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

Afolabi Agbona[1]

1 Texas A&M University - College Station

Potential competing interests: No potential competing interests to declare.

The authors has described a yield prediction approach in maize using derived vegetation index trained using Linear modelling approach. It's indeed an insightful article that describes method that could be used to increase adoption of remote and proximal technology in agriculture.

Please find below my comments regarding the article.

General comments

1. Is there any reason why Normalized Difference Red-Edge index (NDRE) was not used to compared performance with NDVI?. Considering the dense nature of corn leaves and also the Red-Egde band in Sentinel 2 that can penetrate leaves more deeper than red light. I will expect some advantage in using NDRE because NDVI saturates when there is more intense canopy.

2. Generally, references needs to be updated. Use some more recent citations.

3. Use some more specific remote-sensing terms. For example, under the “Yield prediction model” section the data of maize yield was mentioned…this in simple term is the “ground reference data”

4. The author needs to reword the section “Maize yield and NDVI relationship using regression model” appropriately. I will suggest “Prediction accuracy using Linear Regression model for Maize yield and NDVI

Consequently, it should be spelt out in the paragraph that the “metrics for prediction accuracy are “coefficient of determination (R2), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). How is each interpreted (Though, it was mentioned in another paragraph. It will be good to bring it into this section).

5. The article needs to be thoroughly checked for grammatical correctness. There are too many missing verbs, noun etc in between text. For example;

(a.) “expanded potato yield estimations” (remove “s”)

(b.) “The maximum coefficient of determination (R^2) of yield forecasting equation was found to be 0.81 between the mean NDVI and potato yield and and the result revealed ” (There are 2 “and”)

(c.) NDVI. "An NDVI close to 0 corresponds to no vegetation, while Values lies between -0.1 to 0.1 generally corresponded to barren areas of rock, sand or snow. (change Values to lowercase, values “that’ is also missing)

(d.) Under “Code availability” section, “All the used in the analysis of this study are available from the corresponding author on request” ("codes” is missing here)

Specific comments

6. Maize (*Zea mays L.*) is also known as corn, is the world’s fourth major staple food crop after Rice, Wheat and Potato (Reference needed)

7. Maize is initially grown for grain and secondly for fodder and raw material for industrial purpose. (Reference needed)

8. “However, maize growth monitoring and its yield estimation has become a major issue of consideration”….In my opinion, this statement is incomplete. You may want to end it with sentence like “Because in involves the use of traditional methods that are cumbersome and inefficient”.

9. “In recent years, numerous empirical models have been devised to forecast crop production prior to harvest”. (Reference needed for articles where empirical models have been used)

10. “Satellite-based remote sensing is widely recognized as a highly effective technology for acquiring crucial data pertaining to crop distribution and the prevailing growing conditions across extensive regions”…Add a good number of references for this statement

11. Simple Linear Machine Learning is not an algorithm….provide the name

12. How is RMSE and R2 interpreted. Provide an inline sentence on that. Do not assume the reader understand.

13. Under the section “Satellite Image Pre-Processing”, Please expatiate on how TOA was done in the processing of LandSAT. Was it in python, which packages and libraries was used ; provide ref to those libraries as well. OR was it hand calculated.

This applies to radiometric correction. The author has only mentioned that the “radiometrically calibrated images was subjected for DOS-1”.

14. Equation (3) – Every model equation needs to have an error term. Please include it

$Y = B0 + B1*NDVI_{meanxxxx} + e$