DISCRIMINATION OF SUGARCANE CROP AND CANE YIELD ESTIMATION USING LANDSAT AND IRS RESOURCESAT SATELLITE DATA

Sandeep Pandey, N. R. Patel, Abhishek Danodia* and Ranjeet Singh

Agriculture & Soil Department, Indian Institute of Remote Sensing-ISRO, Dehradun-248001

Commission III, WG III/10

ABSTRACT:

The objective of this research work aims at crop acreage estimation at mill catchment level, derivation of sugarcane phenology and yield estimation at field level. The study was carried out in Kisan Sahkari Chini Mill catchment, Nanauta, Saharanpur, Uttar Pradesh. Extensive and systematic field sampling was carried out for ground-truth observations, biophysical measurements (LAI and above/below canopy PAR) and mill-able cane yield through crop cutting experiments. Major emphasis were laid on sugarcane crop discrimination, biophysical parameter estimation, generation of phenological metrics and yield model development for sugarcane crop at mill catchment level. Sugarcane crop discrimination and its acreage estimation was done using multi-sensor satellite data. The sugarcane classification accuracies were >92% for LISS-IV, >86% for Landsat-8 and >83% for LISS-III classified image. The sugarcane phenological matrices at field level derived using time-series of NDVI for a period of 2015-2016 through TIMESAT software. To retrieve the biophysical parameters particularly leaf area index, best predictive function developed with vegetation indices (EVI, NDVI, SAVI) through correlation and regression analysis along this cane yield estimation attempted with multi-date (eight-day) NDVI from Landsat OLI. Yield models developed for ratoon cane and planted cane explained variance in yield significantly with coefficient of determination (R^2) values equal to 0.83 and 0.69, respectively. Similar predictive functions were also established with monthly composite dataset for village-level yield estimates with step wise regression (R^2=0.83) (P=0.00001), Multi linear regression (MLR) (R^2=0.792) (P=0.00081) and Random forest regression (R^2=0.466) (P=0.038).

1. INTRODUCTION

Sugarcane (Saccharum species) is a tall-growing C4 perennial crop that is cultured in the large area of the world, specifically tropical and sub-tropical regions, between 37°N and 31°S, has high efficiency of storing solar energy as well as conversion of the light energy into sucrose in the stem most professionally (Anonymous, 2017: GOI; Grof and Campbell 2001). Sugarcane has 10-15% of fiber content at harvest time and 10-18% of sucrose content and ready to harvest at 10-24 months. Basically it is the plant of warm agroecosystem with high crop water requirement interposed with bright sunshine (Abdel-Rahman & Ahmed, 2008). The water availability is the most crucial factors among the cane variety, climate, soil conditions and availability of water (Inman-Bamber and Smith 2005).

India has an agriculture oriented economy where 58% of its population are involved in agriculture and its allied activities which contributes 17% to India's GDP while it accounts for thirteen percent of country's total export. India has second rank in production of sugarcane after the Brazil with 352.16 million tonnes of production from 4.95 million hectare area in 2015. More than 95% area of sugarcane crop is irrigated. In India, Uttar Pradesh is the highest grower of sugarcane with 2.17 million hectare of area and 145.39 million tonnes of production. It is the second largest agro-based industry after Textile (Anonymous, 2017a, Anonymous, 2013b). It is a generous source of revenue and employment for about forty five million man power (Murthy, 2010). Therefore, sugarcane industry is at prime place in the government policies; so it's monitoring and mapping became prime concern.

Remote sensing is the promising technology for growth monitoring, mapping, phenology and yield estimation of crops with a precise accuracy due to its multi-spectral, multi-temporal and synoptic coverage (Patel et al., 2008; Panigrahy et al., 2009). Crop yield assessment is also important for government as well as for decision makers to finalize the marketing, import-export and policy related to crops for good will of farmers. In India, several methods have been established to estimate the crop yield at regional scale. Navalgund et al. (1991) used IRS-

1A satellite data to define the crop acreage, crop conditioning, crop yield forecasting. Dadhwal et al. (2003) illustrated the significant variation in yield using Vegetation Index (VI) – yield empirical model at farm and district scale. Patel et al. (2006) examined two approaches viz. spectral crop growth profile and biomass production model to assess the crop yield for area of western Uttar-Pradesh. The physical model integrated Crop Simulation Model (CSM) is the best among all, but the limitation of CSM is to their location specific programming, with knowing about lack of data availability as weather, crops and soils of all locations (Radha Krishna Murthy V., 2004). Multi-spectral remote sensing datasets delivers the crop growth analysis as well as the crop phenology using the modeling of spectral crop growth profile (Badhwar, 1980) and yield (Patel et al., 2006a 2006b). However, there are not any attempt has been done to engrossment the phenology based metrics for crop yield modeling or forecasting for Indian conditions where local meteorological condition significantly affect the crop growth and yield. Thus, a better understanding with crop phenological metrics can be used for development of regress empirical model in precise estimation of yield.

In this relation present study monitoring phenology and yield estimation of sugarcane using Landsat and IRS-Resourcesat satellite data is focused on real time monitoring of sugarcane crop, extracting land surface phenological matrices. These are greatly influenced by surrounding factors like climate, topography, cultivation practices, hydrological setup and habitat that has estimated by many workers to monitor phenology of vegetation to assess the effect of climate change, few workers has used phenological matrices in yield estimation (Bolton and Friedl, 2013).

This study deals with land surface phenology monitoring, acreage estimation and yield estimation of sugarcane crop at Kisan Sahkari Chini Mill, Nanuata catchment level in Western Utter Pradesh with the following objectives: (A.) To discriminate sugarcane crop and estimate acreage at mill catchment area using multi-sensor satellite data; (B.) To

---

* Corresponding author

establish the methodology for yield prediction using phenological metrics.

2. STUDY AREA AND DATA USED

Saharanpur district is bound on the two sides i.e. in North by Dehradun and in West by Haridwar districts of Uttarakhand, on East by Ambala and Karnal districts of Haryana state and on South by Muzaffarnagar district of Uttar Pradesh. District has 3689 Sq. km area with 901 mm annual rainfall. Saharanpur has net sown area around 2.742 lakh ha out of which 2.55 lakh ha about (93%) area is irrigated. The core sector of the economy of the district remains to be agriculture, with sugarcane as the main cash crop. The other crops grown in the district include wheat, maize, rice, potato, and tomato. Mango is the main horticulture crop (Pandey, 2017).
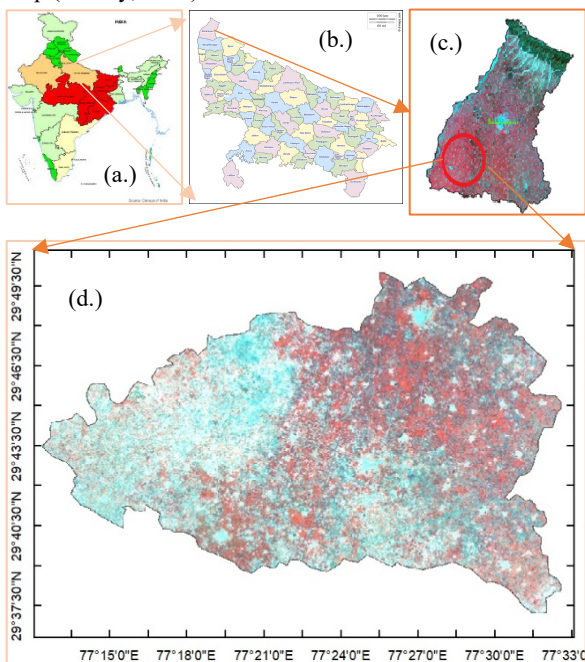

Figure 1. Location map of study area; (a.) India, (b.) Uttar-Pradesh, (c.) Saharanpur and (d.) Nanuata kisan sahkari chini mill catchment area False Color Composite (FCC) of LISS-IV (17 October 2016).

Landsat-8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) satellite data was used in this study for the period from January 2014 to April 2017. Study area falls under the overlapping zone of the satellite path and rows i.e. 146/ 40, 39 and 147/ 40, 39. Thus, Eight days (8 days) temporal resolution basis images were downloaded from USGS earth explorer (https://earthexplorer.usgs.gov). Similarly Landsat-8 also offers HLP (High Level Data Product) the products includes NDVI, SAVI, MSAVI, NDMI and EVI that were downloaded from the same website and were used for phenology metrics based yield modeling. LISS-IV IRS-P6 satellite data (Date of acquisition: Oct 2016) and LISS-III IRS-P6 satellite data (Date of acquisition: February and April 2017) were used for crop classification and acreage estimation.

Field campaigns were done multiple time to observe the crop growth and ground based information from field as well as from farmers. Leaf Area Index (LAI), Absorbed Photosynthetic Active Radiation (APAR), Fractional Absorbed Photosynthetic Active Radiation (fAPAR), soil moisture, GPS locations, Crop Cuttings Experiments (CCE) and plant/sugarcane samples were taken every time of field visit (Figure 2).

Figure 2. Field based data collections during field visit from field to farmer community.

3. METHODOLOGY

The main objective of this study was to discriminate the sugarcane plant and ratoon with reasonable accuracy and to derive phenological metrics based model for yield estimation. Crop discrimination was done using LISS-III, LISS-IV and Landsat-8 OLI satellite data while signature separability analysis was done using ground based GPS information to test the purity of pixels. Empirical relationships between biophysical variables (LAI, fAPAR) and Spectral Vegetation Indices (SVI) from satellite data were developed and compared among them on the basis of correlation coefficients.

The phenological matrices like start time, end time, length of growing period, peak NDVI, base value, peak time, peak value, amplitude, Left derivative, Right derivative, Large integral, Small integral were estimated for sugarcane plant and ratoon both (May to April) from 2014 to 2017 using eight day composite images. Landsat-8 NDVI data curve fitting approach was adopted for deriving the phenological metrics from NDVI data. The extraction of seasonality information was mainly based on "non-linear least squares fits of asymmetric logistic model functions to the time-series". This method is available in a computer program called TIMESAT (JÖnsson and Eklundh, 2004). The curve fitting method was applied to derive phenological matrices. The seasonality information was mainly based on non-linear least square fit. The files were uploaded in TIMESAT software. TIMESAT program run in MATLAB 7.5 software.

Timely and reliable crop yield prediction is an essential component of crop production forecasting system. Satellite remote sensing data for operational yield assessment at mill catchment level is a great significance in providing timely information of the crop health status, yield and other managerial aspects related to crop management, product, storage and market economics as well as crop insurance. Observation site's zonal statistics were calculated where mean NDVI were regressed against sugarcane yield of crop cutting experiment sites (CCE) sites through multiple regression analysis to develop multi date yield prediction model.

Similar more predictive function made by multi linear regression, random forest regression and step wise regression on monthly composite data set the to develop an operational model for sugarcane yield estimation and prediction sugar mill catchment area. The flow chart of methodology is depicted in figure 3 as followed in this study.
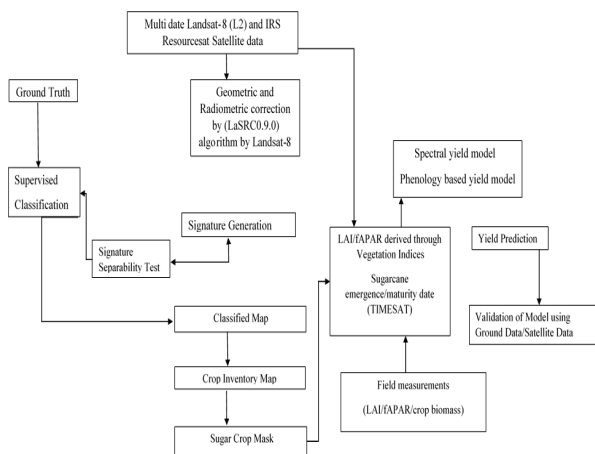
Figure 3. The flow chart of methodology for this study.

4. RESULTS AND DISCUSSIONS

Three sensors dataset with different spatial resolution have been classified and the result shown that as the spatial resolution has playing a vital role in acreage estimation. Crop discrimination has been done using supervise classification technique on LISS IV, LISS III and Landsat-8 satellite data (Fig. 4). Results reveals that LISS IV has the highest overall accuracy followed with Landsat-8 and LISS III datasets. LISS IV, LISS III and Landsat-8 classified map has an overall accuracy along kappa statistics as 85% (0.83), 83% (0.81) and 84% (0.82), respectively. There was found a significant result in sugarcane plant and ratoon acreage estimation as compared to Nanauta sugar mill observed data. There was a relative difference in crop acreage as 0.13%, 8.20% and 7.01% respectively, when remote sensing (LISS IV, LISS III and Landsat-8) based crop acreage was compared with Nanauta sugar mill observed data.

The crop yield model were developed by regression analysis using NDVI at multi-date growth stages of sugarcane. Field level yield was collected by crop cutting experiments (CCE). A positive correlation was established at different Julian days of the growth stages in sugarcane plant and ratoon with the CCE based yield. The statistics revealed that coefficient of determination for the regression model for planted cane determination coefficient (R^2=0.83) with high standard error (P value<0.05). Similarly a regression equation developed for planted sugarcane where coefficient of determination R^2 is 0.69 and P-value (P<0.05) that also infer that model can explain wide variation of data set with higher significance.

To estimate yield from remotely sensed vegetation indices like EVI, MSAVI, NDMI, NDVI and SAVI, a multi linear regression model was developed. The empirical coefficients higher for vegetation indices were used and multi linear regression was performed to establish an empirical relationship between a given spectral indices and yield. For all spectral vegetation indices used (EVI, MSAVI, NDMI, NDVI & SAVI) yield showed significantly positive response with a high coefficient of determination (R^2>0.60). The indices shown best coefficient of determination were MSAVI and NDMI particularly group of soil line vegetation indices.

The relationship developed was linear and showed (R^2=0.79) with (P<0.05) this equation can cover 79.2% of the variation in equation with a very high significance level (P-value=0.00081) (Fig. 5). The soil line based indices which minimize soil back ground effects performed better were found statistically superior over other indices explaining variability of yield in sugarcane crop.
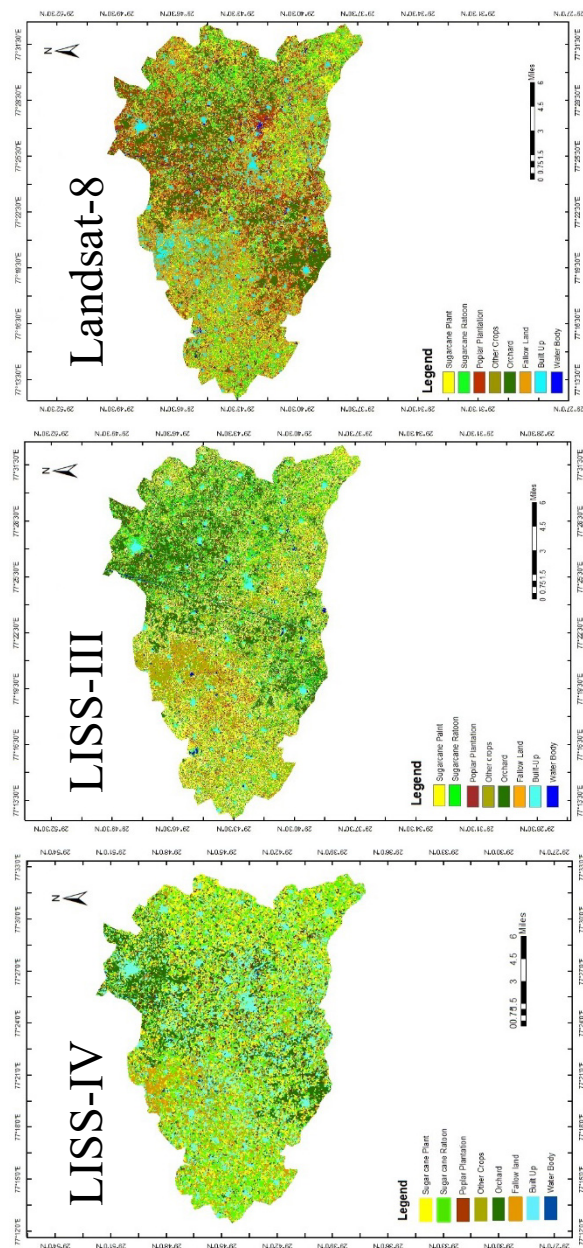

Figure 4. Supervise classified map of LISS IV, LISS II and Landsat-8.

On observing vegetation indices with yield with their different growth stages as we increases the temporal resolution of dataset the relationship became weaker than the coarser temporal resolution. Earlier worked also reported that as we use the Landsat time series data set the additive effects of the errors increases that may be the reason of lowering determination of coefficient value with increasing no. of images in observations.
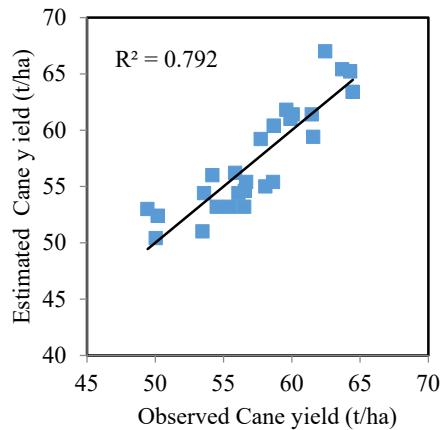
Figure 5. Observed v/s estimated yield using multiple linear regression for sugarcane.
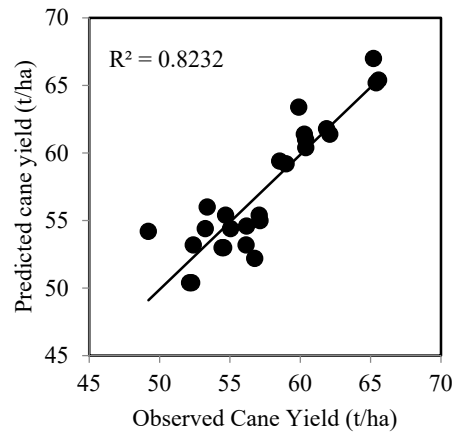

Figure 7. Observed v/s estimated yield using step wise regression based yield model.

To test the multi linear regression model these estimates of yield with best predictive functions were compared with observed yield values obtained in at independent sites (n=26). That showed determination of coefficient value (R^2=0.79), with RMSE value 4.5.

Random forest model for regression analysis was also applied on the same data set in 'R' software to develop better model of sugarcane yield prediction. The statistics found were as coefficient of determination (R^2=0.4665) and corresponding P value = 0.038 (P<0.05). Adjusted R^2 value=0.4318. Model can handle lesser range of variation in data set but with high significance level. Random forest estimated Vs Observed values showed R^2 as 0.43 and RMSE was 19.91.

Stepwise regression model was tried just to remove the contribution of non-significant variables interferences in model output. This model was developed by selection of significant seven high significantly correlated spectral vegetation indices, where the determination of coefficient values (R^2=0.63), with the help of these significant SVIs a predictive equation was developed that's statistics revealed the value of (R^2=0.82) with its very low P value as 0.00001 and adjusted R^2=0.8042 the model can cover >82% of the variable with very high significance level. A yield map was generate using the results of step-wise regression model and depicted in Figure 6. Validation of step-wise regression model showed in fig. 7.
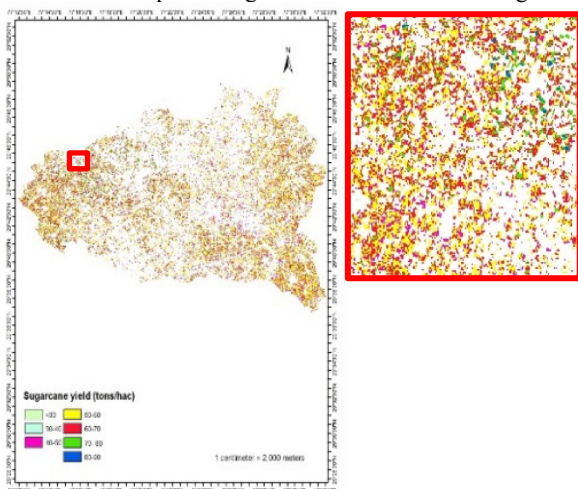
Phenological matrices provide eleven analysis outcomes i.e. SOS, EOS, LGP, base peak, peak amplitude, left derivative, right derivative, large integral and south integral. Empirical relationships between phenological matrices and yield were developed where good correlations were observed particularly with Peak NDVI value and NDVI-amplitude of ratooned and planted cane. The values showing considerable scatter of data due to varied chlorophyll content and canopy characteristics (drooping and erectile).

The crop yield model were developed by regression analysis using Peak NDVI and Amplitude value derived from a time-series of NDVI data set of eight days composite and yields observed by crop cutting experiments (CCE). In this regression equation out of 66 observations, 42 data set were trained in model development (table 1) and remaining 16 for testing purpose.

S.N.	Prediction Equation	n	R^2	SEE	P-value
1.	Ratooned cane yield tons/ha= A + B*peak NDVI + C*NDVI amplitude	26	0.78	6.0	0.0005
2.	Planted cane yield (t/ha) = A+ B *Peak NDVI- C *NDVI amplitude	16	0.75	3.4	0.00047

Table 1. Phenology based yield models and its statistical parameters (A, B & C are coefficient of equations)

This study demonstrated that with phenological matrices use, mill catchment level yield can be predicted up to 78% using landsat-8 OLI imagery. Statistical result revealed that yield prediction using phenological matrices i.e. peak NDVI and amplitude was statistically significant with high coefficient of determination (R^2=0.78) for ratoon cane and (R^2=0.75) for planted cane. The P value <0.05 showed that model can predict a significant wide variation in the data set with high level of significance. Lower value of RMSE (2.5) that predictive function can be work as a model for yield prediction of sugarcane.

The developed model needs to check its applicability for operational use. The estimates of yield with the above predictive function were compared with corresponding observed yield values at independent sites (n=14) left over for

Figure 6. Yield map of sugarcane of Nanuata Kisan Sahkari Chini Mill, Saharanpur (UP).

testing purpose. The estimated yield vales obtained by predictive functions were plotted against the observed yield values. Statistical parameters showed an excellent agreement between ground measured and predicted yield values with R2 of 0.747 and 0.562 for planted and ratoon sugarcane, respectively.

5. CONCLUSIONS

The results depicted that high spatial resolution satellite data gave quality results in classification and its accuracy. Sugarcane crop phenology results enrich more knowledge about crop growth and its spectral variation along spatial scale. Phenological metrics based crop yield model depicted qualitative regression relationships among crop phenology metrics and yield along NDVI based regression model, Multi linear regression model, Random forest based regression model and Stepwise regression model as illustrated in results section.

ACKNOWLEDGEMENT

REFERENCES

Abdel-Rahman, E. M., Ahmed, F. B., 2008. The application of remote sensing techniques to sugarcane (Saccharum spp. hybrid) production: a review of the literature. International Journal of Remote Sensing, 29:13, pp. 3753-3767. DOI: 10.1080/0143116070187460

Anonymous, 2013b. Status paper on Sugar, Directorate of Sugar Development. www.farmer.gov.in accessed on 22/6/2017.
Anonymous, 2017a. Agriculture in India: Information about Agriculture and its importance. https://www.ibef.org.
Badhwar, G. D., 1980, "Crop emergence date determination from spectral data", Photogramm. Engg. & Remote Sens., 46, pp. 369-377.

Bolton, D. K., Friedl, M. A., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agricultural and Forest Meteorology, 173, 74–84.
Dadhwal, V. K., Sehgal, V. K., Singh, R. P., Rajak, D. R., 2003. Wheat yield modelling using satellite remote sensing with weather data: Recent Indian experience. MAUSAM, 54, 1,pp, 253-262.

Duveiller, G., López, L. R., Baruth, B., 2013. Processing of 1-km spatial resolution fAPAR time series for sugarcane yield forecasting and monitoring. Remote Sensing, 5, pp. 1091–1116.Grof, C. P. L., Campbell, J. A., 2001. Sugarcane sucrose metabolism: scope for molecular manipulation. Australian Journal of Plant Physiology, 28,pp. 1–12.

Inman-Bamber, N. G., Smith, D. M., 2005. Water relations in sugarcane and response to water deficits. Field Crops Research, 92,pp. 185–202.

Jönsson, P., Eklundh, L., 2004. TIMESAT - a program for analysing time-series of satellite sensor data, Computers and Geosciences 30,pp. 833-845.

Kumar, A., Roy, P. S., 2011. Effect on Specific Crop Mapping using Worldview-2 Multi-spectral add-on bands- A Soft Classification Approach. Special Sessions on Digital Globe, pp. 1-10.

Lingle, S.E., 1999. Sugarcane. In Crop Yield, Physiology and Processes; Smith, D.L., Hamel, C., Eds.; Springer-Verlag: Berlin, Germany, pp. 287–31.

Murthy, S. R. S., 2010. Economics of Sugarcane Production and Processing. Department of Economic Analysis and Research, National Bank for Agriculture and Rural Development. Mumbai. Occasional Paper. 54, pp. 1-200.

Navalgund, R. R., Parihar, J. S., Ajai, Nageshwara Rao, P. P., 1991. Crop inventory using remotely sensed data. Current science, 61 (3 & 4), pp. 162-171.

Panigrahy, R. K., Ray, S. S. & Panigrahy, S., 2009. Study on utility of IRS-P6 SWIR band for crop discrimination and classification. Journal of Indian Society of Remote Sensing. 37, 2, pp. 325-333.

Patel, N. R., Bhattacharjee, Mohammed, A. J., Tanupriya, B., Saha, S. K., 2006b. Remote sensing of regional yield assessment of wheat in Haryana, India. International Journal of Remote Sensing, 27, 19, pp. 4071-4090.

Patel, N. R., Mohammed, A. J., Rakhesh, D., 2006a. Modeling of Wheat Yields Using Multi-temporal Terra/MODIS Satellite Data. Geocarto International, 21, 1, pp. 43-50.

Radha Krishna Murthy V., 2004. Crop growth modeling and its applications in agricultural meteorology. In proceedings: Satellite Remote Sensing and GIS Applications in Agricultural Meteorology, AGM-8, WMO/ TD-No. 1182, pp. 235-261.