Review of: "Yield Forecasting Model for Maize Using Satellite Multispectral Imagery Driven Vegetation Indices"

Denis Osuna

Potential competing interests: No potential competing interests to declare.

Yield Forecasting Model for Maize Using Satellite Multispectral Imagery Driven Vegetation Indices

Nur Mohammad, Mohammad Amirul Islam, Mohammad Mukhlesur Rahman et al.

Summary of the paper. Propose a corn yield prediction model based on NDVI using high-resolution images from Landsat 8 Operational Land Imager (OLI) and Sentinel “A Multi-spectral Instrument (MSI) at Dinajpur District in Bangladesh. The authors utilized corn yield data from two seasons (2018-2019 and 2019-2020) to generate single-season and combined-season regression models among NDVI and grain yield data. A third season of data was utilized to probe the confidence of the regression models.

Major comments….

Utilizing more current references throughout the document; 40% are from the last five years, 14% are from 2013-2017, and the rest are older than 2012.

Introduction: Eliminate redundancy from the crop surface in hectares and tons; these are the rounded values in millions between parentheses.

In the second paragraph, numerous empirical models have been devised in recent years to forecast yield, but references are missed.

In the second paragraph, the phrase "numerous studies that recognize the high effectiveness of remote sensors in acquiring crop information is discussed," but references are omitted.

The first time the reference BSS is used, the meaning should be specified (Bangladesh Bureau of Statistics, 2020). In the same paragraph, the NDVI is cited; it would be helpful to describe concisely what this index is and highlight it, perhaps because it is one of the most used (simplicity of determination; this description is in the materials and methods, a section in which the description of the formula/procedure of determination could be limited.). In this same paragraph, the machine learning algorithm talks about extracting information, but it uses another index (GNDVI). It would be helpful to exemplify the procedure using NDVI and not enter another vegetation index such as the GNDVI.

The acronym NDVI is repeatedly described in the introduction.

Materials and methods: the description of the region's climate under study can be synthesized by denoting the values of

temperature, humidity, and precipitation during the maize valuation period.

The abstract indicates the study was developed in two cycles, 2018-2019 and 2019-2020, while in the materials and methods, three cycles were pointed out: 2018-2019, 2019-2020, and 2020-2021.

If the maize studied was planted in the Rabi season, describe the environmental conditions of that season. The description of the Kharif season, where other crops are seeded, can be dispensed with since the study focused on maize.

It was pointed out that three images of each mission were used for each growing season, where the criterion was to use the image collected that presented the highest NDVI. Based on this, it is necessary to adjust abstract wording where it is emphasized that a single image was used per cycle per mission, and already in the materials and methods, it could be described that each mission takes images every 16 days in Landsat 8 and every 10 days in Sentinel; the current wording seems to use NDVI values of each date and that based on this, models were generated, mainly because in one part of the intro the reference to Panek et al. (2021) used by the authors privileges those regression models for grain yield that use NDVI time series.

The section describing the calculation of NDVI will only focus on how this vegetation index was calculated and the values that can be obtained for this index, as already done. Equation 1 was referred to in the text in the materials and methods, and the equation (NDVI calculation formula) had its corresponding number. Equation 2 was listed but needed to be fully cited in the text.

Tables 2, 3, and 4 could add average data and standard error for each column showing NDVI values per imaging mission. This could help detect which images (Landsat 8 vs. Sentinel 2A) presented higher NDVI values, in addition to the dispersion of observations. It would also be helpful to add a column indicating the area (ha) that each farmer's field considered.

In Figure 3, put the letters a, b, c, d at the top right of each image. It could be that the direction of the images is located at the upper level (Landsat 8) from left to right in the 2018-2019 cycle, and at the lower level, the Sentinel 2A images are on the left (2018-2019) and right (2019-2020).

Page 11, review the t/ha units... One of these says ha/t. Please correct.

In Table 5, for probability values instead of pure values of 0, I recommend putting “$p < 0.01$” instead.

Table 6, aggregate the average and standard error for each column of NVDI values, predicted yield, and error yield.

The results obtained are not discussed.

Conclusions: Summarize conclusions, highlighting general statements supported by the evidence. Provide a future perspective on the work, i.e., a roadmap recommendation for readers for future work.