estimation, on a field by field basis, might be achieved by including soil fertility information and agrometeorological data. The result are encouraging, and further investigation is recommended especially in modeling the influence of atmospheric and topographic effects on the spectral vegetation indices.
Table 2. Sugarcane yield prediction models and their accuracy figures, where: $Y_{\text{est}}$ is the predicted sugarcane yield; VAR is variety (NA56-79 or SP70-1143), GS is growth stage (first cut or ratoon) which were converted to dummy variables, and RVI is the spectral vegetation index.

<table>
<thead>
<tr>
<th>Crop Year</th>
<th>Yield Prediction Model</th>
<th>Mean Observed Yield (ton ha$^{-1}$)</th>
<th>Stand. Error (ton ha$^{-1}$)</th>
<th>Coeff. Deterrn</th>
</tr>
</thead>
<tbody>
<tr>
<td>85/86</td>
<td>$Y_{\text{est}}=22.3+2.0^{<em>}\text{VAR}+10.4^{</em>}\text{GS}+18.2^{<em>}\text{VAR}^{</em>}\text{GS}+24.6^{*}\text{RVI}$</td>
<td>82</td>
<td>14</td>
<td>$r^2=0.40$</td>
</tr>
<tr>
<td>86/87</td>
<td>$Y_{\text{est}}=5.2+5.7^{<em>}\text{VAR}+18.7^{</em>}\text{GS}+17.8^{<em>}\text{VAR}^{</em>}\text{GS}+16.4^{*}\text{RVI}$</td>
<td>74</td>
<td>13</td>
<td>$r^2=0.41$</td>
</tr>
<tr>
<td>87/88</td>
<td>$Y_{\text{est}}=9.8+7.0^{<em>}\text{VAR}+14.5^{</em>}\text{GS}+22.0^{<em>}\text{VAR}^{</em>}\text{GS}+14.8^{*}\text{RVI}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

CONCLUSIONS

Qualitative variables or dummy variables (variety and growth stage) in the regression model significantly improved the sugarcane yield estimation. The coefficients of determination were 0.20 and 0.13, when using only spectral information (RVI), and increased to 0.40 and 0.41, when qualitative variables were included, for crop year 1985/86 and crop year 1986/87, respectively.

Further improvement for prediction of sugarcane yield estimation, on a field by field basis, might be achieved by including soil fertility information and agrometeorological data. The result are encouraging; however, further investigation is recommended especially in modeling the influence of atmospheric and topographic effects on the spectral vegetation indices.

REFERENCES


