UP-SCALING PADDY YIELD AT SATELLITE-FOOTPRINT SCALE USING SATELLITE DATA IN CONJUNCTION WITH CCE DATA IN SAHIBGANJ DISTRICT, JHARKHAND

Bikash Ranjan Parida [1, *], Avinash Kumar Ranjan [1]

[1] Department of Land Resource Management, School of Natural Resource Management, Central University of Jharkhand, Brambe-835205, Jharkhand, India - bikashrp@gmail.com

Commission III, WG III/10

KEYWORDS: Yield prediction, Remote sensing, NDVI, EVI, AquaCrop model, CCE data

ABSTRACT:

Agriculture plays a vital role in the economy of India as almost half of the workforce dependent on agriculture and allied activities. Rice is an important staple food and provides nutritious need for the billions of population. Mapping the spatial distribution of paddy and predicting yields at district level aggregation are crucial for food security measures. This study has utilized the time-series MODIS-based Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) data in conjunction with CCE data to derive a statistical model for up-scaling paddy yield at satellite-footprint scale over Sahibganj district in Kharif (monsoon) season 2017. The CCE data were collected from ten random paddy plots. In addition, Area, Production, and Yield (APY) data were collected during harvesting period by interacting with eighty farmers belong to eight villages. The AquaCrop model was also used to simulate the paddy yield for Kharif season. The key results showed that based on the farmers-based yield data, paddy yield was observed as ~3200 kg/hectare, whereas, NDVI and EVI-based yield models based on satellite data showed about 2,960 and 3,530 kg/hectare, respectively. Moreover, multi-regression-based yield model showed the mean yield of 3,070 kg/hectare. With respect to farmers-level yield data, the relative deviation (RD) of yield based on NDVI data was -7.5% (underestimation), while EVI was 10.31% (overestimation). The multi-regression-based yield model and AquaCrop model were underestimated by -4.06 and -10.16%, respectively. Thus, it can be inferred that the multi-regression-based yield was close to farmers-based survey yields. It can be concluded that the satellite databased yield prediction can be reliable with ±10% of RD. Nevertheless, remote sensing technology can be beneficial over traditional survey method as the satellite-based methods are cost-effective, robust, reliable, and time-saving than the traditional methods.

1. INTRODUCTION

Agriculture plays a vital role in the economy of India as almost half of the workforce dependent on agriculture and contributes nearly 18% to the national GDP. Rice (Oryza Sativa) has a significant role by providing the nutritious requirement to 1.3 billion of the Indian population. Yield prediction before harvesting using remotely sensed satellite imagery and crop models are significant steps to make a key decision on food storage, procurement, public distribution, export-import (EXIM), and national food security.

Remote Sensing (RS) techniques and crop growth simulation models have been provided a dynamic and robust way to monitor the agriculture system in the current decades. These methods are cost-effective, accurate and reliable; require less manpower and time, as compared to traditional survey methods. In the contemporary world, remotely sensed information collected from various satellite sensors (multi-spectral, hyperspectral, synthetic aperture radar) have facilitated the development of modern agriculture systems by providing a wide range of spatial data to monitor crops and to predict crop yields (Ranjan and Parida 2019; Mondal et al. 2018; Panigrahy et al. 2010). Over the years, Remote Sensing (RS) sensors have been improved the quality of data by enhancing the spatial, temporal, spectral, and radiometric resolutions. Consequently, these RS data have been utilized to extend the accuracy of crop acreage estimation as well as yield forecasting owing to synoptic coverage, a wide range of multi-temporal and multi-spectral data.

In this regard, several studies have accounted that RS satellite data have the potential for crop growth monitoring, acreage estimation, and yield prediction, as the spectral response of crops were strongly associated with crop canopy cover and biophysical parameters (Sehgal et al. 2002; Patel et al. 2006). On account of yield prediction, numerous studies have recognized a virtuous correlation between the satellite-derived Vegetation Indices (VIs) and crop yields (Nuasra et al. 2012; Son et al. 2014).

Nevertheless, crop simulation models have been frequently used across the world to monitor and simulate the crop yield. Many crop simulation models viz. AquaCrop, DSSAT, InfoCrop etc. have been extensively used in diverse climatic condition to monitor and predict the various crop yield at regional to a national scale (Greaves and Wang 2016; Aggarwal et al. 2006). Crop simulation models are also a robust and dynamic way to monitor and estimate the crop yield at various scale in diverse climatic and geographic condition. As simulation, models require a smaller number of input parameters to calibrate the models. Thereby, crop simulation models are also very cost-effective, timesaving, robust and reliable over the traditional crop survey methods.

---

* Corresponding author

The overarching objective of this research was to upscale the point-based paddy production to satellite-based footprint scale using CCE data and various satellite sensors data (i.e. MODIS-based NDVI and EVI). In addition, AquaCrop model was also performed to simulate the paddy yield during Kharif 2017 over the study site.

2. STUDY AREA

This study was carried out over Sahibganj district, situated in the eastern portion of Jharkhand state (India). It is geographically extended from 24° 42′ 30″ to 25° 21′ 27″ N (latitude) to 87° 25′ 10″ to 87° 54′ 35″ E (longitude) with ~16 m of altitude from mean sea level. This is the only district of Jharkhand state from where river Ganga is passed, and most of the accompanying regions are highly fertile and cultivable. Paddy is dominantly cultivated during the monsoon season under the rainfed condition. Except for the river Ganga, there are two more river tributaries (viz. Gumani, and Morang River) flows across the district. Wherein paddy is the primary kharif (monsoon) crops in the District, usually, paddy is sown in the starting of July month and harvested in the late November month.
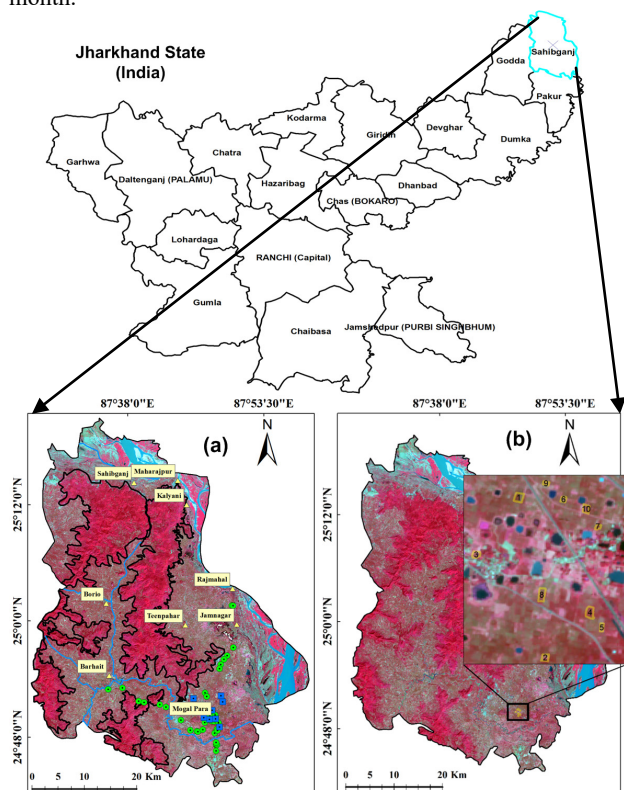

Figure 1. Location map of the study area (a) with GPS points of paddy plots (green colour) and the farmers-based APY data (cretan blue), in (b), ten CCE plots were shown.

3. MATERIALS AND METHODOLOGY

In this study, satellite data, field data (CCE-data, farmers APY data, and paddy management information), and ancillary data (weather data, soil information) were used. The characteristics of the data used have been listed in **Table 1**.

Satellite data processing i.e. yield upscaling, map layout preparation etc. were done using ArcGIS software, developed by Environmental Science Research Institute (ESRI). While the crop yield was simulated using AquaCrop model developed by the Food and Agriculture Organization (FAO).

3.1 Satellite Data used

MODIS-based MOD13Q1: The MODIS-based MOD13Q1 dataset provides 16 days composite NDVI and EVI product at 250 m of spatial resolution. The time-series (from 12th July to 3rd December 2017), MOD13Q1 product corresponding to kharif season (2018) was acquired from the Earth explorer of United State Geological Survey (USGS). The time-series NDVI and EVI data were used to identify the phenology of paddy as well as it was used to upscale the paddy production into satellite-based footprint scale.

3.2 Crop Cutting Experiment (CCE) data used

Several field visits were conducted to acquire the paddy related information. During the last of October month, a field visit was conducted to mark the 10 CCE plots (**Figure 1b**) and paddy yields were collected for yield modeling. After the paddy harvesting, yield was estimated for the corresponding ten CCE-plots. In addition, we have made an interview with eighty farmers (10 farmer's × 8 Villages) from eight different villages over the district to acquire the paddy yield information (**Figure 1a**). The villages are situated in the southern portion of the district as shown in **Figure 1a (**Cretan blue square**).** These datasets were used to derive paddy yield models.

3.3 Data processing

The MODIS-based NDVI and EVI data were originally having a sinusoidal projection, which was re-projected to geographical latitude and longitude projection as WGS 1984 using the MODIS Reprojection Tool (MRT). To up-scale the paddy yield, the paddy acreage map was used from Ranjan and Parida (2019).

Data/Sensor	Date of Acquisition	Spatial resolution (m)	Source
MODIS-based NDVI and EVI	12th Jul.-3rd Dec. 2017	250	USGS
CCE data	After paddy harvesting	-	Field Visit
Rainfall	July-Dec.	District-wise	Giovanni
Temperature (Air) Wind Speed Insolation Relative Humidity	July-Dec.	District-wise	Power NASA

Table 1. Details characteristics of data used.

3.4 Yield up-scaling

Yield up scaling was performed using the well-known regression model viz. linear regression (**Eq. 1**) and multi-regression (**Eq. 2**) analysis. We have used the ten CCE yield data with NDVI and EVI, to perform different regression

analysis. Thereby best-fitted regression equation was used upscale the CCE-based yield to satellite-based footprint scale.

$$y = a+bX \tag{1}$$
$$y = a+bX+cX_2 \tag{2}$$

Where   y = dependent variable (yield);
       a = constant (intercept);
       b and c = coefficient of independent variables (NDVI/EVI)

Relative deviation (RD) was computed using **Eq.3**. The observed yield was obtained from 80 farmers, whereas estimated yield was obtained from various up-scaling methods.

RD (%) = (Obs. Yield – Est. Yield/ Est. Yield) ×100   (3)

3.5 AquaCrop model

The AquaCrop model was performed to simulate the paddy yield by assimilating the various weather and management data. Evapotranspiration (ET_0) was estimated by the AquaCrop model, which uses the air temperature (min.-max.), humidity, solar radiation, and wind speed. The ET_0 was calculated as per the FAO Penman-Monteith method. The above-ground biomass (AGB) and yields are estimated by using **eq. 4 and 5.**

$$AGB = WP \times Tc / ETo \tag{4}$$
$$Yield = (Harvest\ index) \times (Biomass) \tag{5}$$

Where, WP = Water productivity, Tc = Canopy temperature (°C), ETo = Reference evapotranspiration (mm per unit time)

4. RESULTS

4.1 Satellite Data derived Paddy Yield

The time-series (16-days interval) MODIS-based NDVI and EVI data have been processed and the best-fitted yield models have been listed in **Table 2.** As per the literature of various study conducted across the world, best yield model was observed during the peak stage of the paddy. Multi-regression yield model has been derived by using NDVI and EVI data.

Model	Equation	r-square	p-value
NDVI	Y= 17400(NDVI) – 9326.1	0.68	<0.01
EVI	Y= 18412(EVI) – 4387.8	0.64	<0.01
Multi-regression	Y= -8754.64+12561.58(NDVI)+6957.17(EVI)	0.73	<0.01

Table 2. Satellite databased paddy yield model

The NDVI and EVI databased yield model showed that coefficient of determination (r-square) of the yield model was 0.68 and 0.64, respectively. While the R^2 was high (0.73) in case of multi-regression-based yield model. In all cases, the p-value was significant (< 0.01).

By employing the above-mentioned yield model, point-based paddy yield has been upscaled into satellite-based footprint scale. The upscaled paddy map from different yield models have been presented in **Figure 3 (a), (b), and (c).**
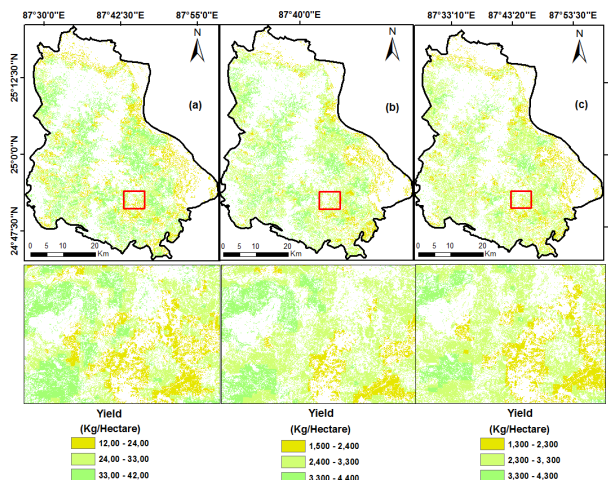

Figure 3. Upscaled paddy yield map into satellite-based footprint scale (a) NDVI model-based, (b) EVI model-based, and (c) Multi-regression model-based.

By quantifying the upscaled paddy yield map, yield range has been prudently observed. The results showed that 1200-4200, 1500-4400, and 1300-4300 kg/hectare yield have been estimated from NDVI, EVI, and multi-regression-based upscaled map, respectively. Apart from this, the mean paddy yield from the abovementioned map was noted as 2960, 3530, and 3070 kg/hectare corresponding to NDVI, EVI, and multi-regression-based upscaled map, respectively.

The mean paddy yield was estimated from the different upscaling model and their respective RD has been tabulated (**Table 3).** The mean paddy yield based on farmers-level has been observed as ~3200 kg/hectare.

Model	Mean Yield (kg/ha)	RD (%)
Farmers-level	3200	
NDVI-based model	2960	-7.50
EVI-based model	3530	10.31
Multi-regression model	3070	-4.06

Table 3: Mean paddy yield (kg/ha) estimated from different yield models and their corresponding RD

The RD as per the NDVI-model and EVI-model was about -7.5% (underestimation) and 10.31 (overestimation). It can be inferred that the RD value (4%) as per the multi-regression-based model was lowest. Hence, it can be stated that the multi-regression-based yield-upscaling model can estimate optimum yields.

4.2 Yield Simulation from AquaCrop Model

The yield has been simulated for Kharif 2017 from AquaCrop model over the Sahibganj district. The paddy yield has been estimated as 2785 kg/ha. The RD was estimated as -10.16% (underestimation).

5. DISCUSSION AND CONCLUSION

Remote sensing based crop yield predictions at various scales have been offered a cost-effective and time-saving way. In this study, we explored the coarser resolution satellite data (i.e., MODIS-based NDVI and EVI) to upscale the point-based paddy yield to satellite footprint scale in conjunction with CCE data. The AquaCrop model has been also performed to predict the paddy yield during Kharif 2017. It can be concluded that coarser resolution satellite data with limited CCE data can be crucial for the yield upscaling at satellite footprint scale in the regional level. The key findings showed that multi-regression model estimated relatively accurate paddy yields as it was very close to farmers-level yields. As per the AquaCrop, the yield was predicted as 2785 kg/ha with RD of -10% (underestimation). In this context, RD was quite large but it was reasonable. Therefore, the AquaCrop model can be also preferred in local to regional level for yield prediction.

ACKNOWLEDGMENTS

REFERENCES

Aggarwal, P.K., Banerjee, B, Daryaei, M.G., et al. 2006. InfoCrop: A dynamic simulation model for the assessment of crop yield, losses due to pest, and environmental impact of agro-ecosystem in tropical environments: II. Performance of the model. Agriculture System 89, pp. 47–67.

Greaves, G., and Yu-Min W. 2016. Assessment of FAO AquaCrop Model for Simulating Maize Growth and Productivity under Deficit Irrigation in a Tropical Environment. *Water,* 8 (12), pp. 557.

Mondal, M. et al., 2015. Simulating Yield Response of Rice to Salinity Stress with the AquaCrop Model. Environmental Science: Processes & Impacts 17 (6), pp. 1118–26.

Mondal, Saptarshi, and Chockalingam Jeganathan. 2018. "Evaluating the Performance of Multi-Class and Single-Class Classification Approaches for Mountain Agriculture Extraction Using Time-Series NDVI." *J. Indian Society of Remote Sensing,* 46 (12): 2045–55.

Nuasra, I.W., Nishio, F., Hongo, C. (2012). Rice yield estimation using Landsat ETM+ data and field observation. *J. Agricultural Sciences,* 3, pp. 45–56.

Panigrahy, Sushma, Gargi Upadhyay, Shibendu Shankar Ray, and Jai. S. Parihar. 2010. Mapping of Cropping System for the Indo-Gangetic Plain Using Multi-Date SPOT NDVI-VGT Data. *J. Indian Society of Remote Sensing,* 38 (4), pp. 627–32.

Patel N.R., Bhattacharjee B., Mohammed A.J., et al. 2006. Remote sensing of regional yield assessment of wheat in Haryana, India. *Int. J. Remote Sens.,* 27, pp. 4071–4090.

Patel, N. R., B. R. Parida, V. Venus, S. K. Saha, and V. K. Dadhwal. 2012. Analysis of Agricultural Drought Using Vegetation Temperature Condition Index (VTCI) from Terra/MODIS Satellite Data. Environmental Monitoring and Assessment 184 (12), pp.7153–63.

Ranjan A.K., Parida B.R., 2019. Paddy Acreage Mapping and Yield Estimation Using Sentinel-based Optical and SAR Sensors Data in Sahibganj district, Jharkhand (India). Spatial Information Research. DOI: 10.1007/s41324-019-00246-4

Sehgal, V.K., Rajak D.R., Chaudhary, K.N., Dadhwal, V.K. 2002. Improved regional yield prediction by crop growth monitoring system using remote sensing derived crop phenology. *In: Internat. Arch. Photogramm. Remote Sens. & Spatial Inf. Sci.* pp. 329–334

Son, N. T., Chen, C. F., Chen, C. R., Minh, V. Q., and Trung, NH. 2014. A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agric. For. Meteorol.,* 197, pp. 52–64.