ISSN 0258-7122 (Print), 2408-8293 (Online)
Bangladesh J. Agril. Res. 49(2): 105-123, June 2024

MAIZE YIELD FORECASTING MODEL FOR USING SATELLITE MULTISPECTRAL IMAGERY AT KAHAROLE IN DINAJPUR DISTRICT

N. MOHAMMAD[1], M. A. ISLAM[2], M. G. MAHBOOB[3]
M. M. RAHMAN[4] AND I. AHMED[5]

Abstract

The cultivation of maize, a high-yielding grain, has seen increased in Northern Bangladesh. Traditional crop yield prediction is costly and error-prone, often delaying its post-harvest activities. This study used remote sensing (RS) techniques for forecasting pre-harvest maize yield to improve the management system. The normalized difference vegetation index (NDVI) was widely used to predict crop yield. The study used Landsat 8(~ 30 m) and Sentinel 2A (~ 10 m) high resolution data for 2018-2019 and 2019-2020 to predict maize yield based on the year 2020-2021 at Kaharole upazila in Dinajpur district. The single cloud free image acquisition date based on maximum NDVI for both satellite images was used for each maize growing period to develop the yield prediction model. A regression model was performed between NDVI values and 20 farmers filed-level maize yields. The absolute mean error of prediction was about 10.15% for Landsat 8 and 8.82% for Sentinel 2A compared to the actual maize yield during 2020-2021. It can be concluded that NDVI data extracted from Sentinel 2A high resolution satellite images can be successfully used to predict the maize yield with appreciable accuracy. This research recommends the implementation of satellite multispectral imagery for accurate maize yield forecasting to enhance agricultural planning and resource management.

Keywords: NDVI, Landsat 8, Sentinel 2A, Maize, Prediction and Satellite image.

Introduction

Timely and accurate prediction of crop yields is necessary for effective agricultural land management, decision making and sustainability of agricultural food production (Masson-Delmotte et al., 2018). Remote sensing technology plays a vital role in the agriculture sector by providing timely and accurate information (Atzberger, 2013). Maize (*Zea mays L.*) is also known as corn, is the world's fourth major staple food crop after Rice, Wheat and Potato. Maize is initially grown for grain and secondly for fodder and raw material for industrial purposes. It is one of the important coarse cereal crops grown in the distinct agronomical conditions of Bangladesh. Maize cultivation mainly increased rapidly in the northern part of Bangladesh. It has the highest potential of per day carbohydrate productivity. During the last decades, maize cultivation in Bangladesh showed an increasing

[1,4&5]Agricultural Statistics & ICT Division, Bangladesh Agricultural Research Institute (BARI), Gazipur, [2]Department of Agricultural and Applied Statistics, Bangladesh Agricultural University (BAU), Mymensingh, [3]Forestry Unit, Bangladesh Agricultural Research Council (BARC), Dhaka, Bangladesh.

trend (1.39%). During 2011-12, total maize cultivation area and yield were 0.2 million ha and 1.3 million metric tons, respectively, while in 2019-20, these figures rose to 0.47 million ha and 4.02 million metric tons, respectively (BBS, 2020). These indicate that the prospects of maize production in Bangladesh is quite bright. However, maize growth monitoring and its yield estimation haves become a major issue of consideration.

Traditional field based crop data collection becomes clumsy, costly, time consuming, and have the possibility of error (Reynolds et al., 2000). In addition, yield estimations with conventional methods are no longer beneficial for planners as they take too much time. During the last few years, many empirical models have been developed to predict crop yield before harvesting, but many of them have become unpractical, especially those are depending on field data collection. As the satellite-based remote sensing is one of the best tools to provide vital information about the distribution of crops and its growing conditions over large areas, it can be applied for maize growth monitoring and yield forecast. The combination of data acquired by Landsat 8 and Sentinel 2 (2A, 2B) remote sensing satellites can provide a high temporal resolution (3–5 days) (Li et al., 2017) which is crucial for several applications requiring dense satellite data time series. A global equal area projection grid defined every 0.05° is used considering each sensor and combined together. The temporal observation frequency improvements afforded by sensor combination are shown to be significant. In particular, considering Landsat-8, Sentinel-2A, and Sentinel-2B together will provide a global median average revisit interval of 2.9 days, and, over a year, a global median minimum revisit interval of 14 min (±1 min) and maximum revisit interval of 7.0 days. Segarra et al., (2020) conclude that Sentinel-2 has a wide range of useful applications in agriculture, yet still with room for further improvements. Even though a combination of Landsat 8 and Sentinel 2 proposes a high frequency of observations, discrepancies in available cloud-free data, however, will still exist.

Several studies have been carried out by the correlation of normalized difference vegetation index (NDVI) with yield (LiuWT et al., 2002). Recent studies took the benefit of Landsat and Sentinel 2 data to approach crop yield forecasting at a moderate spatial resolution. For example, Lambert et al. (2018) applied Sentinel 2 data and a peak LAI (Leaf area index) procedure to predict the yields of cotton, maize, millet, and sorghum in Mali. The multiple coefficient of determination (R^2) ranged from 0.48 to 0.80 for distinct crops on training data. Lai et al. (2018) applied time-integrated Landsat NDVI for wheat yield estimation in Australia. They applied an asymmetric bell-shaped growth model to fit NDVI against time. Shakun et al. (2019) applied the combination of Landsat 8 and Sentinel 2 high frequency of observations (3–5 days) at moderate spatial resolution (10–30 m), which is important for crop yield studies that were executed for the model with near infrared (NIR) and red spectral bands and derived the area under the ROC Curve (AUC), constant, quadratic and linear coefficients of the quadratic model. The best model yielded a root mean square error (RMSE) of 0.201 t/ha (5.4%) and coefficient of determination (R^2) = 0.73 on cross-validation. Rahman et al. (2020) used the Simple Linear Machine Learning (ML) algorithm, the extracted Landsat derived the average green normalized difference vegetation index (GNDVI) values for each of the blocks were converted to Sentinel GNDVI and found strong correlations (R^2 = 0.92 to 0.99) in Bundaberg growing region of Australia. Lima et al. (2019) observed that both satellites showed the same performance in terms of accuracy for Sentinel-2 and Landsat 8, respectively. However, Landsat 8 mapped 36.9% more area of selective logging compared to Sentinel-2 data for mapping small-scale logging in the Brazilian Amazon. In some studies, predict grain yield 2–3 months prior to the harvest. More advanced regression models for yield prediction apply a time series of NDVI which allow one to obtain better forecasting (Panek et al., 2021). In Bangladesh, Bala and Islam (2009) expanded potato yield estimations models by using NDVI, LAI (leaf area index), and fraction of photosynthetically active radiation (fPAR) vegetation indices for Munshiganj District of Bangladesh by applying Moderate Resolution Imaging Spectroradiometer (MODIS) (with lowest resolution greater than 250m) 8-day composite surface reflectance data and noticed that an average error of estimation is about 15% for the study location. Islam et al. (2011) used the NDVI indicator developed from time series MODIS satellite images, the phonological growth of wheat has been monitored during the Rabi season of 2007-2008 for the greater Dinajpur area of Bangladesh. A strong correlation between the wheat production and satellite represented wheat area was found (R^2=0.71) which represents the effectiveness of the remote sensing tools for crop monitoring and production estimation. Rahman et al. (2012) applied NOAA-AVHRR data for prediction of potato yield in Bangladesh. However, a high resolution (~ 30 m) satellite image from Landsat is cost freely available since 1984. The availability of Landsat 8 images contributes an ample opportunity for long-term frequent environmental monitoring (Mandanici et al., 2016). Newton et al. (2018) improved a potato yield prediction model by applying 16-day high resolution (~ 30 m) Landsat surface reflectance data to identify the maximum normalized difference vegetation index (NDVI) value of a potato growing season in the Munshiganj district of Bangladesh. The maximum coefficient of determination (R^2) of the yield forecasting equation was found to be 0.81 between the mean NDVI and potato yield and the result revealed that the difference between predicted and actual filed yield is about 10.4%.

However, very few studies have been conducted on the relationship between high resolution (~ 30 m) Landsat 8 and Sentinel 2A (~ 10 m) satellite data and maize yield in Bangladesh. The objective of the present study is to construct a maize yield prediction model based on NDVI at Kaharole Upazila in Dinajpur district of Bangladesh respectively using high-resolution Landsat 8 Operational Land Imager (OLI) and Sentinel 2A Multi-Spectral Instrument (MSI) surface reflectance data. The combined use of high resolution Landsat 8 and Sentinel 2A images have been applied in this study to improve the yield assessment model for the maize crop.

Study Area

The study was conducted at Kaharole Upazila in Dinajpur district especially in Rabi season which was the highest maize growing areas of Bangladesh. It covered 26.36% and 14.15% maize cultivation areas in the northern part and all over Bangladesh respectively (BBS, 2020). Kaharole upazila lie between 25° 44′ to 25° 53′ N latitude and 88° 30′ to 89° 43′ E longitude respectively (Figure 1). It covers 205.54 km^2 area where 59% areas is cultivable land. The climate condition of this area is hot and humid from April to October (summer) and cool and dry from November to March (winter). This upazila receives an annual average rainfall of 1965 mm and 2417 mm, where 90% of rainfall occurs between May and October. In Kaharole upazila, the maximum and minimum mean temperature during the winter varies between 23.6 to 16.8 and 24 to 16.8°C, respectively. During summer, the maximum and minimum mean temperatures vary between 33.2 to 26.0°C and 29.8 to 25.6°C respectively. The average monthly minimum and maximum humidity varies from 68 to 86%, where the maximum humidity during the summer and the minimum humidity observed during the dry season. The agricultural pattern of this area is categorized by two growing seasons, *Rabi* and *Kharif*. *Rabi* is the main growing season, which is dominated by maize and wheat that starts in late-October or early November and ends in April. Again, *Kharif* is dominated by rice and jute, which starts in May and ends in September. Other food crops which include potato, pepper, onion, pulses, sugarcane, and oilseed are also cultivated in the study areas. The soil condition of this area is prevailed by non-calcareous brown floodplain soils and grey floodplain soils.

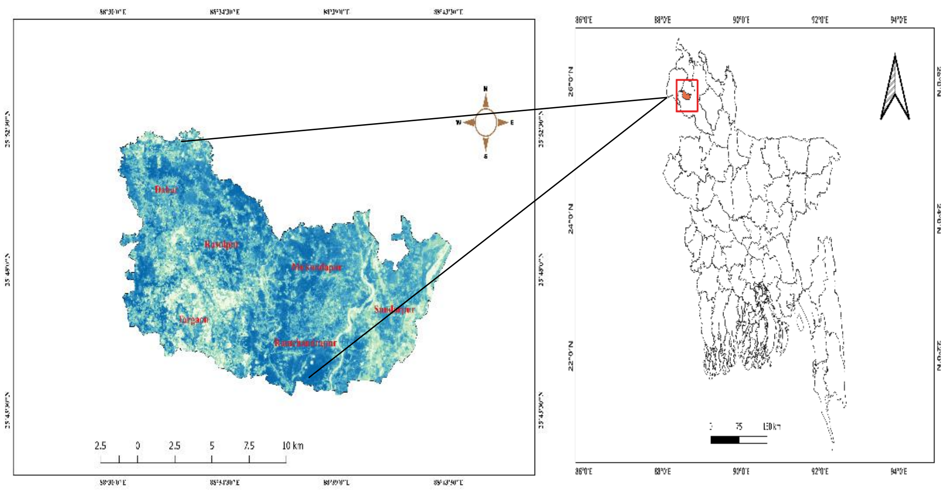


Fig. 1. Map of the study location Kaharole Upazila Dinajpur.

Materials and Methods

Yield data collection from farmer's fields

Twenty maize fields were selected from Kaharole Upazila of Dinajpur district for the three maize growing seasons 2018-2019, 2019-2020 and 2020-2021 with the agreement of the farmers (Figure 2). Data were collected from those 20 maize fields for each season. Crop information data such as field GPS locations, planting and harvesting time and yield were collected from this selected Upazilas farmers' fields.

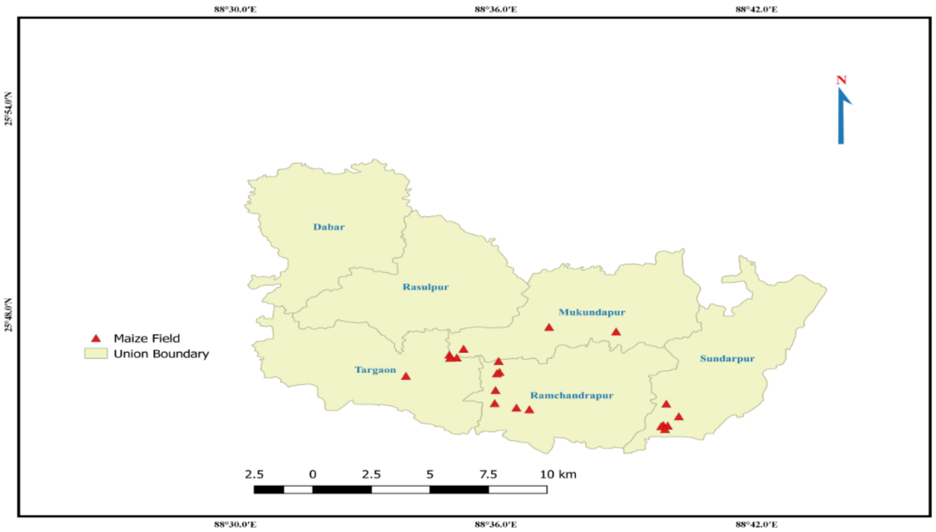


Fig. 2. Location of selected 20 maize fields (red triangles) over Kaharole, Dinajpur district

Landsat-8/OLI and Sentinel-2A /MSI Datasets

Landsat 8 images (OLI) were obtained from the United States Geological Survey (USGS) Earth Explorer website (http:// earthexplorer.usgs.gov/). Landsat 8 (OLI) is a sun-synchronous satellite staying at an altitude of 705 km above the earth with a 16-day repeat cycle. Landsat 8 has two types of sensors, especially the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS). The OLI sensor equips nine spectral bands, including a pan band, and TIRS produces two spectral bands. Sentinel-2A images (MSI) were obtained from the European Space Agency (ESA) Copernicus portal (https://scihub.copernicus.eu). Sentinel-2A carries a multispectral instrument (MSI). Images furnished by Sentinel-2A are publically available for free and have 13 spectral bands with a spatial resolution ranging from 10 m to 60 m (depending on the band) and a current temporal resolution of about 10 days (depending on the latitude). A total of 6 maximum cloud-free images were downloaded: 3 images from Landsat 8 OLI and the remaining 3 images from Sentinel 2A MSI satellite data for Kaharole Upazila in

Dinajpur district, covering the consecutive years 2018-19, 2019-20, and 2020-21, respectively (Table 1). The single date of image acquisition based on maximum greenness was used for each growing period i.e. 2018-2021 for maize cultivation. The maize sowing date was considered to be the last week of November and the first week of December for each growing season 2018-19, 2019-20 and 2020-21 respectively for the entire study site based on the information taken from the location visits. Every single image was calculated from the starting day of the plantation. The dates of image acquisition of Landsat 8/OLI and Sentinel 2A/MSI for this study are presented in Table 1.

Table 1. Model development using Landsat 8 and Sentinel 2A satellite image at Kaharole Upazila, Dinajpur district

Satellite images	Landsat 8			Sentinel 2A		
Growing Season	2018-19	2019-20	2020-21	2018-19	2019-20	2020-21
IAD	10/03/2019	12/03/2020	15/03/2021	16/03/2019	20/03/2020	05/03/2021
DAP	95	97	103	101	105	93

IAD=Image acquisition date; DAP= Days after plantation

Satellite Image Pre-processing

For Landsat data, raw digital numbers (DN) were adjusted to top-of-atmosphere (TOA) reflectance values following reference (Simonetti et al., 2015). Two techniques were used to preprocess the satellite images: (1) radiometric calibration and (2) atmospheric correction. Remote sensing data adopted from satellite sensors are influenced by several factors, such as atmospheric scattering and absorption, sensor-target-illumination geometry, sensor calibration, and data processing procedures (Teillet, 1986). For that, radiometric calibration is needed. Radiometric calibration means a set of correction techniques that are associated with the correction of the sensitivity of satellite sensors, topography and sun angle, atmospheric scattering, and absorption (Kim et al., 1990). The radiometric calibration was done by transforming the digital numbers (DNs) to surface reflectance by radiance conversion. Open source-based Quantum Geographic Information System (QGIS) software 2.18.13 version allows a plugin, which gives a tools for atmospheric correction, known as dark object subtraction (DOS-1) level 1. In this study, this tool was used in the radiometrically calibrated images to the minimize atmospheric scattering effect. DOS-1 searches each pixel of a band to find the darkest value. The scattering is eliminated by subtracting this value from every pixel in the band.

Two approaches were followed to download and process Sentinel-2A imagery. The first was a simplified process for farmers and advisors to monitor the crop status during the season. The free and open-source QGIS 2.18.13 version software together with the Semi-automatic Classification Plugin (SCP) was used for those purposes (Congedo, 2016). The advantage of using the SCP is that the user can preview and download per date and tile single bands and correct the Sentinel-2 images in the same interface. Afterward, the vegetation indices can be computed, stored and compared with other dates within the same QGIS environment. The only limitation is that in the conversion of top-of-atmosphere (TOA) reflectance values into bottom-of-atmosphere (BOA) the image-based Dark Object Subtraction (DOS-1) technique is applied. This process is less accurate than the physically-based correction that could be applied to Sentinel-2 images using the attached metadata (Congedo, 2016).

Normalized Difference Vegetation Index (NDVI)

Prediction and estimation of yield are closely associated with the capability of identifying crop species and certain agronomic parameters, such as maturity, density, vigor and disease, which can be used as yield indicators. Remote sensing can arrange these types of information to a great extent. There are distinct types of vegetation indices (VIs) generated from different spectral reflectances that are specially used to get these types of information. The NDVI is generally applied extensively around the world to monitor the vegetation quality, growth, and distribution over a large area. Differences in the phenological growth stages of different plants are reflected in the temporal NDVI profiles, since NDVI can measure growth conditions (greenness of vegetation) (Belgiu and Csillik, 2018; Croitoru et al., 2012). It is a dimensionless index, which is performed from the ratio between the surface reflectance of the NIR and RED bands of the spectrum as follows (Equation 1) (Rouse et al., 1974).

$$NDVI= NIR-RED/NIR+RED \qquad (1)$$

Where RED (Visible red) and NIR (Near infrared) are reflectance measurements for RED and NIR bands, respectively. Here for Landsat 8/OLI, band 4 and band 5 represented RED and NIR bands and for Sentinel 2A/MSI, band 4 and band 8 represented Red and NIR bands (Tucker, 1979). Factors like strong reflectance in NIR and strong absorption in Visible Red of specific vegetation distinguish the vegetation from bare soil. NDVI for a given pixel can always output in a number that ranges from - 1 to +1; however, for natural surfaces NDVI values are within the 0 to +1 range. Negative values of NDVI i.e. values approaching -1 correspond to water. An NDVI close to 0 corresponds to no vegetation, while Values between -0.1 to 0.1 generally correspond to barren areas of rock, sand or snow.

Maize yield estimation by satellite-based remote sensing technique

The red band and NIR of the calibrated images were selected from each dataset and exported into QGIS 2.18.13. A simple raster calculation was done by QGIS 2.18.13 using Equation 1 to find the NDVI images. Finally, the NDVI images were masked using the shape file from the study area. The field points of the location were imported, and the mean NDVI values for each point were extracted from the satellite image considering a 3 × 3 matrix surrounded by each point on the image.

The relationship between NDVI and the maize growing period was established by plotting the respective values in terms of single days from the start date of the maize plantation to the harvesting period. The day of the maximum NDVI was selected from their relationship with crop yield. To establish this relationship, NDVI data from the growing season 2018-2020 were used. Then, a total of four satellite images viz., each image was collected from Landsat 8 as well as Sentinel 2A from Kaharole Upazila, Dinajpur district depending on the date of the maximum NDVI were selected from two growing seasons, namely 2018-2019 and 2019-2020 to build a relationship between the NDVI values and field level maize yield. This relationship based on the each farmer's field point NDVI values was validated using the satellite image of the 2020-2021 growing season. NDVI values less than 0.25 and more than 0.95 were removed from the listed fields to reduce the influence of reflectance of other objects like bare soil, settlements, water bodies, non-agricultural crops, and infrastructure.

Yield prediction model

The final step is to determine the relationship between NDVI and maize yield from farmers' fields with the equation below:

$$y = f(x) \tag{2}$$

Where y and x are maize yield data collected from farmers' field and NDVI, respectively. The relationship between NDVI and crop-like maize yield has been observed through the linear regression model, where the response variable is denoted by maize yields and the explanatory variables by NDVIs. Several studies applied a linear regression model to describe the relationship between NDVI and crop (wheat) yield in distinct locations (Ren et al., 2008). To develop the maize yield estimation model for both fields, the data of maize yield and Landsat 8 (OLI) and Sentinel 2A (MSI) images were used for 2018-2021.

Results and Discussion

Maize yield from farmers' field and corresponding NDVI values for different locations

Maize yield data and NDVI values from different date satellite images of Landsat 8 and Sentinel 2A for corresponding farmers' fields have been collected from Kaharole Upazila during the maize growing season 2018-19 and 2019-20 respectively. Twenty farmers' field yield data collected from study area and their corresponding NDVI values for two satellite images have been presented in Figure 3 and 4 for consecutive maize growing seasons 2018-19 and 2019-20 respectively.

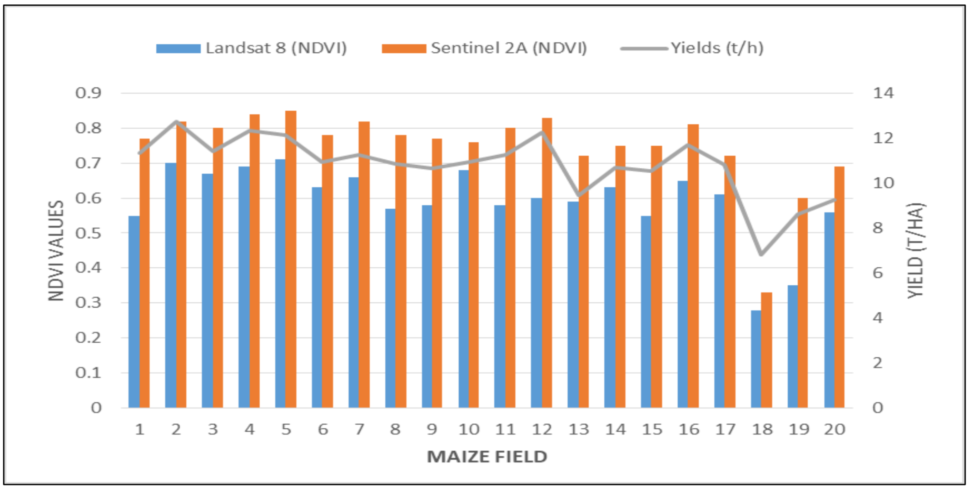


Fig. 3. NDVI values of satellite images and yields of corresponding Farmer's fields at Kaharole, Dinajpur during the season of 2018-19

Figure 3 shows that NDVI values are 0.71 and 0.85 which were the highest for Landsat 8 and Sentinel 2A and the yield was 12.12 t/ha for farmer's field 5 but the yield is a maximum of 12.72 t/ha for field 2; NDVI values are 0.28 and 0.33 that were the lowest and yield was 6.8 t/ha for farmer's field 18 during 2018-19.

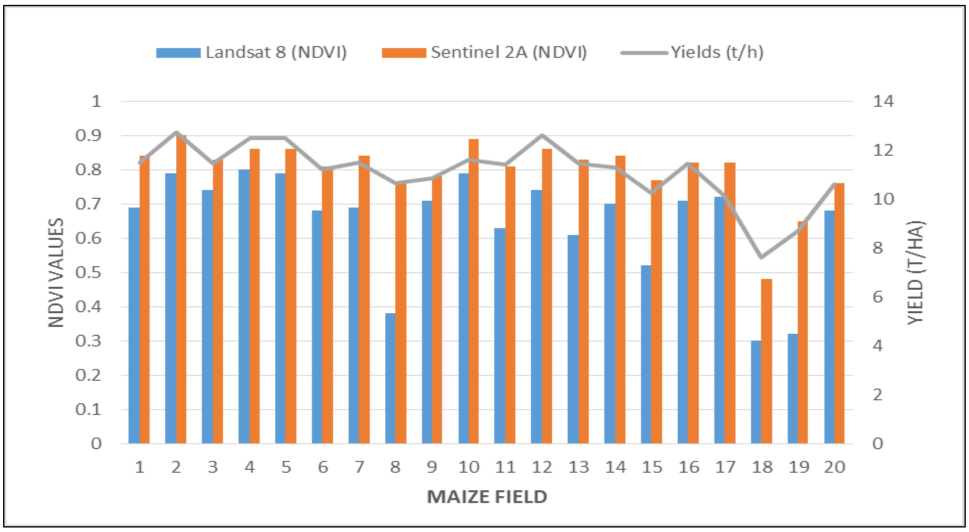


Fig. 4. NDVI values of satellite images and yields of corresponding Farmer's fields at Kaharole, Dinajpur during the season of 2019-20

Maximum NDVI for Kaharole were 0.80 and 0.90 for farmer's field 4 and 2 and minimum NDVI were 0.30 and 0.48 for Landsat 8 and Sentinel 2A for farmer's field 18, but the highest and lowest yield were 12.74 t/ha and 7.6 t/ha for farmer's field 2 and 18, respectively during 2019-20 in Figure 4.

Regression analysis of the NDVI values over the field locations

A total of four satellite images (2 images each for Landsat 8 and Sentinel 2A) from two growing seasons during 2018-2019 and 2019-2020 were selected. Based on available images, those sowing the maximum NDVI in each growing season were found 95th, and 97th days after plantation for Kaharole Upazila from Landsat 8 images as well as 101th, and 105th days after plantation for Kaharole upazila from Sentinel 2A images for 2018-2019 and 2019-2020 growing seasons, respectively. The spatial distribution of the NDVI varies from year to year. Spatial distribution of the NDVI over the selected location for selected distinct satellite images against different growing seasons is presented in Figure 5.

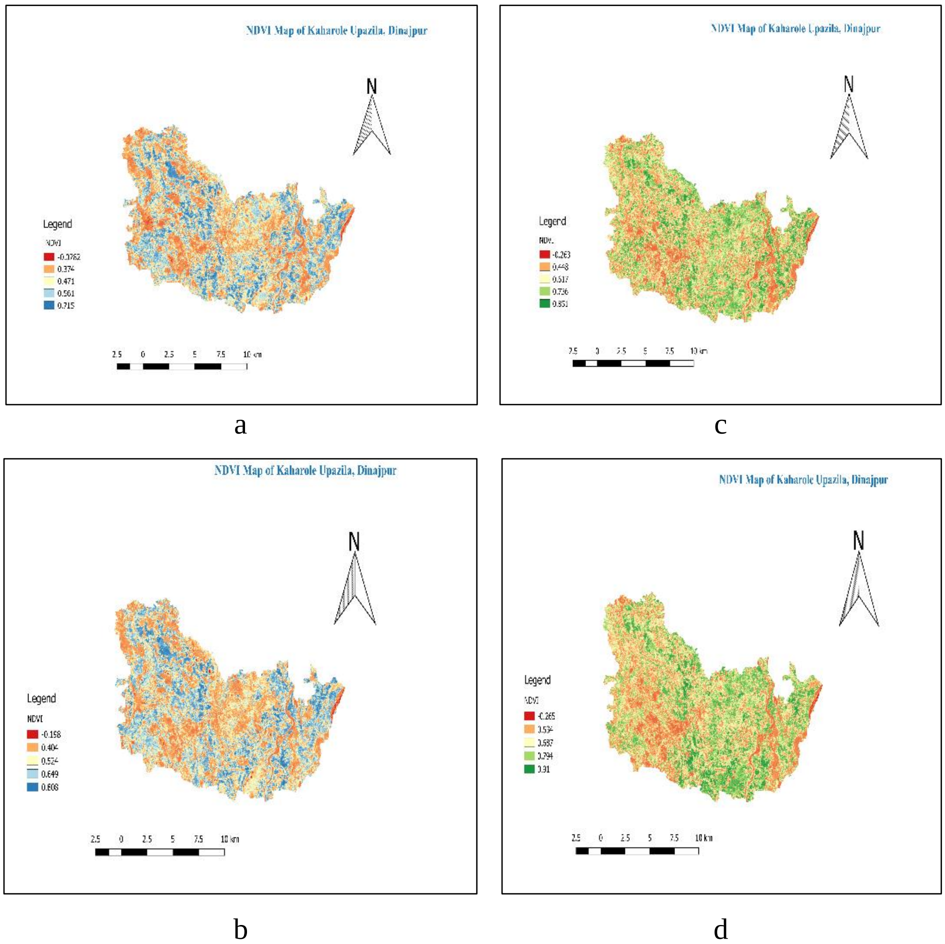


Fig. 5. Spatial distribution of the NDVI for different satellite images during the growing season 2018-19 and 2019-20. a. 95th days after plantation during 2018-19; b. 97th days after plantation during 2019-20 for Landsat 8; c. 101th days after plantation during 2018-19; d. 105th days after plantation during 2019-20 for Sentinel 2A at Kaharole Upazila.

For Landsat 8 data, NDVI distribution was maximum during 2019-2020 and distribution was minimum during the season 2019-20 at Kaharole Upazila in Figure 3. On the other hand, for Sentinel 2A data, NDVI distribution was maximum during the season 2019-2020 and minimum during the season 2019-20 at Kaharole Upazila in Figure 5. The NDVI distribution from different locations of Sentinel 2A data is outperformed Landsat 8 data during the maize growing season 2018-2019 and 2019-2020, respectively in Figure 5.

For two satellites viz., Landsat 8 and Sentinel 2A, the NDVI values and their corresponding yields for twenty farmers' maize field from different locations i.e., Kaharole Upazila during individual maize growing season 2018-2019 and 2019-2020, respectively are shown in Figure 3 and 4. From Figure 3 and 4, the mean NDVI and mean yield for two satellite data, Landsat 8 and Sentinel 2A for this location during the combined season 2018-2019 and 2019-2020 respectively were calculated which are represented below in Figure 6. Mean NDVI and mean yield are calculated for two satellite images for combined two season because maize yield in each season i.e., 2018-19 and 2019-20 is mostly the same for this location. Mean NDVI and mean yield of combined maize season performed better than NDVI and its corresponding yield of individual maize season for each satellite image in this location because according to best model criteria viz. Multiple determination of coefficient (R^2), Mean Absolute Percentage Error **(**MAPE), Normalized Root Mean Square Error **(N**RMSE) and Root Mean Square Error **(**RMSE) fitted best for mean NDVI and mean yield of the combined maize season rather than single maize growing seasons like 2018-19 and 2019-20 (Figure 5). Mean NDVI is the largest which is 0.75 for farmers' field 2, 4 and 5 as well as the smallest is 0.29 for farmers' field 18 for Landsat 8 and for Sentinel 2A , the largest and lowest are 0.86 and 0.41 for farmers' field 2, 5 and 18 for Kaharole Upazila for the combined year respectively. The highest and lowest mean yields of the two satellite data is 12.73 (t/ha) and 7.2 (t/ha) for this Upazila for the combined year respectively in Figure 6.

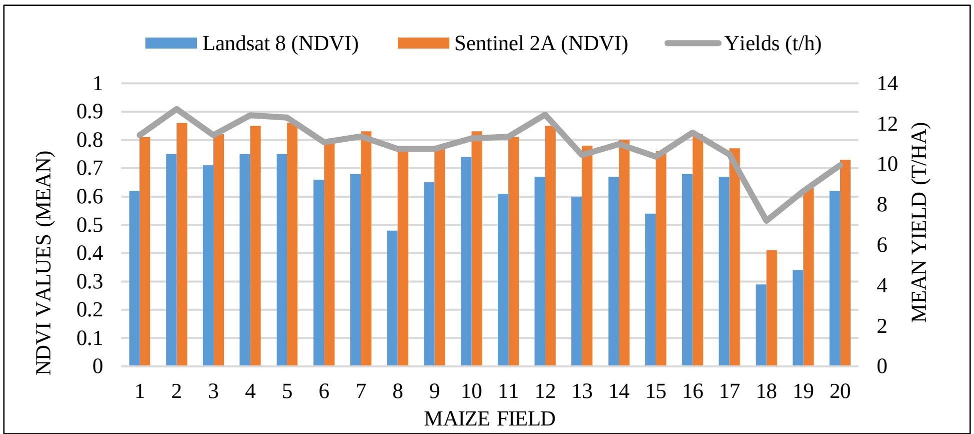


Fig. 6. Mean NDVI values for two satellite images and corresponding mean yields of Farmer's Maize Fields at Kaharole, Dinajpur during the combined season of 2018-19 and 2019-20

Maize yield and NDVI relationship using regression model

Regression analysis of maize yield against the single season basis NDVI and combined season basis mean NDVI for Kaharole was performed for two satellite images, Landsat 8 and Sentinel 2A, and are graphically presented in Figure 7.

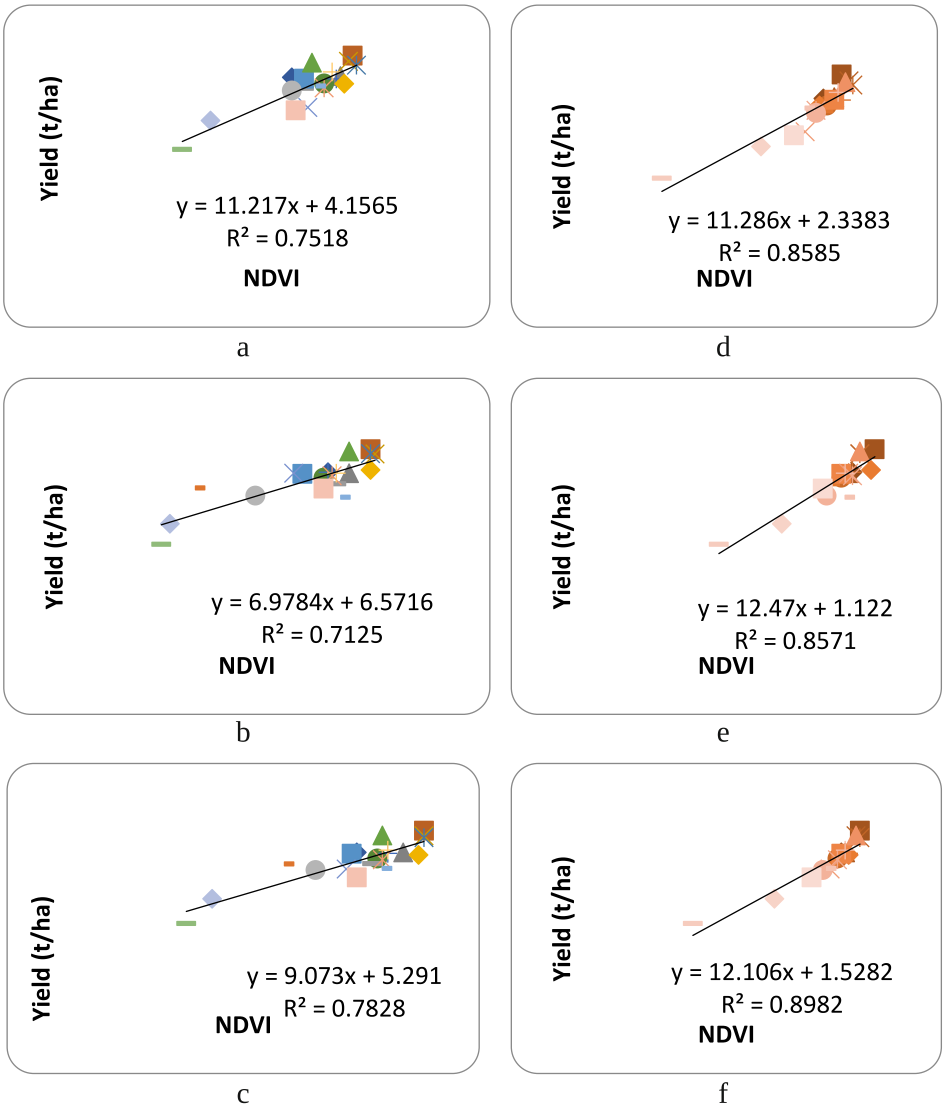


Fig. 7. Yield prediction model established from regression analysis between yield data collected from 20 farmers' maize fields for different images; [a. yield vs NDVI, 2018-19 b. yield vs NDVI, 2019-20 c. yield vs NDVI (mean), combined 2018-19 and 2019-20 for Landsat 8.] ; [d. yield vs NDVI, 2018-19 e. yield vs NDVI, 2019-20 f. yield vs NDVI (mean), combined 2018-19 and 2019-20 for Sentinel 2A] at Kaharole Upazila.

The yield vs. NDVI relationship for Landsat 8 and Sentinel 2A satellite image are shown in Figure 7 revealing that R^2 which are highest and other accuracy criterion like MAPE, RMSE and **N**RMSE of mean NDVI for combined maize growing season i.e., 2018-2019 and 2019-2020 is better fitted than the single maize growing season i.e., 2018-2019 and 2019-2020, respectively for the selected location. The parameters of the regression analysis estimated from the yield vs. NDVI relationship for the combined season, together with the value of R^2, MAPE, RMSE and NRMSE are presented in Table 2. The relationship between mean NDVI for the combined two maize growing seasons and yield is provided almost well compared to the single-season basis NDVI vs. yield relationship for two satellite images.

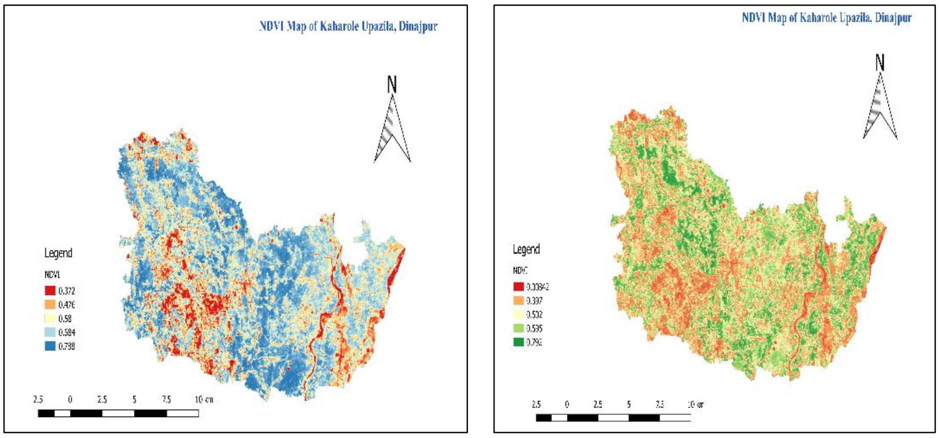
Table 2. Regression parameter and model criterion for the combined season at Farmers' maize field

Location	Satellite Data	Regression parameter for mean value				Best Model Criteria			
		β_0	β_1	SE(β_1)	P-Value	MAPE	RMSE	NRMSE	R^2
Kaharole	Landsat 8	5.291	9.073	1.126	0.0000	0.046	0.589	5.37	0.782
	Sentinel 2A	1.528	12.10	0.961	0.0000	0.031	0.403	3.68	0.898

Here, the regression coefficients of all the fitted model of two satellite images are show a highly significant effect for Kaharole upazila in Table 2. Multiple determination of coefficient (R^2=0.898) along with MAPE (0.031), RMSE (0.403) and NRMSE (3.68) are performed better for Sentinel 2A satellite image than Landsat 8 satellite image for Kaharole Upazila, Dinajpur that are shown in Table 2.

Development and validation of the yield prediction model

The yield prediction model based on the regression analysis was developed based on the yield data collected from the 2018–2019 and 2019–2020 maize growing season. To evaluate the performance of the model validation is essential. Based on the deviation from the estimated and model prediction, model performance can be determined. Hence, the model has been further validated using yield data for the 2020-2021 maize growing seasons. After 103rd and 93rd days after plantation, an NDVI image was selected for two satellite images viz., Landsat 8 and Sentinel 2A from the 2020 to 2021 growing season at Kaharole Upazila, Dinajpur in Figure 8.



a b

Fig. 8. Spatial distribution of the NDVI over the different locations for Landsat 8 and Sentinel 2A satellite images during the growing season of 2020-2021. (a+b): 103rd and 93rd days after plantation for Kaharole Upazila, Dinajpur for different satellite images.

The NDVI value was extracted from each of the 20 farmers' fields for two satellite images from Kaharole Upazila of Dinajpur district during the maize growing season 2020-2021, which were presented in Table 3. As the coefficient of determination (R^2) was found high from the relationship of the mean value of NDVI and the yield of combined season, the validation was done using the general mean value equation (NDVI and the yield relationship of combined maize growing season) for two satellite images. The general mean value equation that was elaborately defined in Table 2 has been used separately in two developed regression models for validation along with two satellite images for the study area. The general regression equation from the mean NDVI (combined maize growing season) is presented in equation 3.

$$\text{Yield} = \beta_0 + \beta_1 * \text{NDVI}_{\text{mean (Combined maize growing season)}} \quad (3)$$

For Landsat 8: $\text{Yield} = 5.291 + 9.073 * \text{NDVI}_{\text{mean (Combined maize growing season)}}$ (4)

For Sentinel 2A: $\text{Yield} = 1.528 + 12.10 * \text{NDVI}_{\text{mean (Combined maize growing season)}}$ (5)

The actual yield (t/ha) of maize and predicted yield (t/ha) using equations 4 and 5 for two different satellite images for Kaharole Upazila during the 2020-2021 maize growing season were also presented in Table 3.

Table 3. Estimated yield and predicted yield from selected Farmer's Maize Fields at Kaharole, Dinajpur during the season of 2020-2021

Farmer's Field	Actual Yield(t/ha)	NDVI Values		Predicted Yield (t/ha)		Error of Yield (%)	
		Landsat 8	Sentinel 2A	Landsat 8	Sentinel 2A	Landsat 8	Sentinel 2A
1	11.25	0.53	0.73	10.1	10.36	10.22	7.91
2	11.65	0.57	0.74	10.46	10.48	10.21	10.04
3	11.95	0.61	0.77	10.82	10.85	9.46	9.20
4	12.55	0.67	0.78	11.37	10.97	9.40	12.59
5	11.53	0.53	0.71	10.09	10.12	12.49	12.23
6	10.67	0.44	0.66	9.28	9.51	13.02	10.87
7	12.37	0.61	0.77	10.83	10.85	12.44	12.28
8	12.2	0.6	0.77	10.73	10.85	12.04	11.06
9	12.15	0.61	0.77	10.83	10.85	10.86	10.69
10	10.89	0.52	0.75	10.01	10.6	8.08	2.66
11	12.35	0.58	0.76	10.55	10.72	14.57	13.19
12	10.35	0.48	0.69	9.64	9.88	6.85	4.54
13	10.87	0.5	0.71	9.83	10.12	9.56	6.89
14	11.42	0.58	0.76	10.55	10.72	7.61	6.12
15	11.1	0.44	0.67	9.28	9.64	16.39	13.15
16	11.25	0.54	0.72	10.19	10.24	9.42	8.97
17	10.34	0.41	0.65	9.01	9.39	12.86	9.18
18	9.1	0.39	0.62	8.82	9.03	3.07	0.77
19	9.5	0.43	0.64	9.19	9.27	3.26	2.42
20	11.04	0.54	0.72	10.19	10.24	7.69	7.25

The highest and lowest observed yield of maize were 12.55 (t/ha) and 9.1 (t/ha) as well as maximum and minimum NDVI values were 0.67 and 0.39 for Landsat 8; 0.78 and 0.62 for Sentinel 2A for farmers' field 4 and 18 respectively during the maize growing season 2020-2021 (Table 3). Largest and smallest predicted yield of maize for Landsat 8 were 11.37 (t/ha) and 8.82 (t/ha) as well as for Sentinel 2A these were 10.97 (t/ha) and 9.03 (t/ha) for farmers' field 4 and 18 respectively during the maize growing season 2020-2021 (Table 3). Maximum and minimum yield gap (%) for Landsat 8 were 16.39 and 3.07 for farmers' field 15 and 18 respectively and also for Sentinel 2A were 13.19 and 0.77 for farmers' field 11 and 18 respectively in selected areas during the maize growing season 2020-2021 (Table 3). The predicted yield of maize for two satellite data was less than the actual yield of maize (under estimated) in each farmer's field for the study areas are shown in Table 3. Depending on the value of Table 3, the validation of maize yield for two satellite images viz., Landsat 8 and Sentinel 2A was presented in Table 4 during the maize season 2020-2021.

Table 4. Validation of maize yield during the season 2020-2021

Location	Actual Yield (Mean)	Predicted Yield (Mean)		Mean Error of Yield (%)	
		Landsat 8	Sentinel 2A	Landsat 8	Sentinel 2A
Kaharole, Dinajpur	11.23	10.09	10.23	10.15	8.82

The estimated farmers' field yield (mean) was 11.23 (t/ha) and the predicted yield (mean) for Landsat 8 and Sentinel 2A were 10.09 (t/ha) and 10.23 (t/ha) respectively in this study areas during the maize growing season 2020-2021. The percentage of mean yield error was 10.15 for Landsat 8 and 8.82 for Sentinel 2A. Error of mean yield (%) of Sentinel 2A was performed better than Landsat 8 during the maize growing season 2020-2021 (Table 4). In Bangladesh, a few researchers used satellite data to forecast the yield of some crops like potatoes, rice, wheat etc. but not for maize. The mean error of potato yield prediction was found 15% by using MODIS (~250m) data at Munshiganj (Bala *et al.*, 2009) and Newton *et al.* (2018) revealed that mean error of potato yield was 10.4% using Landsat 8 (~30m) satellite data at the same location. However, some researchers used the yield-NDVI relationship for maize yield prediction globally. Strong relationships (R^2 from 0.77 to 0.84) were observed between NDVI and the grain yield of maize using RapidEye imagery satellite and promising results for yield prediction were found with a 5% error of estimation in Hungary (Bu *et al.*, 2017).

Conclusion

The study has investigated the prediction capacity of remote sensing NDVI data for maize yield in selected locations viz. Kaharole Upazila, Dinajpur of Bangladesh. It has also investigated the relationship between NDVI and yield for the study region. Here the two satellite images Landsat 8 (OLI) and Sentinel 2A (MSI) which were high spatial resolution data were used in this study in the consecutive years 2018-19, 2019-20 and 2020-21 respectively. Mean NDVI and mean yield of combined maize season performed better than NDVI and its corresponding yield of single maize season for each satellite image and Kaharole Upazila. The yield prediction equations were found based on mean values of NDVI for the combined growing season against the yield of maize. The yield against NDVI relationship for both satellite images showed that R^2 along with MAPE, RMSE and NRMSE of mean NDVI for the combined maize growing season performed better than the single maize growing season for the study area. The absolute mean error of prediction was found to be about 10.15% for Landsat 8 and

8.82% for Sentinel 2A compared to the actual yield during the maize growing season 2020-2021. The predicted yield (mean) of Sentinel 2A which is 1.33% closer to the actual yield than Landsat 8 was observed in this research. The yield prediction model for Sentinel 2A images performed better than the Landsat 8 because of the high spatial resolution (~10m) that was revealed in this study. It was found that NDVI data extracted from Sentinel 2A high resolution satellite images can be successfully used to predict the maize yield for the study area with appreciable accuracy. So, the high resolution Sentinel 2A images can be an effective means for early prediction of maize yield.

References

Atzberger, C. 2013. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. *Remote Sens* **5(2):**949–981

Bala, S. K. and A. S. Islam. 2009. Correlation between potato yield and MODIS-derived vegetation indices. *Int J Remote Sens* **30(10):** 2491–2507

BBS, 2020. *Yearbook of Agricultural Statistics*–2020, 32nd series Statistics and Information Division, Ministry of Planning, Dhaka, Bangladesh

Belgiu, M. and O. Csillik. 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote sensing of environment*, 204, 509–523.

Bu, H., L. K. Sharma, A. Denton, D.W. Franzen. 2017: Comparison of satellite imagery and ground-based active optical sensors as yield predictors in sugar beet, spring wheat, corn, and sunflower. *Agronomy Journal* 109 299–308.

Congedo, L. 2016. Semi-Automatic Classification Plugin Documentation. https://fromgistors.blogspot.com/p/semi-automatic-classification-plugin.html (retrieved 5/12/2016).

Croitoru, A. E., I.-H. Holobaca, C. Lazar, F. Moldovan, and A. Imbroane. 2012. Air temperature trend and the impact on winter wheat phenology in Romania. *Climatic Change*, 111, 393–410.

Islam, A. K. M. S. and S. K. Bala. 2011. Estimation of yield of wheat in greater dinajpur region using MODIS data. 3rd International Conference on Water & Flood Management (ICWFM-2011).

Kim, H. H. and G. C. Elman. 1990. Normalization of satellite imagery. *Int J Remote Sens* **11(8):**1331–1347.

Lai, Y. R., M. J Pringle, P. M. Kopittke, N. W. Menzies, T. G. Orton and Y. P. Dang. 2018. An empirical model for prediction of wheat yield, using time-integrated Landsat NDVI. *Int. J. Appl. Earth Obs. Geoinf.*, 72, 99–108.

Lambert, M. J., P. C. S. Traoré, X. Blaes, P. Baret and P. Defourny. 2018. Estimating smallholder crop production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens*. *Environ*. 216, 647–657.

Li, J. and D. P. Roy. 2017. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sens*. 9, 902.

Lima, T. A., R. Beuchle, A. Langner, R. C. Grecchi, V. C. Griess and F. Achard. 2019. Comparing Sentinel-2 MSI and Landsat 8 OLI Imagery for Monitoring Selective Logging in the Brazilian Amazon. *Remote Sens*. 11: 1-21; doi:10.3390/rs11080961

Liu, W. T. and F. Kogan, 2002. Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices. *Int J Remote Sens* 23(6):1161–1179

Masson-Delmotte, V., P. Zhai, H.-O. Po¨rtner, D. Roberts, J. Skea, P. Shukla, A. Pirani,. W. Moufouma-Okia, C. Pe´an, and R. Pidcock. 2018. Global warming of 1.5 OC: an IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. World Meteorological Organization Geneva, Switzerland.

Newton, I. H., A. F. M. T. Islam, A. K. M. S. Islam, G. M. Tarekul, T. Anika and S. Razzaque. 2018. Yield Prediction Model for Potato Using Landsat Time Series Images Driven Vegetation Indices. *Remote Sensing in Earth Systems Sciences.* 1:29–38.

Panek, E., and D. Gozdowski. 2021. Relationship between MODIS Derived NDVI and Yield of Cereals for Selected European Countries. *Agronomy*. 11, 340. https://doi.org/10.3390/agronomy11020340.

Rahman, M. M. and A. Robson. 2020. Integrating Landsat-8 and Sentinel-2 Time Series Data for Yield Prediction of Sugarcane Crops at the Block Level. *Remote Sens*. 12, 1313; doi:10.3390/rs12081313.

Rahman. A., K, Khan, N. Y. Krakauer, L. Roytman, and F. Kogan, 2012. Using AVHRR-based vegetation health indices for estimation of potato yield in Bangladesh. *J Civil Environ Eng* **2:**111

Ren, J., Z. Chen, Q. Zhou and H. Tang. 2008. Regional yield estimation for winter wheat with MODIS-NDVI datain Shandong, China. *International Journal of Applied Earth Observation and Gioinformation*, 10(4), 403-413. Resolution earth observations: The Global Agriculture Monitoring (GLAM) project.

Reynolds, C. A., M. Yitayew and D. C. Slack. 2000. Estimating crop yields and production by integrating the FAO crop specific water balance model with real-time satellite data and ground-based ancillary data. *Int J Remote Sens* **21(18):**3487–3508 *Sens* 2: 1589-1609.

Rouse, J. W., R. H. Haas, J. A. Schell and D. W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: Proceedings of the ERTS-1 3rd symposium, Vol 1. NASA SP-351. NASA, Washington, pp 309–317.

Segarra, J., M. L. Buchaillot,. J. L. Jose Luis Araus, and S. C. Kefauve. 2020. Remote sensing for precision agriculture: sentinel-2 improved features and applications. *Agronomy*, 10, 641; doi:10.3390/agronomy10050641.

Simonetti, D., A. Marelli and H. Eva. 2015. IMPACTool Box: Portable GIS Toolbox for Image Processing and Land Cover Mapping; *Publications Oce of the European Union: Luxembourg*, ISBN 978-92-79-50115-9.

Skakun, S., E. Vermote, B. Franch, J. C. Roger, N. Kussul, J. Ju and J. Masek. 2019. Winter Wheat Yield assessment from Landsat 8 and Sentinel-2 data: incorporating surface reflectance through Phonological fitting, into regression yield models. *Remote Sens*, 1768:1-19; doi:10.3390/rs11151768. www.mdpi.com/journal/remotesensing.

Teillet, P. M. 1986. Image correction for radiometric effects in remote sensing. *Int J Remote Sens* **7(12):**1637–1651.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ*. **8:** 127–150.

USGS, 2016. Earth Explorer. Available online: http://earthexplorer.usgs.gov/ (accessed on 20 February 2019).