REPORT OF

WHEAT PRODUCTION SURVEY IN NIGERIA

Executive Summary

Wheat farming is experiencing renewed interest from policy makers who see Nigeria's potential to be self-sufficient in wheat production. But as vital as wheat is, we cannot say with certainty the actual quantity of wheat produced in the country as conflicting figures were being published from different sources. There has been no recent survey or study to estimate the wheat production in Nigeria. The last official estimates of the wheat acreage is 60,000 hectares and dates back to 2008. Much has changed since then, and this study provided an overview of the current wheat production in Nigeria.

The wheat production survey was conducted across thirteen (13) states in Nigeria covering 2020/2021 farming season by the National Bureau of Statistics. In this report, we analyze wheat production in thirteen northern states of Nigeria. Wheat yield per hectare was estimated through crop cutting experiment while total land area cultivated was estimated by satellite mapping. In determining the total land area cultivated, we first built an Artificial Intelligence (AI) model which uses remote sensing satellite images to distinguish between wheat and non-wheat pixels. This helps us predict wheat acreage for different Local Government Areas (LGAs). We also estimated the yield for all the pixels predicted as wheat and our results were in close accordance (85%) with Crop Cutting Experiment.

We validated the predicted acreages in multiple ways: 1) Visual Validation of the prediction, 2) Expert Feedback from FMAN and NBS on LGA-wise acreage and 3) On plots of other crops grown during the same season. Through the extensive validation we are reasonably confident about the validity of the estimations, as well as aware of its certain shortcomings which could be handled in future iterations. The total Wheat Production was estimated from the product of yield per hectare and the total land area cultivated. Wheat is cultivated in an area of about 11,820 hectares with a production of 36,943.80 tonnes of grain across the 13 states.

TABLE OF CONTENT

ABBREVIATIONS

CCE	Crop Cutting Experiment
ESA	European Space Agency
Eta	Actual Evapotranspiration
ETc	Crop Evapotranspiration
FAO	Food and Agriculture Organization
FMAN	Flour Mills Association of Nigeria
Ky	Crop Stress Coefficient
Kc	Crop Coefficient based on stage
LGA	Local Government Area
IBM	International Business Machines
MAPE	Mean Absolute Percentage Error
NBS	National Bureau of Statistics
NDVI	Normalized Difference Vegetation Index
RMSE	Root Mean Squared Error
USDA	United States Department of Agriculture
USGS	United States Geological Survey
Ym	Maximum potential Yld

1.0 Introduction

Wheat is a cereal widely cultivated in the northern part of Nigeria. It is typically milled into flour which is then used to make a wide range of foods such as bread, pasta, cakes, noodles, pastries, among others. Since the oil shock of 2015/2016, wheat farming is experiencing renewed interest from policy makers who see Nigeria's potential to be self-sufficient in wheat production. But as important as wheat is, we cannot say with certainty the actual quantity of wheat produced in the country as conflicting figures were being published from different sources. It is on this note that the National Bureau of Statistics (NBS) conducted wheat production survey across twelve (12) states in Nigeria covering 2020/2021 farming season. The survey was sponsored by the Flour Milling Association of Nigeria (FMAN), while Cropin Technology Solutions was contracted for the Satellite Mapping. Representatives of Federal Ministry of Agriculture and Rural development (FMARD) as well as Central Bank of Nigeria (CBN) served as advisors and technical supporters of the Study.

1.1 Objectives

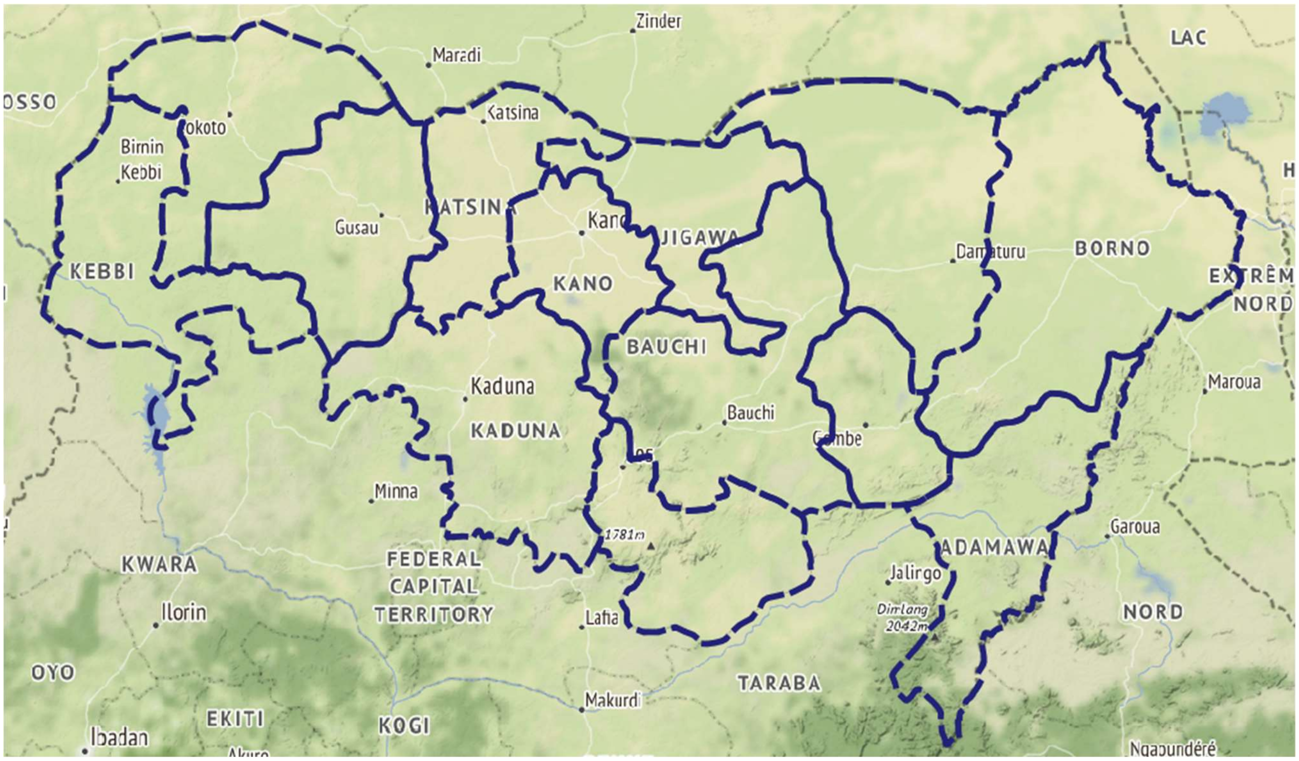
- To determine the yield per hectare of wheat
- To determine total production of wheat in Nigeria
- To determine total land under cultivation with wheat
- Estimate acreage and crop yield for Wheat during 2020 - 2021
- Validate the predicted results with ground collected data.

1.2 Scope:

The study includes two main components: (1) crop cut survey of ~700 farmers in 12 states to determine average yields and (2) satellite mapping of all land in 13 states to determine total land under cultivation. Information gathered include demographic characteristics of Farmers, general Information on Plot/Farm, Satellite mapping and Crop-Cutting/measurement (Wheat Production).

1.3 Coverage:

The field work commenced on the 14th of April, 2021 across the selected states (Adamawa, Bauchi, Gombe, Jigawa, Kaduna, Kano, Katsina, Kebbi, Sokoto, Plateau, Yobe, and Zamfara). The data collection lasted for twelve (12) days. It is also expedient to note that Borno State was initially considered to be covered, but it was realized that the LGAs where wheat is cultivated were having serious security challenges and could not be accessed by the enumerators but was covered by the satellite mapping.



1.3 Survey Instruments/Equipment Used for Crop Cut

The survey instruments/equipment utilized include paper Assisted Personal Interviewing (PAPI) Computer Assisted Personal Interviewing (CAPI) based questionnaire, Listing/Selection Forms, Manual of Instruction, Hand-held GPS, Weighing Scale, one square meter wooden box and Sacks. The instrument used was specifically designed for this survey and was subjected to series of reviews by the technical committee to see that it captures what it intends to measure.

1.4 Sample Size/Design for Crop Cut

A multi-stage random sampling method was adopted. Two (2) wheat growing LGAs was planned initially to be surveyed per state but this was not the case in most of the states as majority of the farmers had harvested their wheat as at the time the study commenced. We ended up surveying more than two LGAs in some states in order to be able to cover the required number of wheat farms per state. Again Five (5) EAs were supposed to be canvassed in each of the selected LGA as planned in the design but this was not feasible in some states. A total of10 EAs was initially planned to be canvassed per state but in the course of fieldwork we increased the number of EAs to 12. We started working from the farms to the households. We went to the farm, identified area that was not harvested and interviewed the farmer right on the farm. Fifty (50) wheat farming households was planned to be canvassed per state but a maximum of 60 wheat farms was covered per state since we did not carry out listing exercise in the entire EAs. A Total of 4 field staff conducted the survey in each state. Two teams were constituted per state.

1.5 Design for the Satellite Mapping

We first built an Artificial Intelligence (AI) model which uses remote sensing satellite images to distinguish between wheat and non-wheat pixels. This helps us predict wheat acreage for different Local Government Areas (LGAs). We also estimated the yield for all the pixels predicted as wheat. We validated the predicted acreages in multiple ways: 1) Visual Validation of the prediction, 2) Expert Feedback from FMAN and NBS on LGA-wise acreage and 3) On plots of other crops grown during the same season. Through the extensive validation we are reasonably confident about the validity of the estimations, as well as aware of its certain shortcomings which could be handled in future iterations. The data used for this study are Satellite Data, Weather, Ground Data which includes multiple datasets such as Optical data (Sentinel 2A/B) and Microwave satellite data (Sentinel-1) for crop detection, Smart sampling and sowing calendar generation. Landsat 8 OLI/TIRS and weather data were utilized for formulation of the yield model. Reference Evapotranspiration was calculated using weather data and crop specific coefficients. Auxiliary datasets used in the pilot study such as crop calendars, seasonality parameters etc. were utilized from Government sources.

Sentinel-2 processed data with cloud removal algorithms were used. Sentinel-2 data for a given season which had been corrected for atmospheric conditions were used in the analysis. Daily weather data at a resolution of 5 km were collected from IBM weather server which gave complete weather variables including precipitation, mean temperature, wind speed, relative humidity and other auxiliary data which has been utilized for estimation of Evapotranspiration and yield proxies

1.6 Training for Crop Cut

The training was conducted at two levels. The first level was a national level training-of-trainers (ToT) held at NBS headquarters in Abuja. The ToT took place over two days (9th - 10th of April, 2021) inclusive of field practice. In attendance were representatives of the Federal Ministry of Agriculture (FMARD), Lake Chad Research Institute (LCRI), Flour Milling Association of Nigeria and the National Bureau of Statistics (NBS).

The training went on smoothly with class practical on the use of hand-held GPS and crop cutting procedures. The second day of the training was basically field practice on the method of carrying out crop cutting.

The Field practice was held on an un-harvested rice farm in Kuje Area Council. The farmer took the trainers to his farm where live field practice took place, two different farms were measured and the crop cutting carried out on the sample boxes.

1.7 Training of Enumerators

The state level training, also referred to as the training of enumerators was held over a period of two days (12^{th} - 13^{th} of April, 2021) inclusive of one day of field practice across the twelve selected states. All the expected participants attended the training. During the training, participants were trained on the objectives and methodology of the wheat survey, the use of GPS devices, weighing scale, laying wooden box on the farm and computer assisted personal interviewing (CAPI) questionnaires in the survey solution platform. The field practice was carried out across the 12 states in un-harvested wheat farms.

1.8 Field Work for Crop Cutting/Objective measurement

In this survey, we utilized objective measurement (use of GPS) to measure farm size. The Enumerator and the Crop Farmer/Crop Farm Manager walked along the edges of the plot and identified the boundaries and cleared all obstacles that could hinder effective measurement. Thereafter, the enumerator walked round the plot with the GPS to take the measurement of the farm and recorded accordingly. This was followed with the measuring of yield plot. In carrying out crop cutting two methods were adopted by the enumerators.

1.8.1 Measuring a farm that was below 4000sqm

Two crop yield measurements were taken from the wheat farms of size below 4000sqm. The enumerator entered the farm, took 15 steps right into the farm and laid the yield plot at the spot (measured 1st sample box of 1sqm) and labelled it sample one. The enumerator then located the extreme (at the back) of the farm and took 15 steps right into the farm and laid the yield plot at the spot and labelled it sample two. These two samples were harvested, threshed separately, weighed and recorded accordingly.

1.8.2 Measuring a farm of 4000 sqm and above in size:

Four sample boxes were taken from the wheat farm of size 4000 and above. The enumerator entered the farm, took 15 steps right into the farm which served as the starting point (spot 1), then took another 15 steps to the right to set 1sqm sample box (1st sample box). Walked back to the starting point (spot 1) and took another 15 steps to the left to setup the second 1sqm sample box (2nd sample box).

The enumerators walked to the other end of the farm, entered the farm from the extreme and took 15 steps right into the farm which served as another starting point (spot 2), then took 15 steps to the right to set 1sqm sample box (3rd sample box). Walked back to the starting point (spot 2) and took another 15 steps to the left to setup the second 1sqm sample box (4th sample box).

The four sample boxes mapped for crop cut were separately harvested, threshed, weighed, recorded and entered into CAPI accordingly.

The field work lasted for 12 days (14th – 25th of April, 2021).

1.9 Monitoring/Coordination of fieldwork

In order to ensure quality data collection, monitoring/coordination of field work started concurrently with field work. This was to ensure that the enumerators adhere strictly to the guidelines and did not deviate from the work plan. Daily plausibility checks were carried out by the ICT department in order to identify and summarize key quality issues from data that was sent to the NBS server.

Problems identified through the daily review was discussed with the appropriate teams and with corrective measures enacted to mitigate future occurrence of such problems. Monitoring exercise officially lasted for 5 days (14th-18th April, 2021) for NBS Monitors but they continued to keep in touch with the enumerators throughout the period of fieldwork. Staff of the LCRI, IAR, FMARD and FMAN took part in monitoring and coordination of fieldwork.

1.9.2 Area and Production

The total land area cultivated was estimated from the satellite mapping conducted across the thirteen (13) states surveyed while the wheat yield per hectare was estimated from the crop cut. Crop cut was not carried out in Borno and Adamawa states but we imputed the yield per hectare in Yobe state to Borno state and Gombe state to Adamawa state since they are neighbouring states instead of estimating the two states at zero yield per hectare. The total Wheat Production was estimated from the product of yield per hectare and the total land area cultivated.

Table 1: Estimated Area (Ha) by Satellite Mapping and Production in (Kg) Crop Estimate by NBS

SN	State	Acreage (Satellite Mapping)	Yield/HA in kg (crop cut)	Yield/HA in MT (crop cut)	Production (Kg)	Production (Metric Tonnes)
1	Kano	2,326	2,800	2.8	6,512,800	6,512.8
2	Jigawa	2,091	2,800	2.8	5,854,800	5,854.8
3	Kebbi	1,340	3,300	3.3	4,422,000	4,422.0
4	Bauchi	1,157	3,400	3.4	3,933,800	3,933.8
5	Kaduna	1,140	2,700	2.7	3,078,000	3,078.0
6	Gombe	687	3,700	3.7	2,541,900	2,541.9
7	Yobe	626	3,200	3.2	2,003,200	2,003.2
8	Katsina	509	3,600	3.6	1,832,400	1,832.4
9	Plateau	505	3,800	3.8	1,919,000	1,919.0
10	Sokoto	393	3,200	3.2	1,257,600	1,257.6
11	Zamfara	192	3,500	3.5	672,000	672.0
12	Borno	487	3200	3.2	1,558,400	1,558.4
13	Adamawa	367	3700	3.7	1,357,900	1,357.9
Total		11,820	3125.53	3.1	36,943,800	36,943.8

Wheat is cultivated in an area of about 11,820 hectares with production of 36,943.8 tonnes of grain across the 13 states. Maximum area under wheat cultivation is in Kano State (19.68%), followed by Jigawa State (17.69%) while Zamfara state accounted for the least (1.62%). Similarly, state wise comparison of wheat production for 2021 shows that Kano state is the major producer of wheat with production of 6,512.8 tonnes, followed by Jigawa state (5,854.8 tonnes), Kebbi state (4,422 tonnes) while Zamfara state ranked last (672 tonnes). However, in terms of average yield per hectare major wheat-growing states is low compared to other states, Plateau state recorded the maximum yield (3.8 tonnes/hectare), followed by Gombe state (3.7 tonnes /hectare) while Kaduna state has the least (2.7 tonnes/hectare) followed by Kano and Jigawa states with 2.8 tonnes per hectare each.

Table 2: Number of Wheat Farm where a crop cut yield survey was conducted per State		
S/N	State	Number of Wheat Farm Covered
1	Adamawa	60
2	Bauchi	55
3	Gombe	60
4	Jigawa	60
5	Kaduna	60
6	Kano	60
7	Katsina	60
8	Kebbi	65
9	Plateau	59
10	Sokoto	60
11	Yobe	60
12	Zamfara	55
	NATIONAL	714

Table 3: Average of the wheat Farm size covered for crop cut survey per State			
S/N	State	Average Farm Size in square metre	Average Farm Size in Hectares
1	Adamawa	6464.81	0.65
2	Bauchi	3437.73	0.34
3	Gombe	2236.17	0.22
4	Jigawa	2864.31	0.29
5	Kaduna	2224.20	0.22
6	Kano	3339.50	0.33
7	Katsina	2683.03	0.27
8	Kebbi	6132.34	0.61
9	Plateau	1723.20	0.17
10	Sokoto	6688.11	0.67
11	Yobe	3172.35	0.32
12	Zamfara	4267.33	0.43
	NATIONAL	3787.67	0.38


Figure A: Picture of farmer doing Crop Cutting


Figure B: Farmer using weighing scale for Wheat yield

Figure C: Recording of wheat yield and interviewing of wheat farmer enumerators

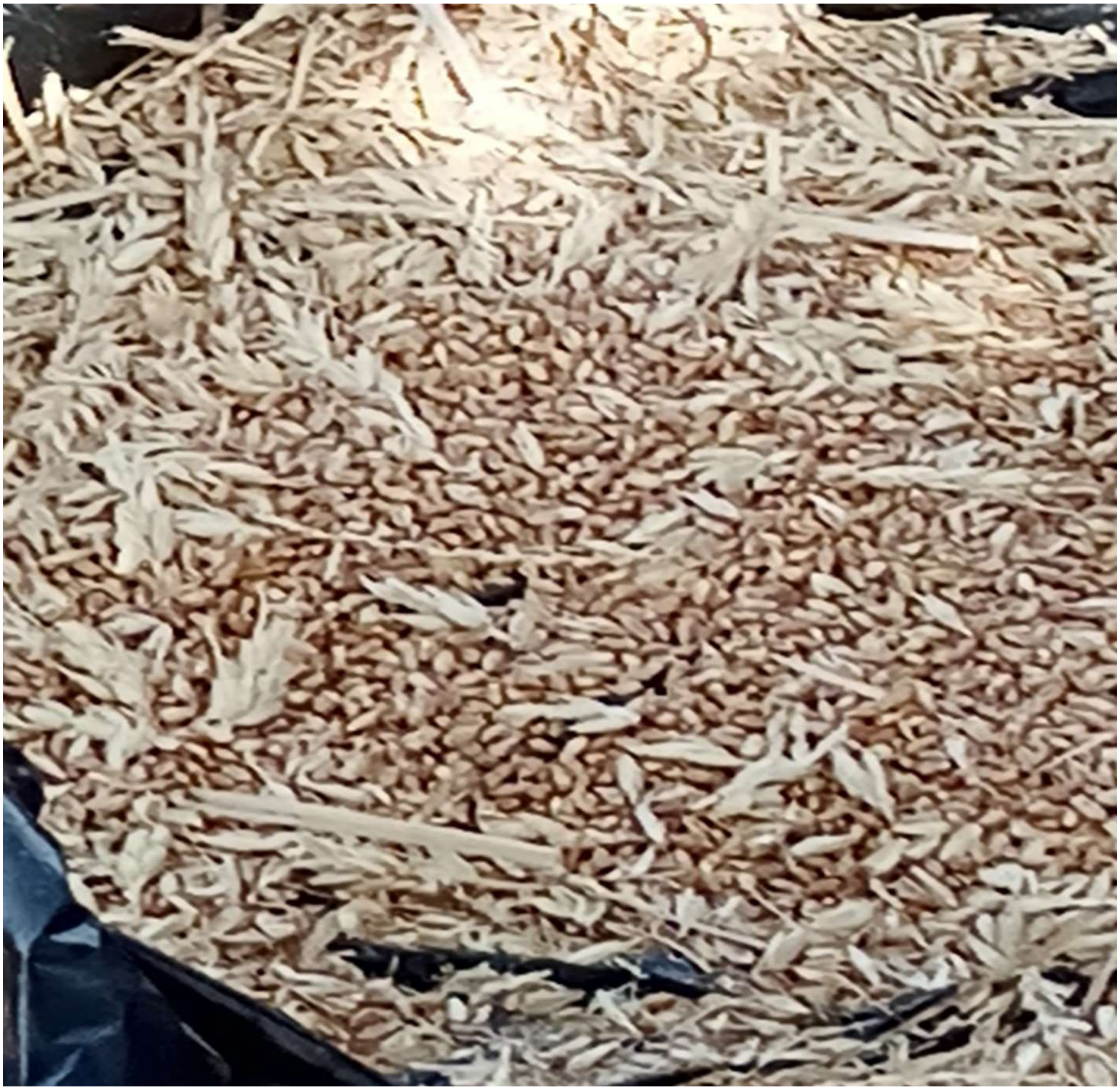


Figure D: Wheat seeds threshed by enumerators


Figure E: NBS Monitor on monitoring exercise to a wheat farm

Figure F: Reading and recording of data from GPS


Figure G: Field personnel with wheat farmers

Figure H: Threshing of wheat seeds by field personnel

Figure I: Laying of wooden box for crop cut

2.0 Satellite Mapping for Crop Acreage and Yield wheat Production in Nigeria

The data source for this study includes multiple datasets such as Optical data (Sentinel 2A/B) and Microwave satellite data (Sentinel-1) for crop detection, Smart sampling and sowing calendar generation. Landsat 8 OLI/TIRS and weather data were utilized for formulation of the yield model. Reference Evapotranspiration was calculated using weather data and crop specific coefficients. Auxiliary datasets used in the pilot study such as crop calendars, seasonality parameters etc. were utilized from Government sources. Data used in the study and their specification are illustrated in table 2.

Table 4 : Data Used: Satellite Data, Weather, Ground Data

The study includes multiple datasets such as Optical data (Sentinel 2A/B) and Microwave satellite

	Sentinel - 2	LANDSAT	Evapotran-spiration	Rainfall	Acquisition dates
Specification	**Sentinel 2A/B** **Pixel size:**10m **Bands:** All optical bands	**Landsat 8 OLI/TIRS** **Pixel size:** 30m	**Model:** Energy balance Model **Parameters:** Weather, Location information etc. **Pixel size:** 5 Km **Provider:** IBM **Frequency:** Daily	**Pixel size:** 5 Km **Provider:** IBM **Frequency**: Daily	Nov Month Stack Dec Month Stack Jan Month Stack Feb Month Stack Mar Month Stack

data (Sentinel-1) for crop detection, Smart sampling and sowing calendar generation. Landsat 8 OLI/TIRS and weather data were utilized for formulation of the yield model. Reference Evapotranspiration was calculated using weather data and crop specific coefficients. Auxiliary datasets used in the pilot study such as crop calendars, seasonality parameters etc. were utilized from Government sources. Data used in the study and their specification are illustrated in table 4.

Sentinel-2 processed data with cloud removal algorithms were used. Sentinel-2 data for a given season which had been corrected for atmospheric conditions were used in the analysis. Daily weather data at a resolution of 5 km were collected from IBM weather server which gave complete weather variables including precipitation, mean temperature, wind speed, relative humidity and other auxiliary data which has been utilized for estimation of Evapotranspiration and yield proxies

Table 5: Dataset / Model Description used in the analysis

S/No	Dataset	Specification	Resolution (m)	Source
1	Plot geotags	Crop info	-	FMAN & NBS
3	Sentinel-2	Optical	10	ESA
4	LANDSAT 8 OLI/TIRS	Optical	30	USGS
5	Rainfall, Temp, Humidity etc	Weather	5000	IBM
6	Ancillary information		-	Literature Survey * (USDA Report, FAO)

*http://www.fao.org/land-water/databases-and-software/crop-information/wheat/en/

Steduto, P., Hsiao, T. C., Fereres, E., & Raes, D. (2012). *Crop yield response to water* (Vol. 1028). Rome: Food and Agriculture Organization of the United Nations.

2.0.1 Methodology

The overall methodology of the pilot study flows through three phases viz.

1. Crop detection using distance-based model
2. Yield model deployment using Energy Balance Model
3. Validation and Statistics

2.1 CROP DETECTION MODEL

To perform crop classification, we analyze Normalized Difference Vegetation Index (NDVI) time series data for the duration of the crop. NDVI is calculated using two bands from the Sentinel-2 data, NIR and Red. It helps in identifying vegetation growth, with a high value corresponding to higher density of vegetation. By looking at the time series profile of NDVI, one can identify crop cycles as there is a gradual increase in NDVI, followed by a sharp decrease post-harvest. Each crop has a specific NDVI signature, and we leveraged this fact to build our model to detect wheat.

To have the reference NDVI crop signature, we analyzed the wheat farm plots shared by NBS and FMAN for the 2019-2020, and 2020-2021 season.

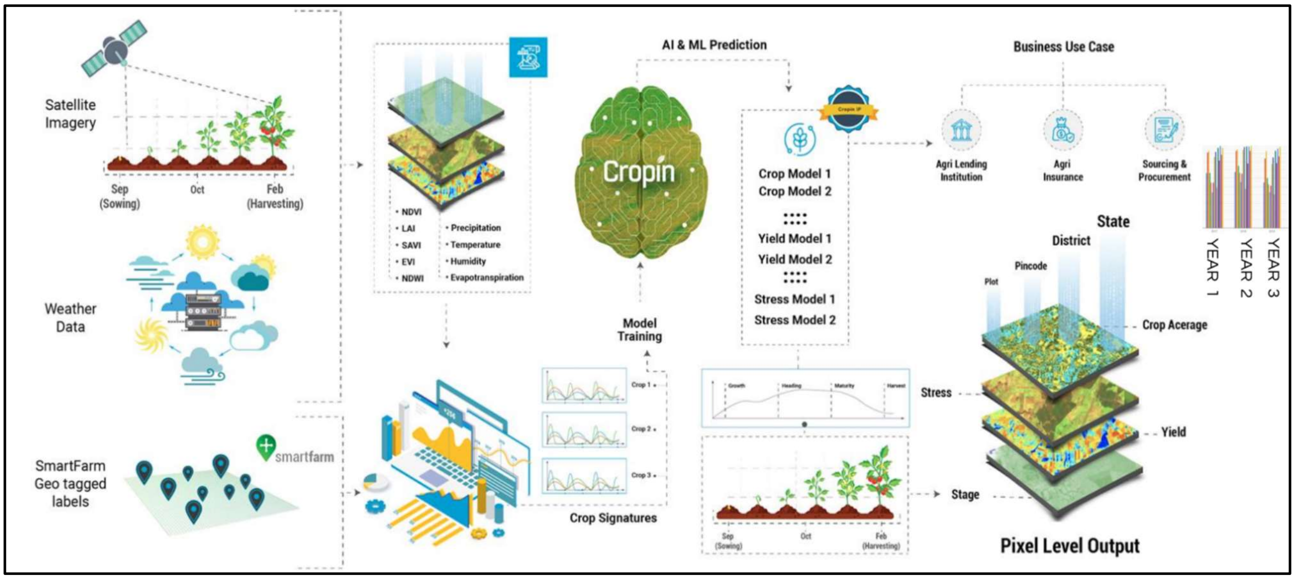


Figure 0:J: NDVI crop signature for Crop detention Model

Table 6: S**ummary of Data Used and retained to extract Wheat Signature**

Source	Year	No. of States covered	Total plots	Retained	Outliers
NBS	2020_21	Kano, Jigawa, Yobe, Gombe, Kebbi, Bauchi, Kaduna, Katsina, Sokoto, Plateau, Zamfara	760 Plots - 22469 Pixels	601 Plots - 17426 Pixels	159 Plots - 5043 Pixels
FMAN	2020_21	Jigawa, Kano, Kebbi, Kaduna	648 Plots - 36657 Pixels	564 Plots - 22747 Pixels	85 Plots - 13910 Pixels
	2019_20	Jigawa, Kano, Sokoto	835 Plots - 58588 Pixels	349 Plots - 16691 Pixels	486 Plots - 41897 Pixels

We first extracted the NDVI time series data for all pixels in the farm plots. All the pixels were then clustered (using K-Means algorithm) based on their profile. In the following two figures, you can see the profile of clusters which were retained as well as rejected.

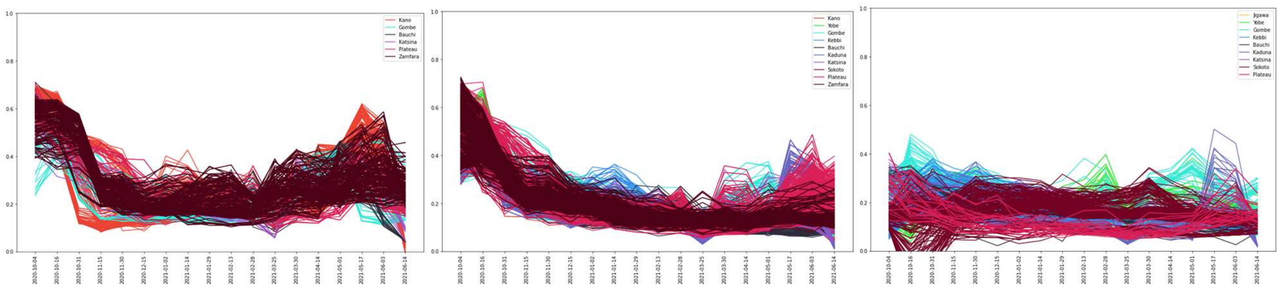


Figure K: NDVI Profile of <u>rejected</u> clusters. These are flat-lines which indicate no crop was grown there.

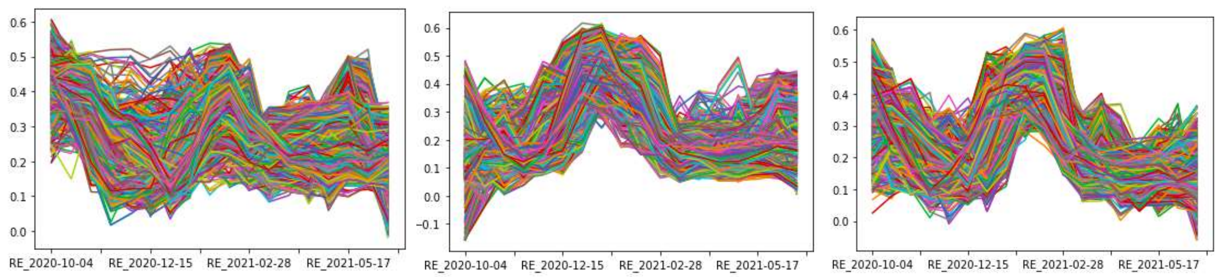


Figure L:** **NDVI Profile of <u>retained</u> clusters. NDVI increases and then decreases, indicating growth of wheat..

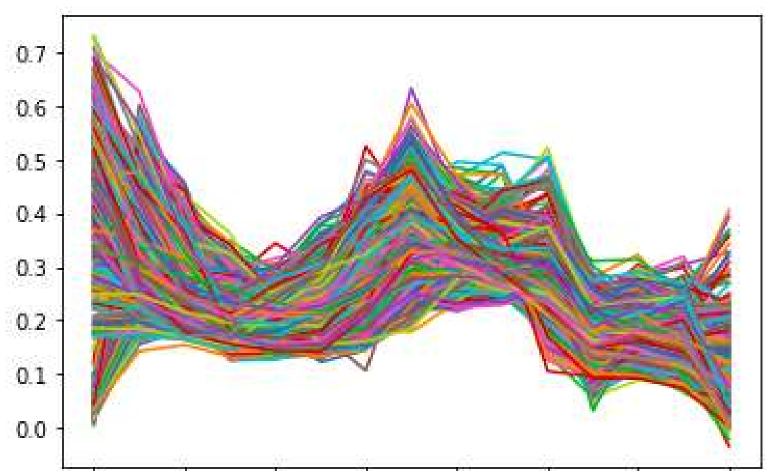


Figure M:** **NDVI Profile of pixels classified as wheat by the model. All the pixels resemble closely the wheat crop cycle and duration.

The retained clusters helped us extract NDVI crop signature for wheat in Nigeria (November - April). We compared each pixel for the entire region of interest with these crop signatures to see how similar they were. If the distance between the two profiles was smaller than the threshold value, which could be tuned, then we classified the particular pixel as wheat. Extensive testing was done to tune the model to not pick forests or barren land as wheat. We can see the type of NDVI profiles which was classified by the model as wheat.

2.2 YIELD MODEL

Crop yield has been estimated based on FAO yield model which consumes Evapotranspiration as a major parameter. Crop Evapotranspiration (ETc) takes into account water transpired from the crop as well as evaporation from the soil. It accommodates the water use efficiency of the crop and the atmospheric demand for moisture. ETc is an important parameter for assessing the water stress which a crop undergoes and according to the amount of stress which it experiences during the entire growing period which has a significant impact on crop yield.

In this study, FAO Model has been used to estimate Reference Evapotranspiration (ETr) which utilizes weather variables such as Mean temperature, Relative Humidity, Wind speed and Solar radiation. Then ETc has been calculated based on crop coefficients at different growing stages of crop and Reference Evapotranspiration data.

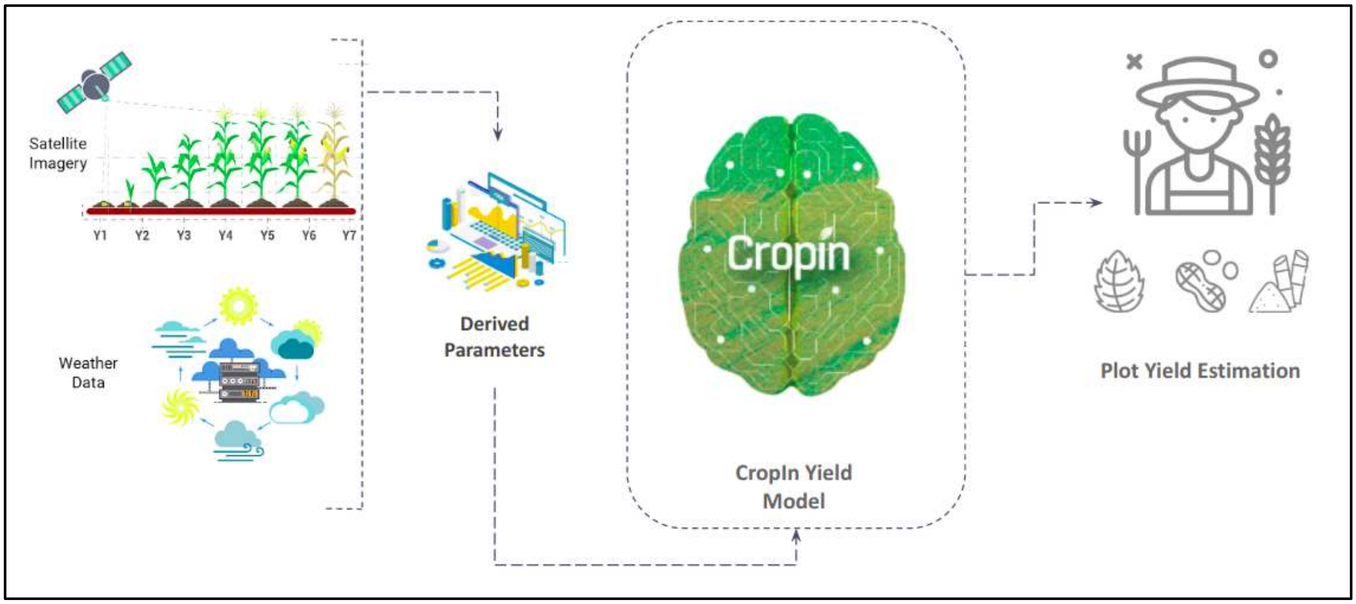


Chart 1: Yield Model flow chart

Yield Model flow chart

Yield forecast is based on a series of simulations using a weather data set combining real time weather and historical weather data and conducted on a weekly basis. Crop yield was computed for the period of sowing to harvest at an interval of month from Nov, 2020 to Feb, 2021 which was considered as crop season for field crops. As given in the sowing calendar, the stage model was taken into consideration to predict the exact maturity date for the crop yield prediction. Each plot constitutes a maturity date, using which it picks the crop agnostic parameters defined to maturity stage such as Kc values and other supplementary data.

The approach is to use all the satellite data of LGAs for a yield model with dynamic Kc values and crop growth specific Ky has taken into consideration and Ym values has been estimated from CCE data collected by the NBS team. The dynamic Kc of Wheat ranges between 0.6 - 0.9 which indicates that all the regions in respective months have reached maturity and have been utilized for the deployed yield model.

3.1 CROP DETECTION RESULTS

Using the model mentioned previously, all the 13 states were processed pixel-wise, and the pixels were classified as wheat or non- wheat. The data was then aggregated at state level. The total acreage predicted for the 13 states was **11,820 Hectares.** The state-wise breakup can be seen in the following figure:

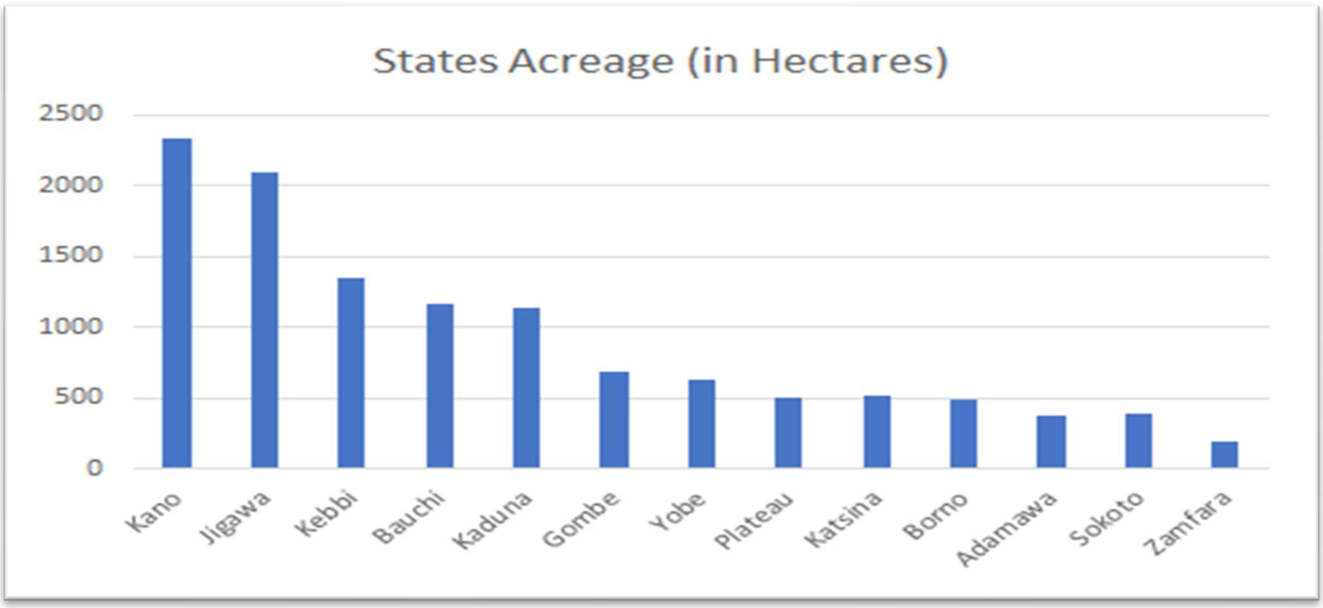


Chart 2: Bar - Plot of State-Wise Acreage

Table 7: State-Wise Acreage

State	Acreage	State	Acreage
Kano	2326	Yobe	626
Jigawa	2091	Katsina	509
Kebbi	1340	Plateau	505
Bauchi	1157	Borno	487
Kaduna	1140	Sokoto	393
Gombe	687	Adamawa	367
		Zamfara	192

3.2 PROBABILITY OF OVERPREDICTION AND UNDERPREDICTION

3.2.1 OVERPREDICTION

In the model developed, over prediction could happen due to two reasons: *1)* Non Agricultural Land being predicted as farmland, and *2)* Other crops being mistaken as wheat. To avoid error due to the first reason, the developed wheat detection model was first validated by looking at pixel level validations overlapped with Google Earth images. This provided us with a sanity check. All model parameters were tweaked till predictions were neatly coming **only** over farmlands, which was reassuring for the model.

Pixel-wise wheat prediction (Bright Green) overlapped with Google Earth. 1) All predicted pixels fall farmland. Not all the farmland is being predicted as wheat, which is reassuring, since our model is not predicting all crops as wheat. 2) Most of the predicted wheat was close to rivers, which makes sense as growing wheat in that season requires irrigation.

This leads us to have confidence in our model that all the pixels being predicted as wheat are having at least some crop growing during that season. To ensure that other crops are not being predicted as wheat, we validated our predictions against maize and rice plots shared by FMAN. We chose these crops as they grow in the same season and could get confused with wheat.

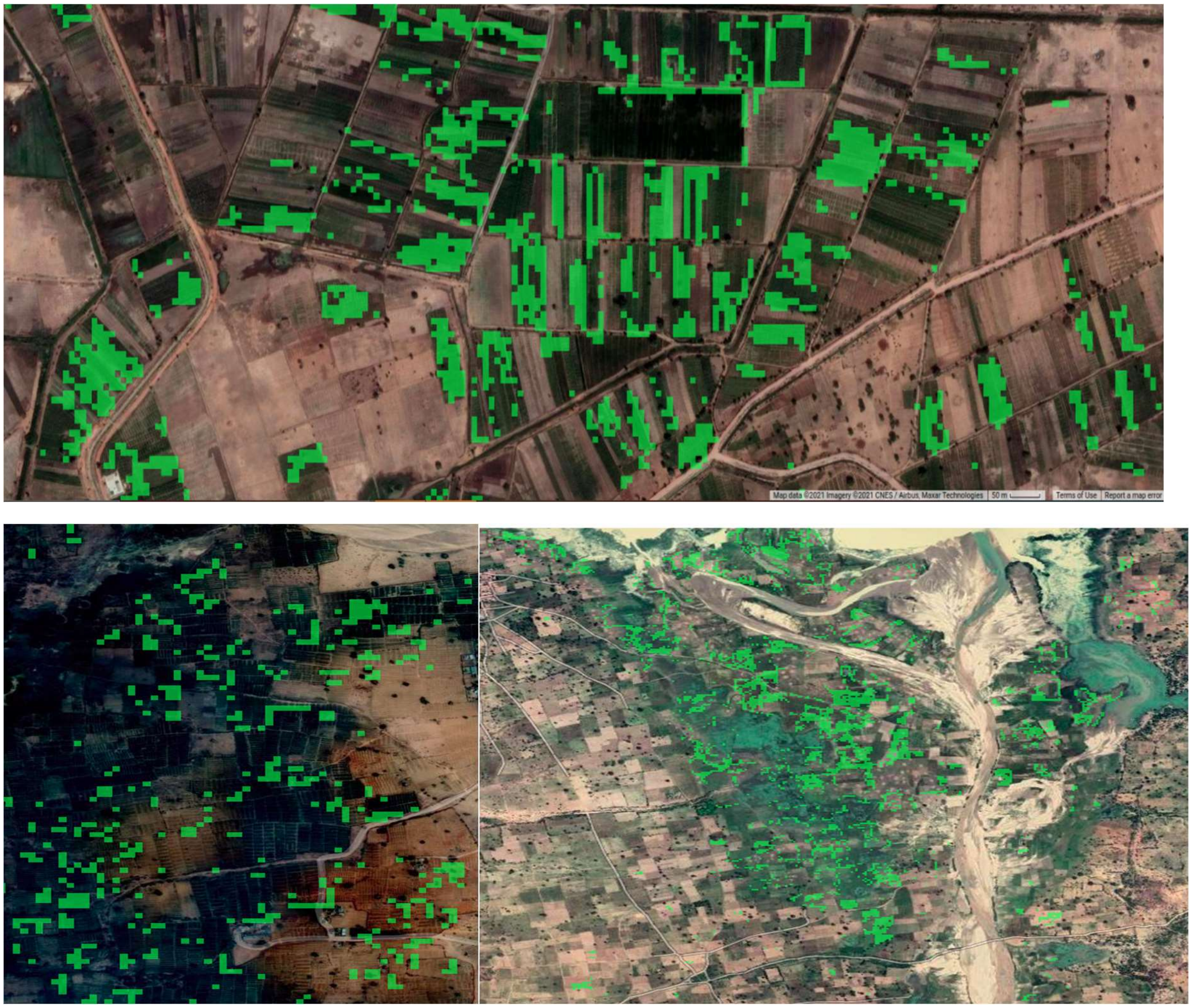


Figure N: Maize and Rice Validation

Maize and Rice Validation

The most likely crops to get predicted as wheat were dry season maize, rice, as well as some vegetables. This is because they have a similar sowing and harvesting time and thus some pixels might get falsely predicted as wheat. Following are the results

Crop	State	Percentage Error
Maize	Jigawa	0
Rice	Kebbi	0
Rice	Jigawa	2-10% (for different plots)

*Sample plots have been provided in the Appendix

On further investigation, the rice pixels being predicted as wheat had a very similar NDVI signature and thus the model could not distinguish between them and wheat. There is also the possibility of some vegetables being classified as wheat, but we were not able to obtain any confirmed vegetables plots to confirm if that is indeed the case.

Nonetheless, the error percentage is quite small and does not exist in all states. More importantly, while total annual Rice Acreage might be very big, the amount of rice grown during the dry season must be substantially lesser. Thus, we don't expect this to affect the wheat acreage results in a significant way.

3.2.2 UNDERPREDICTION

On the other hand, the design of the model ensures that it is only classifying the pixels which are very close to the wheat signature as wheat. This means that some of the wheat pixels might be missed as they might not have a very strong signal. Indeed, a lot of pixels were discarded from the training data which were flagged as wheat by NBS as shown in the NDVI profile of the rejected clusters.

To make sure that our model is not under predicting, we asked for feedback from field experts at FMAN and NBS on LGA wise acreage estimates. The estimates were close to the expected values by experts, and the few discrepancies which were flagged were in both directions (over prediction and under prediction). This gives us confidence that the model is not systematically under predicting.

Nonetheless, these estimates are best viewed in terms of relative values and not the absolute acreage predicted.

3.2.3 YIELD ESTIMATION

Yield distribution of LGAs at 13 State levels. Date of prediction [Feb-Mar 2021]

The yield of Wheat in Nigeria ranges between 2000 - 4900 Kg/Ha with standard deviation and mean of 852 - 1181 Kg/Ha and 2670- 4856 Kg/Ha respectively follows normal distribution curve over the area which shows that the yield has been very well distributed across LGAs in 13 states.

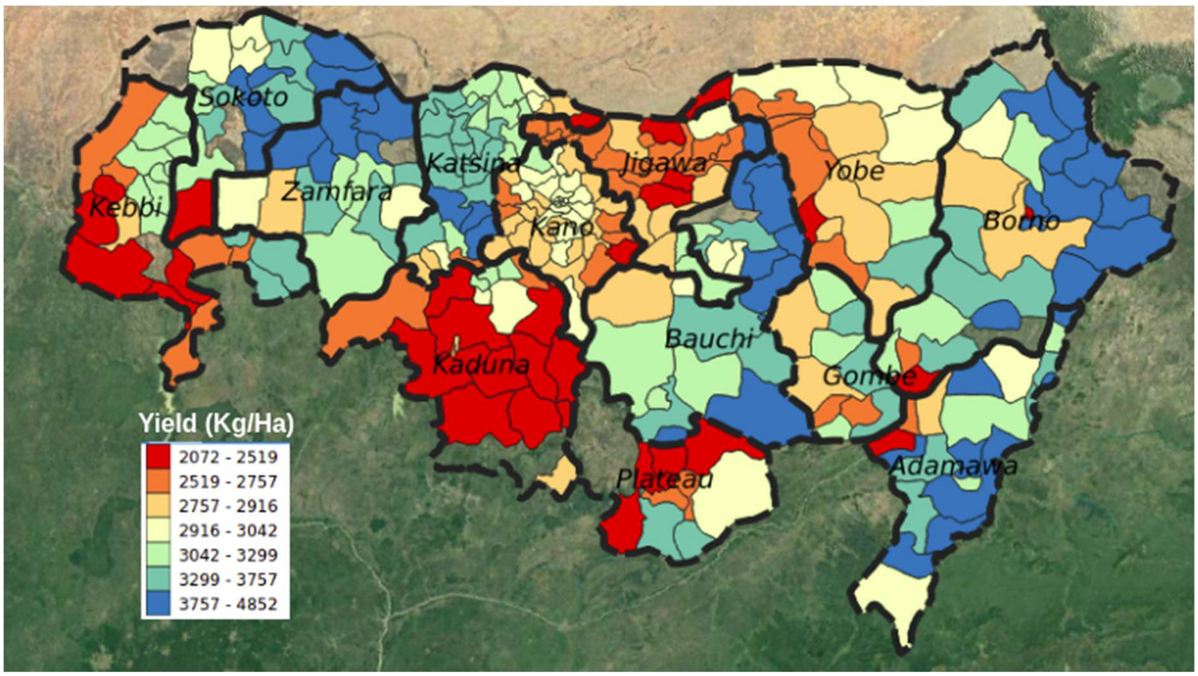
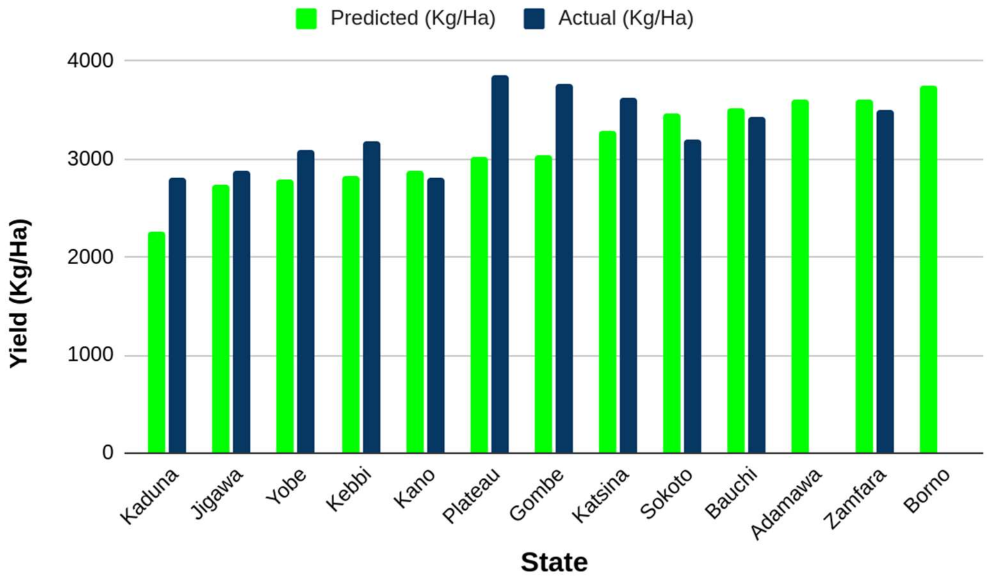


Figure 0: Bar - Plot of Aggregated Yield Mean of States [Predicted & Actual]



* Predicted - based on FAO model and Actual - NBS crop cuts

Table 9: Aggregated Yield Mean of State [Predicted & Actual]

State	Predicted (Kg/Ha)	Actual (Kg/Ha)	State	Predicted (Kg/Ha)	Actual (Kg/Ha)
Adamawa	3606	NA	Kano	2879	2803
Bauchi	3518	3431	Katsina	3294	3632
Borno	3755	NA	Kebbi	2821	3182
Gombe	3033	3770	Plateau	3022	3864
Jigawa	2734	2878	Sokoto	3462	3192
Kaduna	2265	2810	Yobe	2789	3087
			Zamfara	3608	3500

3.2.1 STATISTICAL ANALYSIS

The statistical test has an important role in any validation approach, in this study we used metrics such as MAPE and RMSE. In order to understand in a holistic way, the yield model results and yield data collected from the ground has to undergo validation processes at population level (i.e LGA/State level)

MAPE (Mean Absolute Percentage Error)

The mean absolute percentage error (MAPE) is a statistical measure of how accurate a forecast system is. It measures this accuracy as a percentage, and can be calculated as the average absolute percent error for each time period minus actual values divided by actual values. Where At is the actual value and Ft is the forecast value, this is given by:

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

The mean absolute percentage error (MAPE) is the most common measure used to forecast error, and works best if there are no extremes to the data (and no zeros).

Table 9.1. Interpretation of MAPE

MAPE	Interpretation
<10	Highly accurate forecasting
10-20	Good forecasting
20-50	Reasonable forecasting
>50	Inaccurate forecasting

Source: Lewis (1982, p. 40)

RMSE

Root mean squared error (RMSE) is the square root of the mean of the square of all of the errors. The use of RMSE is very common, and it is considered an excellent general-purpose error metric for numerical predictions

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2}$$

where Oi are the observations, Si predicted values of a variable, and n the number of observations available for analysis. RMSE is a good measure of accuracy, but only to compare prediction errors of different models or model configurations for a particular variable and not between variables, as it is scale-dependent.

Table 10: **Accuracy Analysis using RMSE and MAPE - State wise**

State	Total plot collected	No. of plots (after cleaning)	RMSE	MAPE (%)	Interpretation *
Kaduna	60	49	742.3	16.4	**Good**
Jigawa	60	58	712.32	15.0	**Good**
Yobe	60	43	735.15	15.6	**Good**
Kebbi	65	54	673.49	11.3	**Good**
Kano	60	51	683.68	12.0	**Good**
Plateau	59	16	747.72	19.2	**Good**
Gombe	60	38	741	18.2	**Good**
Katsina	60	54	684.11	11.3	**Good**
Sokoto	60	49	695.6	12.1	**Good**
Bauchi	55	40	706.81	12.4	**Good**
Zamfara	55	43	704	13.1	**Good**
	654	**495**	**711.47**	**14.23**	

Accuracy Analysis using RMSE and MAPE - State wise

*Source for interpretation - Table 9.1

Accuracy Analysis using RMSE and MAPE - Overall 11 states

Country	No. of LGA Mean Taken into analysis	RMSE	MAPE
Nigeria	32	711.47	14.23

3.4.1 Recommendation

Listing of Wheat farmers across the wheat producing states should be carried out to compile a frame for wheat farmers

There is need to investigate the time when wheat is grown and harvested across the wheat growing states, this will guide us appropriately in deciding the right time to conduct wheat survey.

3.4.2 Conclusion

It is very pertinent to state that the timing of the crop cut survey was a bit late as most farmers had already harvested their crops, however, the satellite mapping timing was not an issue as it covered data starting from October 2020. Additionally, the model built from the Satellite mapping can be further improved on in subsequent analyses considering a base model now exists which can guide other analyses.

Overall, the exercise was a huge success across the 13 states. It is also important to note that Adamawa State had finished harvesting before the commencement of the field work, thus, only area cultivated was measured for the wheat farms.

3.4.3 Challenges

Two LGAs where wheat is intensely cultivated were selected for coverage in all the states. On getting to the field, a different scenario was encountered which slightly differed from the methodological frame work. Most of the farmers have harvested their wheat while others were seriously harvesting. In some states, calls were put through to the Wheat Farmers Association executives to identify their members that had harvested their crops. This formed the bases for selection of farmers in those states instead of going through the enumeration areas (EAs) and the households (HHs)

As a result of the challenge outlined above, quick listing could not be carried out in the selected EAs. Quick listing was stopped, and number of wheat farms covered per state was increased.

Discussion on

1. Distance based model performs well in terms of predicting wheat only over farmlands. Moreover, it does not predict all farm plots to be wheat, which shows it can identify wheat signature as separate from most other crops
2. The model is sometimes (2-10%) not able to distinguish between rice and wheat grown in the same season. A further refinement of the model using SAR data can be considered in the next iteration.
3. Based on the statistical measures, it proves that yield prediction are with good accuracy (MAPE of 14.23 - Good Forecasting and RMSE of 711.47 kg/ha shows reasonable deviation from actual)
4. It is proposed that the CropIn yield model can be implemented for supplementing CCE points in different scenarios when CCE activity is not possible including flood affected areas, rapid harvesting due to harvesting machines which harvest large areas in a short span of time.

REFERENCES

Cropin Technology Solutions Pvt. Ltd : Crop Acreage and Yield report of wheat production in Nigeria, Aug 2021

USDA - Grain Production in Nigeria - https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Grain%20and%20Feed%20Annual_Lagos_Nigeria_03-15-2021.pdf

FAO - Wheat Production in Nigeria - http://www.fao.org/faostat/en/#data/QCL