Final Report

Researching efficient and affordable fertiliser products for increased and sustained yields in the maize belt of Nigeria

A cooperation between AfSIS Ltd., OCP-Africa and IITA

Sponsor/Donor: *OCP-Africa* S.A.

Executing institution: International Institute of Tropical Agriculture (IITA), Ibadan

Project code (IITA): PJ-002375

Partner institutions: AfSIS, *OCP-Africa*

National Partners: BUK, IAR, NAERLS

Author: E. Jeroen Huising

Date: October 4th, 2019

Acknowledgement and contributions

OCP-Africa is recognised for having launched, and technically and financially supported this project which was conducted with international partners and research and development organisations from Nigeria. This was quite an extraordinary undertaking in terms of the scale of the work, as well as in terms of the practical and integrated approach that was chosen. This refers to the soil characterisation using new spatial approaches and methodologies focussing on functional properties of the topsoil only and making use of new techniques for soil analysis and digital soil mapping, making it possible to do this exercise at the scale and within minimum amount of time. It also refers to having the design of new fertiliser formulations as part of the project and the validation of these fertilisers done at scale through the validation trials. These activities would normally have been done as separate projects and having them integrated into one project is unique. *OCP-Africa* has been actively and directly involved in the project, not only through the design and production of the fertilizers, but also through Mrs Bola Awotide, Idris Azare and others who were instrumental to the project. OCP sent representatives to attend workshops organised; Dr. Cisse was part of the management team that held meetings quite regularly, especially during the earlier phases of the project. The commitment and support to the project is much appreciated.

The involvement of AfSIS, has been crucial in providing the knowledge and know-how, the tools and techniques that made the soil characterisation possible. We like to acknowledge the contribution of Dr Markus Walsh and Alex Verlinden, who sadly passed away while the project was still being implemented. Without their in-kind contributions in terms of time spent on the project, the soil analyses conducted this project would not have been possible. We also like to acknowledge the contribution of Keith Shepherd from the ICRAF lab in this respect.

The IITA-Kano team has been responsible for organizing, coordinating and overseeing the field campaign for the soil characterisation and the field trials, organizing training workshops and meetings with the partners. The team has been able to adequately respond to critical developments, whether related to the FAM (Fall Army Worm) infestation for example, or to organisational and administrative challenges that are likely to arise for projects at this scale (e.g. facilitating the transport in the field, as well as communication for the field teams). We like to thank Dr Alpha Kamara and especially Reuben Solomon and Sadiq Bako. Without their support we would never have been able to make any progress with the project.

We would like to thank our national partners, BUK, IAR and NAERLS. Together with their colleagues from the ADPs they have made a tremendous effort in executing the soil survey and sample collection and in executing the validation trials. They travelled long distances, spent a lot of time in the field and made long hours to make sure everything was done correctly. And much of the work was done under the pressure of time.

We want to thank Julius Adewopo and Kamaludeen Tijjani, both stationed at IITA-Kano for their technical support.

At IITA we had the support from the various service units like Supply Chain (e.g. for clearing and transport of the fertilisers), PDAU and finance department. The Analytical Services Lab had a challenging task of processing the thousands of samples that came from the field, and Joseph Uponi together with his staff is acknowledged for that. The GIS unit by name of

Tunrayo Alabo was of great support and I also like to thank Sam Ofidele for further statistical support.
All persons below are acknowledged for their contributions:

IITA
Dr Alpha Kamara
Sadiq Bako
Reuben Solomon
Kamaludeen Tijjani
Dr. Julius Adewopo
Tunrayo Alabi
Mr. Sam Ofidele
Dr. Dries Roobroeck
Dr E. Jeroen Huising

OCP-Africa
Dr. Limamoulaye Cisse
Dr. Bola Amoke Awotide
Idris Azare

AfSIS Ltd.
Dr. Markus Walsh
Mr. Alex Verlinden (†)

NAERLS
Dr. Adamu A. Yakubu
Dr Chris Daudu

BUK
Prof Musa Ayuba Samndi
Prof Jibrin Jibrin

IAR
Prof Ado Yusuf
Abdulsalam Shobayo

Facilitators from the various ADPs
Sati Dachal
Ibrahim Y. Ogiri
Idris Umar
Yakubu Stanley
Abdulwahab A. Panti
Isaac Y. Mamman
Abdullahi M.Gabi
Ibrahim Shehu
Haruna B. Adamu
Mohammed Ndafogi
Hassan Bobi Saidu
Abdulkadir Ismaila
Habibu Nalado
Dio Tanko
Nehemiah Bala
Ahmad Abdullahi
Haggai Turba
Adamu Kunini
Timothy Ibrahim
Ventur Danvong
Dabang James K.
Sati Dalyop Dagai
Blessing Pyenya Wuyep
Isuwa G. Wenyang
Idris Gambo
Abdulmumuni Aminu
Abbas Ibrahim Bawa
Ganbo Linus Mbone
Alananao. Ibrahim
Nasir Abdullahi
Nuhu Manasseh
Candida N. Dagan
Gambo Haruna
Idris Harazumi
Ishaq Yakubu
Ishaq Abubakar Garba
Hussaini Abubakar
Kehindo Abdullahi
Abdullahi B. Mai Bago
Liman Umar Ahmed
Stephen Shuatlak
Nansa j. Pyelang
Yusuf h. Na ango
Mustapha A. Iliyasu
Suleman A. Rijau
Alanana M. Emmanuel
Yahaya A. Danjanka
Luka Kefas
Lawan Shirema
Saleh S. Umar
Jafar Adamu
Ibrahim Haruna
Muhammed Abubakar
Alkaleri
Ahmad Dalladi
Abubakar Mukail

Content

List of Tables

List of Figures

List of maps

Acronyms

AoO	Area of observation
ADP	Agricultural Development Programme
AfSIS	Africa Soil Information Service
BUK	Bayero University Kano
CV	Coefficient of Variation
DSM	Digital Soil Mapping
EthioSIS	Ethiopian Soil Information Servicer
FAO	Food and Agricultural Organisation of the United Nations
FAW	Fall Army Worm
FMS	Facility Management Services (IITA)
ICRAF	World Agroforestry Centre
IAR	Institute of Agricultural Research (Zaria)
IITA	International Institute of Tropical Agriculture
LGA	Local Government Area
MIR	Mid-Infrared
MSV	Maize Streak Virus
MT	Metric Tonne
NAERLS	National Agricultural Extension and Research Liaison Services
OCP	Office Chérefien des Phosphates
PDAU	Project Development and Administration Unit (IITA)
ROI	Region of Interest
SEM	Standard Error of the Mean
SOC	Soil Organic Carbon
SOP	Standard Operating Procedure
TAMASA	Taking Maize Agronomy to Scale in Africa
VT	Validation Trial
XRF	X-Ray Fluorescence

1. General Introduction and objectives

1.1 Introduction

Nigeria is the largest maize producer in Africa. There was an increasing trend in production from 2000 up to 2014/2015, recovering from a severe drop in production after 1995. Production in 2010 was not that much higher than the production in 1995. Production seems to have stabilized after 2015. Changes in production volumes are explained by the change in the area under maize production, rather than change in yield/ha (Table 1). We see an strong increase in the area of production in 2010 and 2011 and again in 2014 and 2015 to lesser degree, and has declined a little since. Yields have remained the same and the average yield level in 2010 was slightly higher than in 2017.
Nigeria imports a relative small amount of maize grain of around 100,000 to 200,000 MT yearly. Increase in production (and area under cultivation) is likely to be a response to international commodity prices, that does influence the price for the commodity on the national market. These production and yield figures are indicative, as it is notoriously difficult to establish accurate figures on production volumes and the area harvested. Nevertheless the figures do portray the actual trend.

*Table 1. Maize production and maize area harvested in Nigeria 2010-2017**

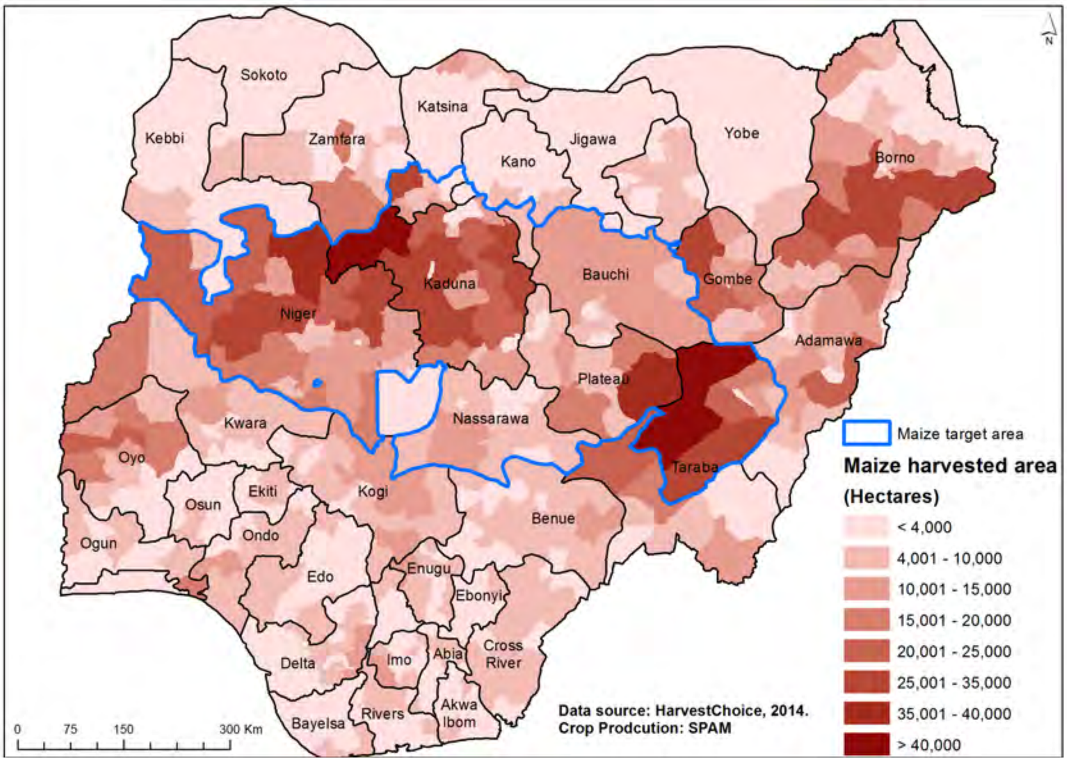
Year	**Production (x 1000 MT)**	**Area Harvested (x 1000 Ha)**	**Yield (t/ha)**
2010	7677	4149	1.85
2011	8878	5457	1.63
2012	8695	5751	1.51
2013	8423	5763	1.46
2014	10059	6347	1.58
2015	10562	6771	1.56
2016	10415	6601	1.58
2017	11000	6540	1.68

*source: http://www.indexmundi.com/agriculture

Maize is the third important food crop in Nigeria in terms of production after cassava and yam. According to FAO around 55% of maize produced is for food, 31% as feed and 12% is processed. According the figures of *Indexmundi* the percentage of maize used for feed is probably more around the 25%, which may indicate that the figures are not very accurate. The demand for maize is expected to increase with 3.2 per cent per year, due to a perspective growth of urbanization and population (FAO, 2013).

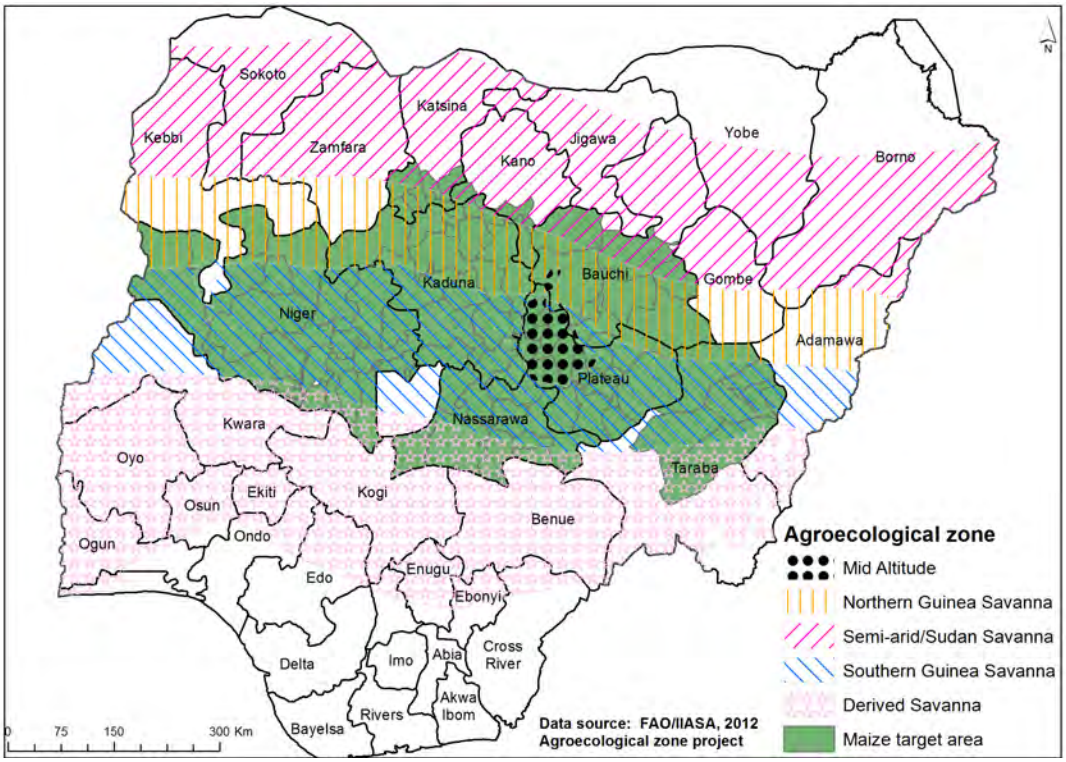
Maize is among the primary food staples, and is more widely and most frequently used in the preparation of traditional foods compared to other grains. Maize for feed is almost exclusively used for chicken feed. As far the industrial processing is concerned an important part is used in the production of beer.

North-Central region is the main maize producing area. Map 1 shows the location where most of the maize is cultivated (based on the area harvested by Local Government Area). It includes large parts of Borno and Adamawa state that were, however, not included in this project for security reasons. The Northern and Southern Guinea Savanna are the most conducive for the maize crop, but we still find much maize cultivated in the semi-arid Sudan Savanna and the Derived Savanna areas (See Map 2).



Data source: HarvestChoice, 2014.
Crop Prodcution: SPAM

Map 1. Maize harvested area per Local Government Area (2014)



Data source: FAO/IIASA, 2012
Agroecological zone project

Map 2. Agro-Ecological Zones of the maize growing area in Nigeria

Seventy percent of farmers are smallholders, with an average 5 ha area of cultivated land accounting for 90 percent of total farm input (NAIP, 2010). Maize in Nigeria is usually intercropped, with yam, cassava, guinea corn, rice, cowpea, groundnut, and soybeans. Average maize yield has increased to just above 2 t ha^{-1} in 2010 (FAOSTAT), but this masks the large variation in maize yield obtained in farmers' fields, which may range between 0.5 to 4.0 t ha-1, depending in part of whether fertiliser is used or not, but there is very little data available on this. The yield levels are well below attainable yields levels that can be obtained in farmers' fields with good agricultural practices and adequate use of inputs.

According to the FAO report maize farmers have received mainly disincentives to grown maize under the prevailing cost structure in the value chain, and despite input support polices that were in place. We do have to consider low yields as result of the poor response to fertiliser application as one of the underlying causes of the disincentives for farmers to grow maize.

Fertiliser use in Nigeria has steadily increased over the past four years, but is still below 20kg per hectare on average (http://data.worldbank.org/indicator/AG.CON.FERT.ZS). A large proportion of farmers that do not use any fertiliser explains the low average fertiliser use. Farmers, in general, seem to be quite receptive to the idea of using fertilisers, and the use of fertiliser is quite common among a large group of farmers. The farmers that do not use fertilisers may not have access to (subsidized) fertilisers, may not have the resources to buy fertilisers (at market price) or indeed consider the risks too high because of the limited response to the fertilisers that are available on the market. All in all this signifies that there is a huge potential for putting effective and price -competitive fertilisers on the market.

The types of fertiliser commonly used in Nigeria include urea, NPK, and Superphosphate (SSP). The most common NPK blends are: 15-15-15, 20-10-10, 12-12-17+2Mg0, and 25-10-10 (NSSP, 2010). We have little specific information on the fertiliser use by the various categories of farmers. We do have an indication, from a few studies conducted, that the response to NPK fertiliser is highly variable which seems to be related to the specific soil nutrient limitations and soil health constraints. This implies that to stimulate the use of fertiliser by the farmers the appropriate fertilisers need to be available to farmers that give an economic rate of return and not only serve to improve maize yields. We do have enough documented evidence that yields of 4 t/ha and above can be reached consistently with use of the right seed and application of the correct fertiliser blends (see The Guardian, 4 December 2015 for the story of Kaboji Farms)

1.2 Objectives

The main objective of the project is to investigate and develop new fertiliser products (formulations) that are cost competitive and efficient in increasing maize yields in a sustained manner for the Nigerian major maize production areas. The new fertiliser products need to give a marked improvement in crop response compared to the current available fertilisers in Nigeria and under the conditions that prevail on farmers' fields. That is, the fertilisers should adequately address existing nutrient limitations that prevail in the soils of the maize belt and it should perform better under the current agro-climatic conditions and with the commonly used agronomic practices. The new fertilisers should not be tested under highly controlled

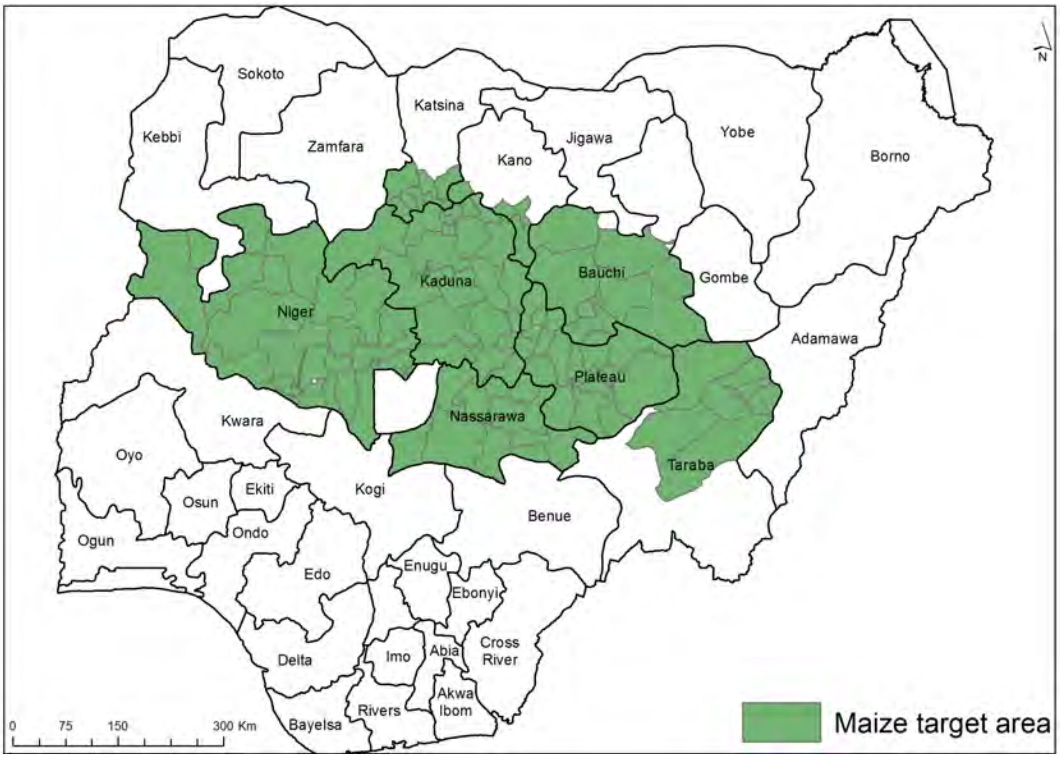
and optimum management conditions. The objective of this study is also not to investigate optimum fertiliser application rates.
The specific objectives are therefore:

1. Identify and characterize the soils of the main maize growing areas in Nigeria in terms of soil nutrient status and other soil parameters and identify possible limiting nutrients and other possible soil related crop production constraints;
2. Design and develop new fertiliser formulations to address the major nutrient limitations within the target area and with expected improved economic returns compared to the common available fertilisers and with commonly recommended applications rates.
3. Test and evaluate the new fertiliser products on farmers' fields in terms of response to standard recommended fertiliser application rates.
4. Identify the conditions under which the new developed fertiliser formulation perform best and provide guidelines for targeting of the application of the specific fertilisers.

2. Scope and methodological approach

2.1 Scope and general approach

To find alternative fertiliser solutions for maize production in Nigeria this project is implemented at scale. We focussed on the main maize growing areas, or the maize belt of Nigeria. We had to leave out the maize growing areas of Borno, Adamawa and Gombe states because of security concerns. The region of interest (ROI) is comprised of following states: Niger, Kaduna, Bauchi, Nasarawa and Plateau, and part of Katsina, Plateau and Taraba states. The area is depicted in Map 3 (see below). The area measures in total about 240,570 squared kilometres, of which approximately 212,642 km2 is cultivated land. The total population of this area is around 23.2 million (Afripop, 2014; http://www.worldpop.org.uk/).



Map 3 Maize production area targeted for this study

For the soil characterisation, as well as for validation of new fertilisers in the farmers' fields, we planned to cover the complete area. For the implementation of the activities we relied on our national partners (BUK, IAR, NAERLS), each providing a project leader and supervisor to supervise the various teams on the ground, with the team leader and the facilitators recruited generally from the Agricultural Development Programs (ADPs) from the Local Government Area (LGA), in general. For both the soil survey and the implementation of the validation trials (VTs) use was made of the same people as much as possible. While for the soil characterisation we could move around in cars to visit subsequent sample locations, for the implementation and execution of the trials we had to facilitate the facilitators by providing motorcycles to enable them to visit the fields in their LGA. The whole operation was coordinated by our IITA-Kano office.

The project necessarily adopted a phased approach, since the one activity could only start after the other had finished. The first phase of the project related to the soil characterisation exercise. The second phase was the design, production and shipping of the new fertiliser formulations. The third phase was the validation of the new fertilisers by conducting VTs in farmers' fields and the fourth phase was about compiling all the data and doing the data analysis.

2.2 Soil characterisation

The soil survey and sample collection started with a training conducted in February 2016 ([1] and [2]) and was to be completed with the presentation of the results of the laboratory analysis weeks before the start of the growing season to allow OCP to design, produce and ship the new fertilisers in time for the establishment of the trials planned for July. Carrying out the soil characterisation in such a limited period of time can only be done if the aim is to map soil functional properties, rather than to map soils (soil series), which would require description of soil profiles. The latter approach would be time consuming and would not provide information that is directly relevant for the purpose of the project, *viz.* to design new fertiliser formulations. On the other hand, it would make it more easy to look at the spatial distribution of the soils (soil series) within the area. In our approach we only sampled the top soil (0-20cm) and the subsoil (20-50cm) and the observations in the field are limited to the minimum (no recording of soil colour, texture, soil structure or other that are typically recorded in soil surveys. Also, no use is made of any prior soil information (e.g. existing maps, soil profile data, or other). The soil characterisation effort was directed to the cultivated land only.

The field work for soil characterisation was done in three-months' time. The analysis of the soil samples using spectral techniques could be done in time as well, but the confirmation of the results using wet chemistry took a while. This was mainly because we decided that both the spectral analysis and the wet chemistry analysis were to be done by both the IITA lab in Ibadan and the ICRAF lab in Nairobi for validation purposes. Certain type of analyses had to be done at ICRAF because of the specialized equipment they have (e.g. for the XRF measurements). The predicted soil parameter values based on the spectral analyses were good enough for Dr. Cisse from OCP to start working on the formulations for the new fertilisers.

Based on the predicted soil parameters for the top soil samples, OCP tried to get insight in the prevailing nutrient limitations, using critical values for soil nutrient concentrations and other soil properties like soil organic carbon, soil acidity, soil texture and others, reported in literature and used elsewhere (e.g. by Ethiopian Soil Information Service, EthioSIS) to which were added nutrient requirements for maize to yield around 5 t/ha of grains. Based on this information OCP developed two new fertiliser formulations they thought would best address the varying most limiting nutrient constraints of the soils in the area (we talk about nutrient constraint envelopes because nutrient constraints generally occur in combinations and in different ratios). See Appendix 2 for the formulations.

The soil data from this project, in combination with data from earlier campaigns (e.g. AfSIS sentinel site data from phase 1 and AfSIS-ISRIC soil legacy data), was used to map soil properties for the whole of Nigeria. This was done using digital soil mapping (DSM) techniques. The results were not present until 2017 when the validation trials were already established. The soil maps can be retrieved from Africasoils.net and the *AfSoilGrids250m* (https://www.isric.org/projects/soil-property-maps-africa-250-m-resolution, http://africasoils.net). This data is later used to evaluate the response to the fertiliser application in the VTs.

2.3. Approach to OCP fertiliser validation

The new fertiliser formulations were produced in OCP Morocco and shipped to Lagos, Nigeria. The fertilisers could be cleared from customs only after a considerable period of time and at a considerable expense, and far too late to be still used for the validation trials for that same year. The trials were subsequently planned for 2017 and this also gave us time to prepare for the establishment of the trials. The trials were planned to be conducted on 1500 locations randomly selected from the 3000 locations where the soil samples had been taken. The VTs were to be conducted on farmers' fields and we therefore had to confirm the proposed trials site locations whether these had conditions suitable for the trial and whether we had consent from the farmer. Alternative locations were identified when needed. The philosophy of the project was to validate the new fertilisers under the conditions that farmers experience in the field, and as such no specific requirements were set for the type of soil or soil condition. The selection of trial site location was governed by protocol. From the planned 1500 proposed trial site locations 1324 could be confirmed for the implementation of the trials, and in most cases these were not on the exact same location as were the soil sample was taken.

We conducted a training in May 2017 to instruct the supervisors, team leaders and facilitators on the protocol for the VTs [4]. Ample time was taken during that training to review the selected trials site locations [5]. Questions were raised if certain proposed trial site locations were not covered, or when the distance to the proposed trial site location was too large, or indeed on suitability of the location for conducting the trials. The fertiliser application rates specified in the protocol was based on the general recommended and commonly applied fertiliser rates of three bags per ha of the NPK fertiliser for basal application and two bags of urea per ha for top dressing, irrespective of the soil condition and in line with the objectives of the project.

The treatments were a Control treatment, and treatment with NPK15-15-15 (NPK triple 15) and two treatments for the OCP fertilisers, one for OCP-F1 and one for OCP-F2. The control treatment (no fertiliser application) was added to be able to determine the response to fertiliser application, which we consider and important evaluation criteria, also for economic evaluation. We also decided to include treatments for both OCP newly developed fertilisers for each trial, rather than just including one treatment with the one of the OCP fertilisers and then target that trial to a-priori defined areas or regions where we expect that specific fertiliser to do best. This decision was based on practical as well as theoretical considerations. Otherwise it would have required to know the specific soil conditions of the location where the trial is implemented and it would have assumed we know the conditions (or the soil related factors) that determine the response to the specific fertiliser.

We used the same maize variety throughout, in order not to introduce another factor (or another explanatory variable) that determines yield.

Furthermore, we adopted an unusual large plot size (compared to what is typical for research projects) to get representative results, considering the expected large within-plot variability as result of the less controlled conditions on the farmer-managed trial fields. It requires the distribution of larger amounts of inputs (seeds and fertilisers), which has logistical implication for getting the inputs to each of the trial sites. And that also required us to provide each of the facilitators with a motorbike to facilitate transport to and in the field.

The establishment of the trials in the field was, like all activities, guided by standard operating procedures. The land preparation and planting was mostly done by the farmer and overseen by the facilitator, but quite regularly the facilitators had to organise labourers for land preparation and planting, if one could not rely on the farmer to provide the labour or to provide it on time. The land preparation could either be by hand (how), animal-drawn plough or tractor-drawn plough, depending on the practice the farmer commonly deployed and depending on the services available. This resulted in some deviation from the specifications in the SOP, especially in relation to spacing of the rows (and consequently plot size) and orientation of the field. However, we included observations in the SOP and on the E-forms that would allow us to determine the situation in the field and later to make corrections to the data and adjustments in the calculation of the yield.
As reported by the supervisors and team leaders finally 928 trials were established in total [5]. The reasons why less trials were established than planned varied: VT sites very close to each other, or in the same field, were merged, poor accessibility and problems with security, some locations which as yet proved to be unsuitable, disputes over the ownership of the land, or where ownership had changed, and farmers that had withdrawn their consent. From the forms that were uploaded to the ODK server, we learned that 872 trials had been established, with the exact location known. (This thus can have been more in practice, but this is the number of trials for which the facilitators have been able to successfully upload the data.)

The management of the crop was also in principle left to the farmer, or defaulted to the facilitator in case the farmer did not follow up. This, likewise, had implications for the quality of the crop management (thinning, gapping, weed management). However, we assume the result from the trials are still valid, as long as the management practices were systematically applied to the different plots. We had an outbreak of fall army worm (FAW) in that season that required a coordinated response. The infestation of FAW was detected in an early stage and IITA-Kano station adequately responded in a timely manner by distributing the chemicals and applying it to all trials in the affected areas. There has not been a major impact on the results of the trials. The management of the trials greatly benefitted from the *Whatsapp* group that was set up and through which facilitators could raise questions to which they would get immediate answer from their supervisors, and through which also request and reminders could be put through to the facilitators. Nevertheless there were trials that were lost due to flooding, cattle feeding on the maize (fields lost due to Fulani herdsmen), army worm infestation and farmers have been reported to neglect the fields due to which the trials were lost.

Training of the supervisors and team leaders in the harvesting of the trials and data collection was conducted from 19-21 September 2017 [6]. Facilitators were not invited because this proved logistically difficult to organise. Moreover, the supervisors and team leaders were supposed to directly oversee the harvesting of all trials, in order the assure the quality of the data. Nevertheless, supervisors and team leader were still expected to instruct the facilitators in the field. The time of the training was planned just before the start of the actual harvesting of the trials in order to include the actual harvesting of one of the VTs in the training programme. The SOP for the harvesting and yield measurements included various observations that would allow for validation of the yield at harvest measurements in the field. (For example, photo were to be taken of each plot with a person holding a measuring rod for assessing plant height.) Part of the SOP was to take samples of the cobs harvested for further measurement in the lab, to determine moisture content of the grain, threshing percentage and number of grains amongst others. The data recording was, like in all cases, done using ODK forms and uploaded to the ODK server, through which the data could be easily accessed. All data is made available; see appendix 5 for a description of all data files.

Data from 705 VTs has been collected through the ODK forms, and this was after the data management workshop conducted in February 2018 at BUK in Kano [7], in which we helped the facilitators and supervisors to upload data that they had so far not been able to upload. Data from an additional 70 VTs had been recorded using the fieldbooks, but this lacked consistency (e.g. in relation to the identification of the trials in both the form for trial establishment and the harvesting of the trial) and reliability of the data was considered such that is could not be included in the final data set. Reasons why the number of trials harvested is less than the number of trials planted were already given earlier. In addition, there were quite a number of trials that had not been maintained by the farmer and therefor lost and there were trials that had already been harvested. The location of the 705 VTs that have been harvested can be found here, if you select the 'OCP VT Harvest 2018' map layer. Also, the maps with the proposed trial site locations and the map with the locations of the established trials can be found there. We see that the spread of the locations of the validation trials for which the data is available is quite good and only very few clusters, notably those to the west in Niger state, are not represented.

2.4 Processing and analyses of the data from the VTs

The data has undergone quite an elaborate process of quality control and data editing before being able to calculate the yield at harvest (see [7], [10]). Data recorded for the circular plot included number of plants harvested, number of cobs harvested, total weight of the cobs harvested, stover weight, number of rows harvested, distance between the rows in the circular plot, etc. It allows to do quality checks on the cob/plant ratio, the number of cobs harvested (in relation to the number of rows harvested), average cob weight, total weight of the cobs harvested (in comparison to the stover weight), and other. The data has been edited and adjustments have been made to the equations used for calculation of the yield based on the data recorded for each particular field. These equations have been retained in the files that were originally used for calculating the yield.

We believe the yield at harvest data for the C-plot is more accurate, because of the enhanced possibilities to check and correct the data and because of the smaller plot size through which likely less errors are made in the counting and weighing of the cobs. But it is less representative considering the smaller area it represents (10m^2).

For the total plot (T-plot) the cobs are harvested in batches, with the number of cobs and total weight of each batch being recorded. It allows for checking on the consistency between the average cob weight of the various batches harvested. It allows in some instances for correction of the data (of individual batches) it there are obvious typos and if corrections significantly improve the consistency.

Then there are 5 cobs sampled from the C-plot that are weighed separately and that are sent to the lab where they are weighed again. The average weight as determined in the lab is used as reference for the measurements in the field and provides an opportunity to for an additional check on consistency. This also allows to determine whether there are errors made in the unit of measurement and scaling of the data and to make corrections based on the assumed unit of measurement. If the data is being scaled this needs to be done in the same way for each plot in that particular trial, such that the (relative) difference in the performance of each treatment can be correctly assessed. Furthermore, checks are done on the average cob weight, that must be within acceptable limits. If the number of cobs harvested is too low, the results are considered unreliable and the data is rejected. Also, the ratio of the number of cobs harvested to what is expected is calculated and if these differ too much between the treatments/plot within that particular trial it may be a reason to reject the data for that trial all together, or for one of the treatments.

The files with the workings as part of the quality control process have been saved, with as much as possible annotations added to the data that has been edited. Also, the files show the calculations of each of the yield variables for all plots. With the quality control completed we remain with 600 trials for which we have reliable data, though not always complete records.

The results are presented as 'yield at harvest', which is based on the weight of the cobs harvested for the C-plot or T-plot, without making corrections for the shelling percentage or moisture content. For the purpose of evaluating yield difference between the different treatment plots, this is most adequate. At the time of harvesting the grain was dry and moisture content proved to be less than the 12.5% or 13.5% that is generally taken as reference for calculating the dry weight. Converting 'yield at harvest' to grain yield for each plot, based on the observed moisture content and shelling percentage of just a small sample of the cobs harvested in the field, may introduce additional variation and error. Rather, to get an indication of the grain yield the data can be converted using the average shelling percentage obtained for the particular region of the for the whole data set. We have calculated the yield adjusted for the number of plants, or the number of cobs, harvested. We observe that the number of plants harvested may vary strongly between the plots within a trial and between the trials and may affect the outcome of the trials. The 'adjusted yield' is the yield that would have been obtained with an assumed 53,333 plants harvested per ha (which would have been the number of plants harvested if plant densities are observed as specified in the protocol, and no plants would have been lost). The 'adjusted yield' actually reflects the average weight of the cobs harvested, and it has therefor a different quality and stands as a variable on its own.

We present the results for the various subregions, as they were serviced by the various teams, separately. This is done because we do not want the additional variation that is associated with the different teams that implemented the trials to confound the analysis of the treatment effects. The difference in timing of the activities (like the time of establishment of the trials) and consequently the difference agro-climatic conditions, as well as the difference in the diligence with which the protocol for the VT was implemented and in the management of the trials will have an effect on the results of the trials. The areas that were serviced by the various teams is indicated in the Table 2 below. We distinguish between the various teams of our national partners (IAR, NAERLS and BUK) and between the three teams operated by BUK. NAERLS operated four teams, two teams for Plateau state and one team for Nasarawa and one for Taraba state, but it proved not to be necessary to make further distinction between these teams. The same applies to IAR that operated two teams.

Table 2. State and Local Government Area serviced by the various teams for the implementation of the Validation Trials

Table 2 The teams and areas they were operating for the implementation of the VTs

Team	State	Local Government Area (LGA)
IAR TEAM	Niger	Rafi, Kontagora, Mariga, Katcha, Wushishi, Paikoro, Bosso, Mariga, Lavun, Mashegu
	Kaduna (part)	Birnin Gwari, Kachia, Kauru, Kubau, Kagarko, Kajuru, Sanga, Zango Kataf
NAERLS TEAM	Nasarawa	Karu, Lafia, Akwanga, Keana, Obi
	Taraba	Ardo Kola, Karim-Lamido, Lau, Yorro
	Plateau	Barkin Ladi, Jos East, Kanam, Langtang North, Langtang South, Mangu, Riyom, Qua'an Pan, Shendam
BUK TEAM 1	Katsina	Kafur, Faskari
	Kaduna	Giwa
BUK TEAM 2	Kaduna	Lere, Igabi, Ikara
	Kano	Rogo, Doguwa
BUK TEAM 3	Bauchi	Toro, Bauchi, Tafawa Balewa, Ningi, Ganjuwa, Alkaleri

For the analysis of the data we in first instance concentrated on evaluating the difference in response to the various fertiliser treatment, and then especially on the performance of the two new OCP fertiliser formulations. In second instance we looked at the soil parameters that possibly determine the response to either of the two fertilisers and to investigate the spatial pattern in the response to each of the two OCP fertilisers, that would help in targeting the application of either of the two fertilisers to areas where they perform best. The data allows for further analysis, like of the nutrient use efficiencies for examples, but this falls outside the scope of the project.

3. Soil characterisation

3.1 Sampling frame

The main geographical Region of Interest (ROI) for this survey was identified on the basis of 133 Local Government Areas (LGAs) that are thought to represent the main maize-producing region of central Nigeria, excluding those LGAs that present security concerns. The sampling frame (Map 4) combines the LGAs with a 1 km^2 resolution cropland area map developed by AfSIS in 2015 using observations from high-resolution satellite images in GeoSurvey. It covers an area of 239,504 km^2.

Map 4 Central Nigeria ROI (239,504 km2) and sampling plan. Areas rendered in dark green are high probability croplands based on GeoSurvey model results. Areas in grey have a low probability. Red crosses indicate the planned sampling locations for the current c

For this study we used a stratified multistage sampling approach on a 1 km^2 resolution Discrete Global Grid (DGG) with a Lambert Azimuthal Equal Area (LAEA) Coordinate Reference System. The main reason for adopting a stratified sampling approach over a simple random sampling approach is the considerable improvement in the efficiency of the survey, both in time spent as well as in the number of kilometres that you need to travel. "Cropland sentinel sites" were defined as 10x10 km^2 areas that contained more than 80/100, 1 km^2 cropland "sites". Selection of 3,000 sample "locations" was accomplished in stages as follows:

- Stage 1: Initially, 60/133 preselected Local Government Areas (LGA's) were selected at random, stipulating that they contained more than 3 unique cropland sentinel sites.
- Stage 2: In each of the 60 selected LGA's, a single 10x10 km sentinel site was selected at random from the available set of potential sentinel sites.
- Stage 3: In each of the 60 selected sentinel sites 10, 1x1 km cropland grid cells were then selected at random. These are the red crosses in Map 4.

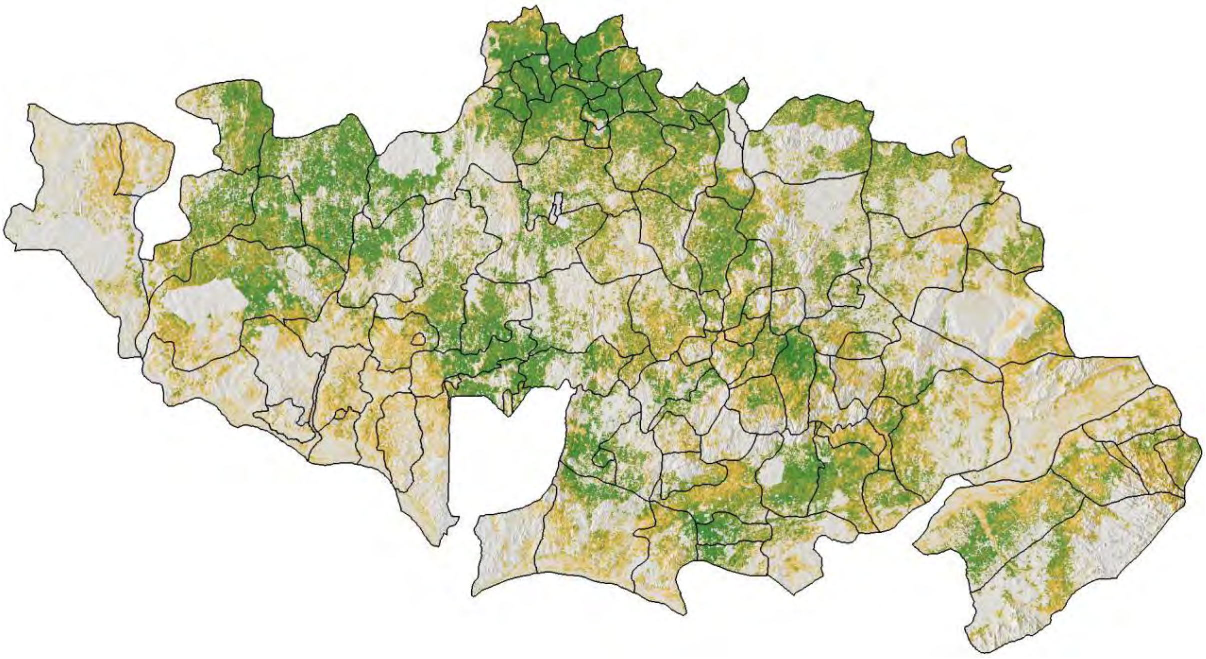
- Stage 4: The geographical centroid of each selected 1x1 km cropland grid cell was used to randomly sample 5/100, 100x100 m grid cells resulting in a total of 3,000 unique, geo-referenced survey locations.

The R-script and grid layers for randomization and location selections are available here.

3.2 Cropland and maize distribution prediction

The field teams observed the crops being grown at each survey location. The presence of the main cereal, legume, root and tuber crops and livestock were recorded as checklists (presence/absence). The standard AfSIS operating procedures for crop distribution surveys (Crop Scout) are available here. We used GeoSurvey to record observations of the occurrence of croplands, buildings and woody vegetation cover based on high-resolution satellite images, which resulted in an additional 5,000 GeoSurvey observations within the ROI on a 100 m resolution, randomized sampling grid. The 2,916 sampled MobileSurvey locations were also used in GeoSurvey to obtain training data about cropland building/rural settlements that are co-located with geo-referenced ground observations for validation.

Based on our GeoSurvey observations, cropland occurs in ~56.9% (56 – 58%, 95% CI) of our 1 ha pixel observations. If we assume that on average cropland actually occupies between 0.71 – 0.81 ha of any given cropland positive hectare in GeoSurvey, the total cropland area for this ROI would fall somewhere between 16.9 – 19.4 Mha of land which could be targeted for the use of the improved fertilizers. These estimates are supported by spatial predictions of croplands based on GeoSurvey and AfricaGrids data. We have an excellent agreement between our spatial predictions and an independent validation sample (AUC = 0.92).



Map 5 Joint Maize & Sorghum distribution probability map for central Nigeria based on 2,916 MobileSurvey observations. Areas rendered in dark green have a high probability of being classified as having Maize and/or cassava present based on ensemble model results. Areas rendered in grey have a low probability. Polygons in black outlie delineated LGA’s

Spatial predictions for the distribution of Maize (and those for other crops) are more complicated than those for croplands. Based on the aggregated, sentinel site-level summaries for MobileSurvey cropland and Maize area proportions, the prediction model could be adjusted. Based on these calculations, we estimate that Maize occurs on ~12.5% (8.6 – 18.1%, 95% CI, eqv. to ~2.0 – 4.4 Mha) of the area of this region of interest, which would be in agreement with the data on maize area harvested presented in Table 1. The most prevalent cereal crop in this environment is Sorghum, which by our estimates occupies ~32.2% (25.7 – 40.2%, 95% CI, eqv. to ~6.2 – 9.6 Mha). It should be noted that the maize area might change substantially from year-to-year in this environment, as input prices for fertiliser and output prices for maize are said to be volatile.

What we can differentiate at present is the joint distribution of Maize and/or Cassava from other cropland areas. For further details on the methods, please see the original report on the soil nutrient status of the croplands of central Nigeria [3]

3.3 MIR prediction of soil properties

The AfSIS Bruker-Alpha KBr based spectral models were used to predict individual Mehlich-3 extractable nutrient contents as well as pH, Hp, EC, SOC and total N of all available soil samples. Examples of the results for nutrient predictions for topsoil (0-20 cm) samples relevant for determining the fertiliser formulations under consideration by OCP are shown in Figure 8.

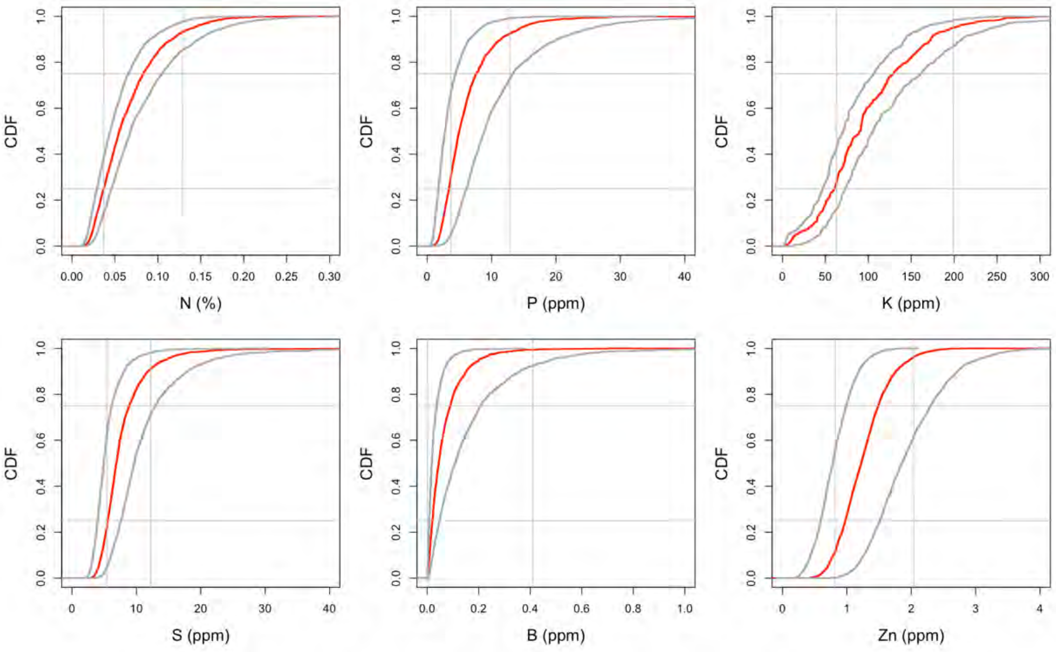


Figure 1 Cumulative distribution functions (CDFs) of total N and Mehlich-3 extractable P, K, S, B & Zn topsoil nutrients predicted from MIR. The red curve shows the CDF of the median prediction. The grey curves show the CDF values at their 5 and 95% credible value

For each of the predicted soil properties we have defined low, medium and high reference ranges based on the 25th and 75th percentiles of the relevant AfSIS data for Africa. These are the areas demarcated by the grey horizontal and vertical lines in Figure 1. Table 3 provides a summary of the AfSIS Mehlich-3 reference levels (low < 25th or high > 75th percentiles) and the geographically weighted regression (GWR, Fotheringham et. al., 2002) estimates of the percentage of the topsoil samples from central Nigeria that fall above or below those levels. (The R-code for calculating these values, producing the cumulative distribution graphs and the GWR regressions is available here.) The results show that the majority of samples fall into the medium AfSIS reference range for all of the predicted soil properties. We also note that by simply predicting individual soil properties that we are ignoring the inherently compositional nature of the data (van den Boogart & Delgado, 2013).

Table 3 AfSIS reference levels for African (0-20 cm) topsoils and geographically weighted regression (GGWR) estimates for the percentage of topsoil samples falling above or below those levels in central Nigeria based on MIR predictions.

	Reference levels		**% of samples**	
Soil property	*low*	*high*	*< low*	*> high*
pH (in H_20)	5.6	6.7	9.5	4.1
EC (dS/m)	37	64	7.2	2.9
Organic C (ppm)	5,160	17,150	19	2.7
Total N (ppm)	370	1,285	20	5.2
M3-B (ppm)	0.001	0.41	0	1.0
M3-Mg (ppm)	86	385	14	4.5
M3-P (ppm)	3.7	13	31	7.1
M3-S (ppm)	5.4	12	22	7.6
M3-K (ppm)	63	200	19	3.6
M3-Ca (ppm)	343	1,710	21	2.3
M3-Mn (ppm)	30	150	14	1.6
M3-Fe (ppm)	63	151	1.8	32
M3-Cu (ppm)	0.50	2.4	23	5.3
M3-Zn (ppm)	0.82	2.0	16	8.4
M3-Al (ppm)	446	1,050	18	1.7
M3-Na (ppm)	21	49	7.5	16
M3-Hp (ppm)	0.14	0.56	1	1

3.4 Determining probability of nutrient limitations of soil in central Nigeria

When applying critical limits to the predicted soil nutrient values, we can determine the probability of the nutrient limitations occurring within the central maize growing region in Nigeria. For the critical nutrient values, we used those that have been adopted by EthioSIS. Typically, you would want to have the critical values that apply to the region of interest and that are confirmed by experimental data, but these do not exist for Nigeria. We have made some few changes to adjust the criteria to apply to the units of measurement used in our data set and to adjust the class definitions in a few cases. For example, we added a class for pH less than 4.5. as ‘Very low’ rather than calling everything below pH 5.5 ‘strongly acidic’. Otherwise we do not have major problems with the critical values and levels adopted by EthioSIS.

In the data file with the predicted values for all the topsoil samples we have colour-coded the values for each soil parameter according to the status or class they belong to, to give a quick impression (by visual interpretation) of the status of the soil nutrients. How that looks is shown in Table 4, in which some rows of that very large table are displayed. In Table 4 we also present the summary statistics. For the nutrients it shows that especially P, B, Zn, K and S are critical, apart from N. Also, Mg and Na are very low. The SOC% is critically low in 86% of the cases. We would have preferred 0.7% as the lower critical limit and 1.5% as the limit for low SOC% rather than 1.2% and 1.75% used by EthioSIS. In that case we would have 46% being rated ‘very low’ and 48% being ‘low’ and 6% rated as ‘adequate’. It, nevertheless, shows that SOC is a general concern and one of the major constraints to improving production in general.

Based on these observations OCP decided to develop two new fertiliser formulations designed to provide balanced plant nutrients and alleviate nutrient limitations. Both contain elevated P_2O_5 levels compared to ‘NPK triple 15’ and have S, B and Zn. One of the new formulations (OCP-F1) has K added in contrast to the second formulation (OCP-F2) that does not have any potassium added. The exact formulations are presented in Appendix 2.

3.5 Spatial distribution of the soil parameters

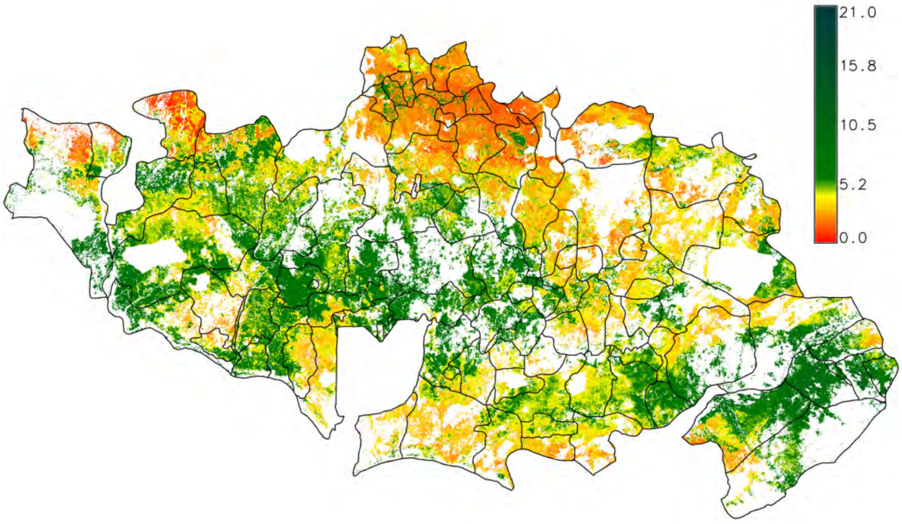
The soil data collected for this project has greatly contributed to the mapping of soil properties for Nigeria and to the soil property maps of Africa of 250m resolution. The data for Nigeria is obtained from the five sentinel sites surveyed during AfSIS phase 1 (800 observations), 1200 observations from the legacy data obtained during the AFSIS project and 2100 observation points from this project. Information about the methods used for producing these maps, specifications as well as access to downloads can be obtained from the ISRIC website (https://www.isric.org/projects/soil-property-maps-africa-250-m-resolution). Information on the map accuracy (RMSE of the difference between the predicted and observed values) indicates that this is limited, but these may have greatly improved for Nigeria with the later updates based on the inclusion of the data from this project (high point density and much better spread of the OCP data), but we do not have the exact figures.

The maps below show the distribution of Available Zinc (Zn-av) in ppm and of Available Boron (B-av) in ppm. The data is extracted for the ROI and shows data only for the cultivated land. The range of the data values in the map is larger than the range of the data values for the predicted available Zn we have from this project, but it shows a clear spatial pattern in the distribution of the Zn-av. Extreme low values occur within the northern part of the ROI in particular.

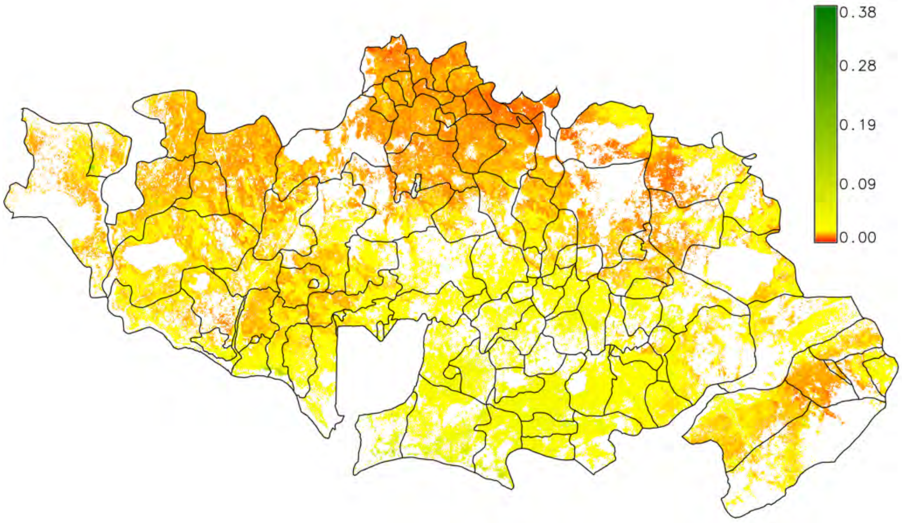
For Boron, the map shows a smaller range of values compared to the results we have for the predicted values of the topsoil samples. However, it confirms that B is within the critically low range for the whole of the ROI, with the tendency of the lowest values occurring towards the north of the ROI.

Table 4. Example of rating nutrients status of individual topsoil samples and the percent distribution of the sufficiency classes for the various pant nutrients, SOC and N based on the predicted values from the MIR scanning for the 2804 topsoil samples

Barcode	C%	Al (ppm)	B (ppm)	Ca (ppm)	Cu (ppm)	Fe (ppm)	K (ppm)	Mg (ppm)	Mn (ppm)	Na (ppm)	P (ppm)	S (ppm)	Zn (ppm)	pH	N%	Ecs	Hp
NG-ODK-PL-udHFUAm0	0.75	348	0.090	139.9	1.938	189.9	27.5	75.6	60.3	17.07	3.13	9.43	0.739	6.11	0.070	68.8	0.418
NG-ODK-PL-VGmhdxVP	0.50	264	0.080	114.9	0.756	115.6	85.6	41.0	35.0	17.63	7.66	7.10	0.691	6.29	0.043	97.5	0.236
NG-ODK-PL-vniLQMFq	1.52	1050	0.113	162.0	1.106	70.6	61.7	236.1	19.2	11.38	4.74	13.93	0.748	5.56	0.100	81.9	0.474
NG-ODK-PL-Vt17Tydt	0.90	703	0.194	561.0	1.468	148.4	95.6	159.0	103.2	13.35	2.93	9.30	1.226	5.78	0.089	86.5	0.284
NG-ODK-PL-w1PLIF5Z	0.90	550	0.091	26.7	1.763	158.6	56.0	52.8	15.3	11.55	7.31	13.01	0.577	5.82	0.093	51.7	0.313
NG-ODK-PL-W4br8DTP	1.32	1404	0.169	215.2	1.835	119.4	131.3	384.4	53.1	16.56	2.96	13.57	1.042	5.31	0.098	59.1	0.669
NG-ODK-PL-w8JDpip9	0.90	659	0.141	215.8	1.329	154.5	57.0	75.8	21.0	18.22	2.83	10.72	0.807	5.46	0.069	45.0	0.499
NG-ODK-PL-wEwg9Vzv	0.55	302	0.030	169.2	0.076	78.1	51.6	51.5	31.3	15.31	5.99	5.70	0.774	6.64	0.030	40.7	0.258
NG-ODK-PL-wRSwxmg6	1.55	692	0.077	52.4	1.177	123.6	14.6	37.0	22.1	11.88	2.08	22.02	0.673	5.15	0.124	47.5	0.743
NG-ODK-PL-xdA5Oakc	0.89	284	0.070	207.4	0.256	83.4	66.9	113.0	48.8	16.75	9.43	6.47	1.072	6.72	0.048	60.4	0.269
NG-ODK-PL-xq6vQ0X0	0.98	401	0.055	361.3	0.526	149.4	65.9	83.9	44.5	9.48	5.77	9.30	0.886	6.03	0.072	76.2	0.311
NG-ODK-PL-XtSPHocl	1.21	792	0.116	181.1	1.015	127.7	87.5	106.1	67.8	10.86	4.83	8.08	1.283	5.50	0.093	64.4	0.348
NG-ODK-PL-xzxRCOk9	2.30	676	0.133	526.3	1.413	198.4	145.4	105.2	23.6	0.47	5.18	31.70	1.165	4.93	0.146	97.8	0.674
NG-ODK-PL-YTYwqe3U	1.38	1158	0.183	343.3	1.941	108.2	127.3	198.8	51.2	11.83	3.09	15.30	0.621	5.52	0.144	59.3	0.514
NG-ODK-PL-yVSQIBSG	2.15	909	0.163	692.9	1.221	102.4	71.0	247.7	41.6	24.24	1.62	18.39	0.809	4.97	0.167	84.8	0.469
NG-ODK-PL-zBYXazJq	1.01	706	0.110	58.7	2.073	157.9	18.1	52.0	44.1	6.81	1.53	19.21	0.613	5.15	0.093	67.1	0.894
NG-ODK-PL-zeLg8ip5	1.89	1710	0.198	266.2	1.584	61.0	118.8	93.3	46.2	28.39	1.83	14.38	0.801	5.16	0.135	72.8	0.676
NG-ODK-PL-ZomKgfLB	2.10	1105	0.161	392.5	2.260	137.5	173.5	64.3	41.2	17.59	2.80	15.00	0.739	5.10	0.124	77.4	0.712
Very low	86%		99%	35%	24%	4%	75%	89%	44%	100%	100%	65%	36%	0%	83%		
Low	11%		1%	60%	28%	7%	25%	11%	41%	0%	0%	33%	49%	11%	13%		
Adequate	3%	0%	0%	5%	47%	89%	0%	0%	15%	0%	0%	2%	15%	89%	3%		
High	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
Very high	0%	100%			0%	0%	0%		0%		0%	0%	0%		0%		



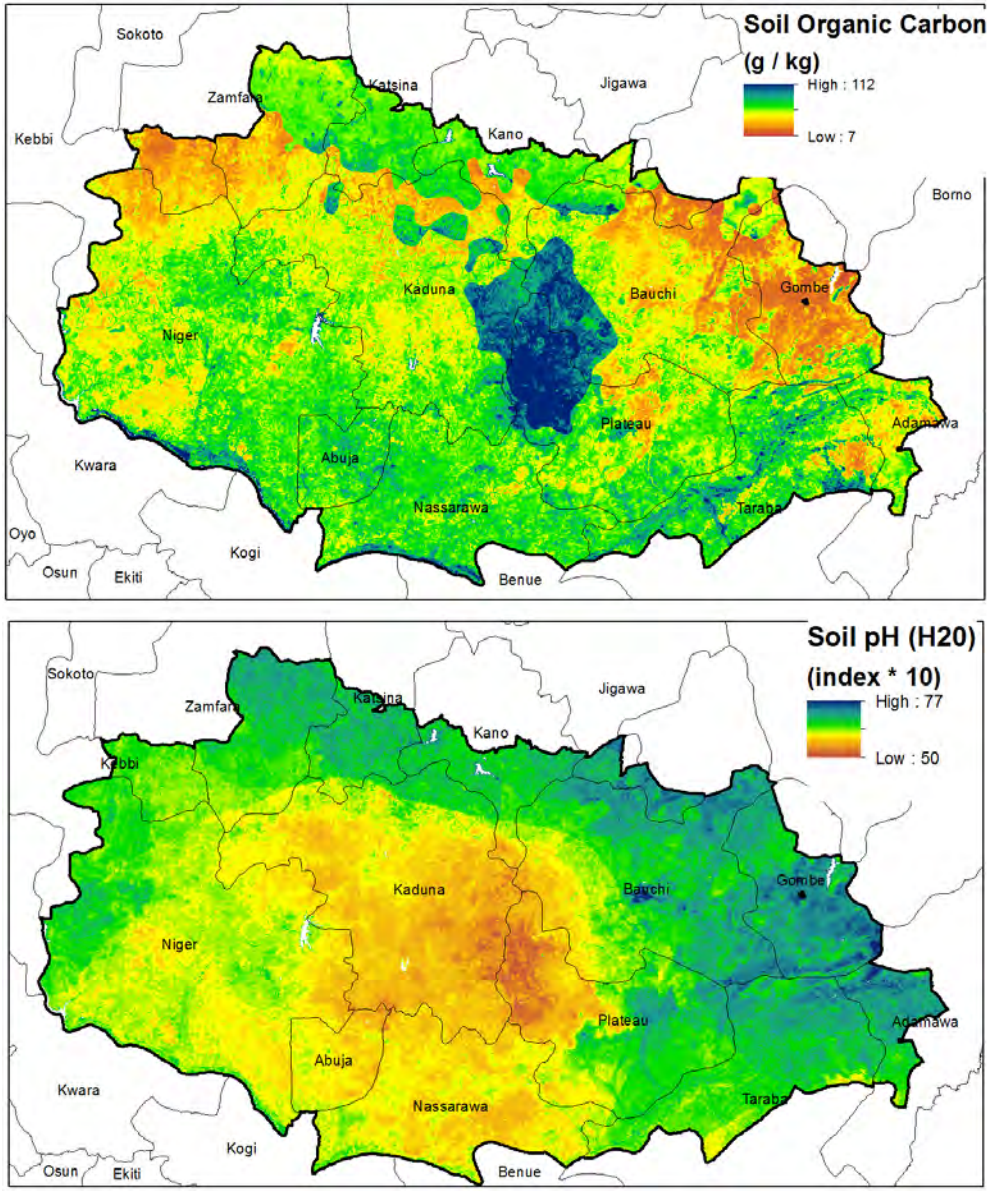
Map 6. Soil property map of Zinc (ppm) in 250m resolution for the ROI with the non-cultivated areas masked.



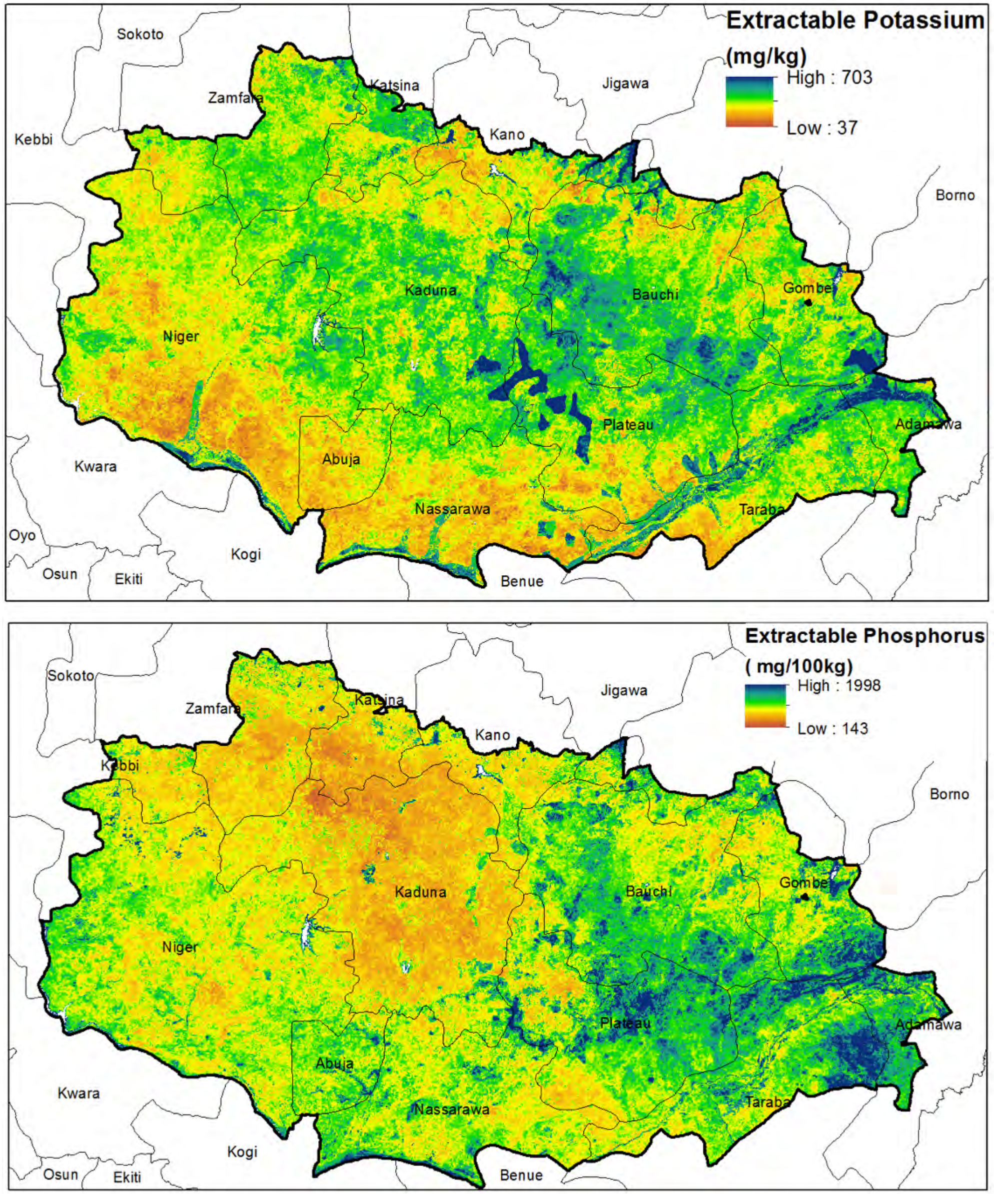
Map 7. Soil property map of B (ppm) for the cultivated land within the ROI.

Below the maps are presented for available P (ppm) and exchangeable K (ppm). The available P seems to be within the range of values we have for the predicted available P of the topsoil samples within the ROI. For our ROI the values are critically low, but we see some differentiation with relatively high values in the southern part of Kaduna and norther part of Plateau state.

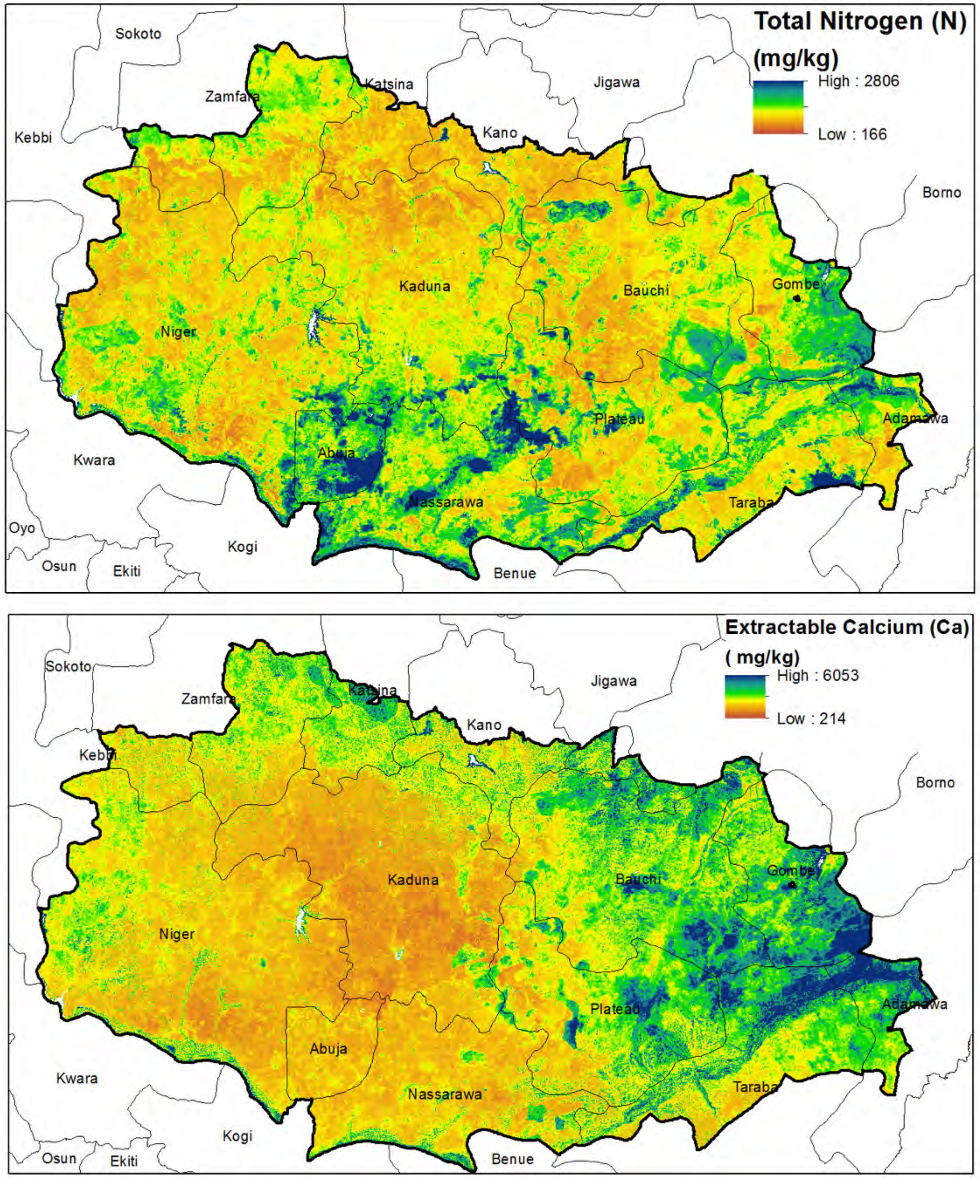
For exchangeable K the situation is similar. The range of values on the soil property map for the ROI is the same as the range of values we observe for the predicted available K (ppm) for the topsoil samples from the region. All values are rates ‘low’ or ‘very low’, but we see a clear spatial pattern with the higher values concentrated in the northern part of Plateau, the southern central part of Plateau state and some areas of Nasarawa state and further east in the floodplains of the Benue river.



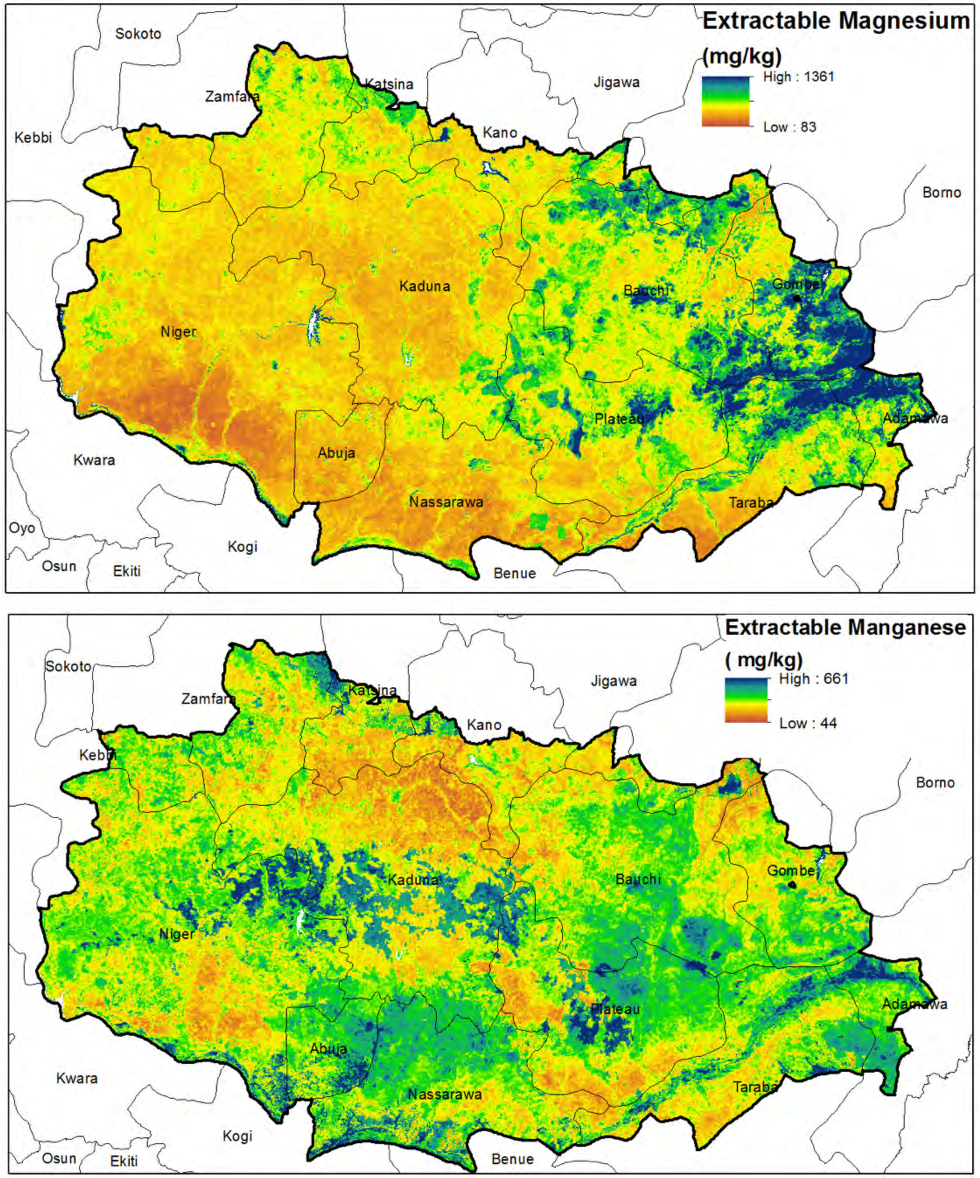
Map 7 Soil property map of Soil Organic Carbon and pH extracted for the ROI from the Africa SoilGrids 250m data



Map 6 Soil property map of extractable potassium and extractable phosphorus for the ROI extracted from the Africa SoilGrids 250m data



Map 9 Soil property map of the ROI for total Nitrogen and extractable Calcium extracted from the Africa Soil Grids 250m data



Map 8 Soil properties maps of extractable Magnesium and extractable Manganese for the ROI, extracted from the Africa SoilGrid 250m data

4. Validation of the new OCP fertiliser formulations

4.1 Validation trial site locations

We explained above the procedure for the selection of the trial sites. The 1500 original proposed trials site locations were taken from the 3000 soil sample locations determined by the stratified multistage sampling approach, in such way that all clusters would be represented by at least on VT, assuring a proper spread of the locations of the VT throughout the ROI. The teams went to the field to determine the location in the field and get the consent from the farmer. Sites for which consent from the farmer had been obtained in 2016 had to be revisited in 2017 to get reconfirmation. These VT locations were reviewed and evaluated based on a few criteria, like how close they are to the PTS locations if the PTS location is actually within a suitable cropland area. It sometimes still required adjustment of the location of the VT at the time of the establishment of the trial. The information on the PTS locations, the locations for which the consent of the farmer was obtained and the final location of the 705 VTs that have been harvested can be found here. The maps allow to follow the process of site selection and to verify that the ultimate locations of the VT is generally very close to the location of the PTS (check “Proposed trial site” to display the locations of the PTS and check “OCP VT Harvest 2018” map layer to display the actual locations for the VTs; zoom in to verify the distance between the VT locations and the corresponding PTS location). The criteria used was that the VT location should be within a circle of 200m radius from the PTS, if the conditions are the same (as determined from inspecting the satellite imagery). Otherwise, the preferred distance should be within 100m form the PTS location. The VT locations that do not fulfil that criteria are considered new points. The facilitators were instructed to take soil samples at the new locations, such that we could have soil properties determined from soil samples for each of the VTs, if that would be needed.

4.2 Response to OCP-fertiliser application

Statistical analysis of the yield at harvest for all the trials shows that there is significant difference between the states as well as the treatments. Results of the Duncan’s multiple range test differ slightly between the analyses for the C-plot and the T-plot yield. But considering the results for the mean ‘yield at harvest’ for both methods shows that Kano stands out, Niger stands out, Kaduna stands out (though Kaduna groups with Kano for the T-plot yield); Plateau, Nasarawa and Taraba stand out as group, and for Bauchi and Katsina we do not find significant differences. This strongly suggest that the teams that we have been operating (see Table2) is an important factor in determining the yield. This is expected because of the difference in the implementation and management of the trials (e.g. time of establishment of the trials). We therefore present the data and results of the analysis for the different teams separately, for better interpretation of the results.

Figure 1 and Figure 2 present the average ‘yield at harvest’ for the fertiliser treatments for the various regions serviced by the different teams. Figure 2 presents the results based on data obtained for the whole plot (T-plot), whereas Figure 3 presents the results based on the data obtained for the 10m^2 sub-plot, or the ‘circular plot’. The error bars represent the standard error of the mean (SEM). The corresponding data is presented in Appendix 3, which includes the number of observations (number of trials), the standard deviation, minimum and maximum observed values and the coefficient of variation (CV).

Results for the region coordinated by BUK is split in three, corresponding to the mandate areas for the three teams operating under BUK (Katsina-Kaduna, Kano-Kaduna and Bauchi). The number of observations is therefore relatively low compared to the number of observations for the Plateau-Nasarawa-Taraba and the Niger-Kaduna region. The standard error (SE) and coefficient of variation (CV) are high especially for Bauchi, indicating that there have been some problems with the trials in that region. It is also for this data set that a relatively high number of records were discarded during the process of quality control. There seems to be a relatively high number of trials that failed.

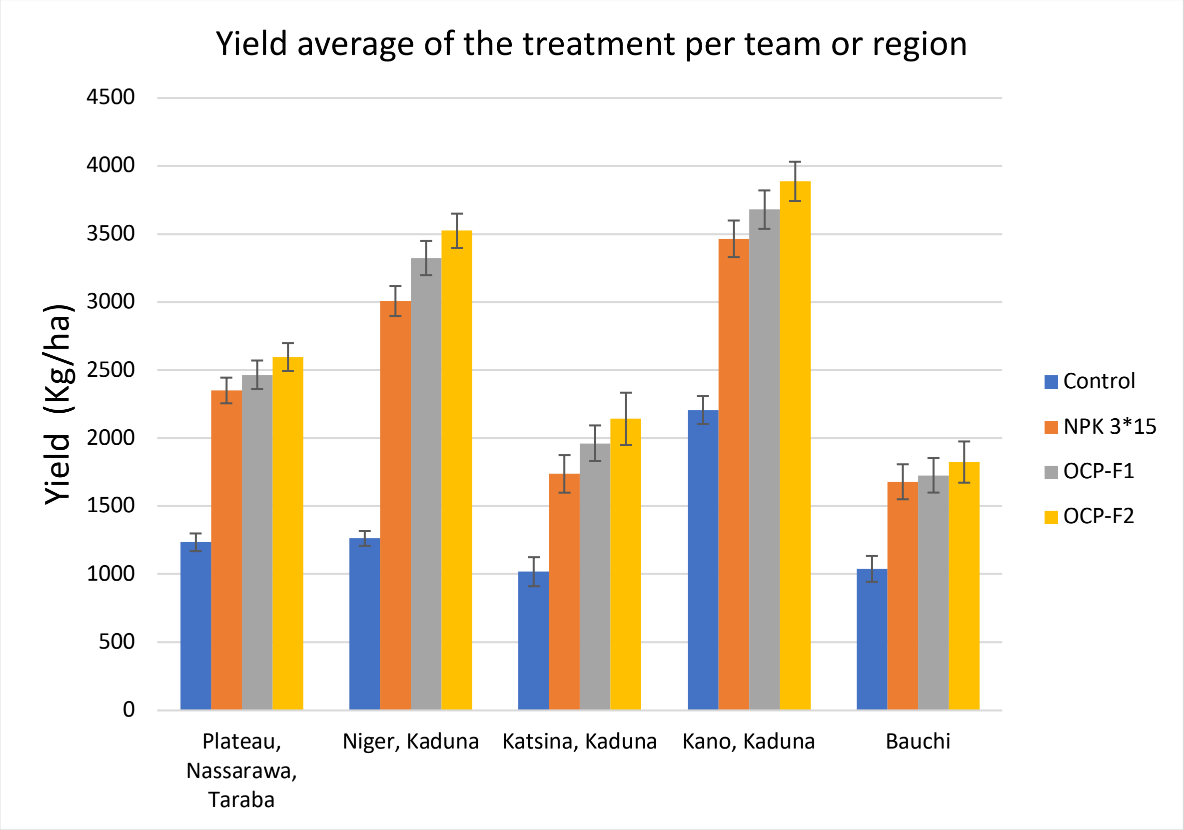


Figure 2 Yield at harvest average for the various treatments based on whole plot data of the validation trials, differentiated for the various regions and teams. The error bars represent the standard error of the mean (SEM)

The statistical analysis indicated that we have significant differences between each individual treatment with OCP-F2 having the highest yield, followed by OCP-F1, NPK 3*15 and the Control treatment yielding lowest. This trend is consistent for the various regions apart from Bauchi and it is consistent for the results obtained from the circular plot, as well as from the whole plot.

The yield levels recorded for the C-plot are generally higher than those recorded for the T-plot, and this seems especially the case for the fertiliser treatments and less so for the Control treatment. This is explained by the relatively high number of plants harvested for the circular plot compared to the whole plot. Where the relative number of plants (q.q. cobs) harvested (the number of plants harvested as percentage of the number of planting stations) for the C-plot is typically between 71% and 87%; the relative number of plants harvested for the whole plot is typically in the range between 57% and 65%. This is different for 'Niger state' in which the relative number of plants harvested is higher for the T-plot than it is for the C-plot, explaining the relatively smaller difference between the yields reported for the C-plot and the T-plot.

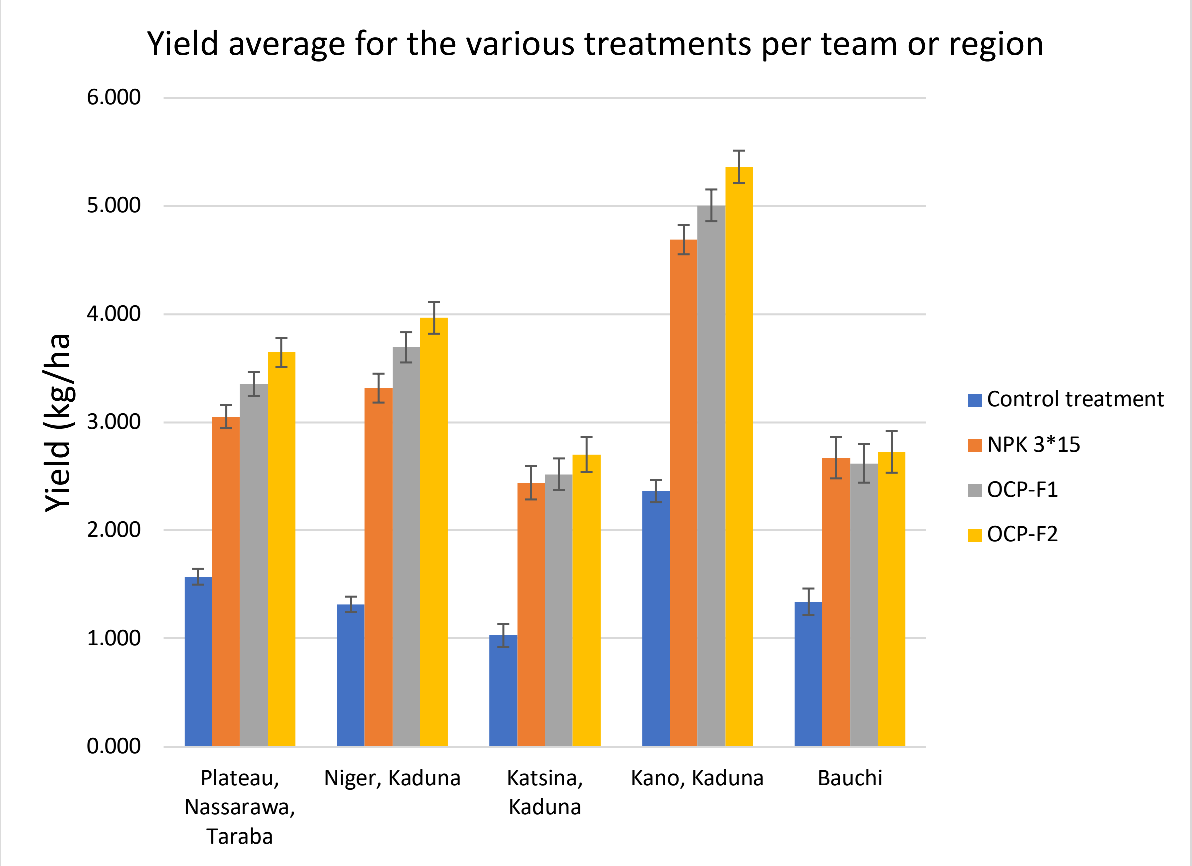


Figure 3 Yield average for the various treatments calculated for the circular plot with results presented for the different regions and teams responsible for the collections of the data. The error bars indicate the Standard Error of the Mean

The number of plants harvested varies strongly between the various regions and between the teams and explaining partly the difference we find in average yield between the regions. Bauchi and Katsina have significantly lower number of plants harvested for the T-plot. Kano, Taraba, Nasarawa and Plateau have comparable numbers of plants harvested and Niger (IAR team) is significantly highest.

The number of plants (cobs) harvested also significantly differs for the various treatments, suggesting that the fertiliser treatment has an effect on the number of cobs harvested (6% and 5% increase in the number of cobs harvested for OCP-F2 and OCP-F1 respectively over the number of cobs harvested for the NPK 3*15 treatment; see Table 5).

For evaluating the response to the fertiliser application, we can look at the ratios of fertiliser treatment yield over the control yield. For the C-plot this ranges between 1.94 and 2.52 for the NPK triple 15 response rate and between 2.27 and 3.01 for the OCP-F2 response rate and between 2.12 and 2.80 for OCP-F1 (leaving Bauchi data outside consideration). It is a considerable difference and the highest response rate is found for the Niger–Kaduna region (the area operated by the IAR team), followed by the Katsina-Kaduna area (BUK team1). The lowest response rates are found in Kano.

Table 5 Mean number of cobs harvested for the various treatments for the C-plot and T-plot and the Duncan grouping.

	Circular plot		Whole plot	
	Mean no cobs	Duncan grouping	Mean no cobs	Duncan grouping
OCP-F2	46.34	A	316.14	A
OCP-F1	45.93	A	310.18	A
OCP-NPK3-15	43.78	B	304.54	A
Control	31.73	C	116.74	B

When considering the data for the T-plot, response rates vary between 1.57 and 2.38 for NPK-3*15, between 1.76 and 2.79 for the OCP-F2 treatment and between 1.67 and 2.63 for the OCP-F1 treatment. The results confirm that we find a significant higher response rate for the Niger-Kaduna AoO than for the other regions and Kano region giving the lowest response rate irrespective of the fertiliser treatment.

4.3 Response to fertiliser application adjusted for plant density

We have calculated the ‘yield at harvest’ adjusted for the number of plants harvested. The number of plants harvested for the C-plot as well as the T-plot may vary strongly between the plots/treatments of the trial. It may influence the outcome of the VTs (it provides an additional source of variation), especially considering the low relative number of plants harvested in these farmer-managed fields, as we already reported in the previous section. The ‘adjusted yield’ is obtained by taking the average weight of the cob of the T-plot and multiplying by 53,333 which is the plant density (number of plants per ha) that is obtained when adhering to plant spacing defined in the protocol. In this report we refer to this variable as the ‘attainable yield’, being the yield that would have been actually obtained under optimum management and if no plants would have been lost. Because we only counted the number of cobs for the T-plot, the conversion is made based on the number of cobs, assuming a 1:1 cob-to-plant ratio. It overestimates the ‘attainable yield’ level with approximately 10%, because the average cob-to-plant ratio for our trials was determined at 0.9 approximately. The ‘attainable yield’ corresponds to the average cob weight and the difference between the attainable yield level reflects the effect of the treatment on the cob weight.

The ‘attainable yield’ levels for the fertilised plots are between 4t/ha to 5t/ha apart for the Kano area where the attainable yield level is around 6t/ha (not considering the Bauchi region for which the data is less reliable). Overall, the mean attainable yield for the OCP-F2 treatment is significantly higher than the attainable yield level of the NPK-3*15 treatment. The ‘attainable yield’ for OCP-F1 treatment takes an intermediate position. The mean attainable yield is slightly lower than for OCP-F2 and not significantly different. The OCP-F1 ‘attainable yield’ is higher than the mean ‘attainable yield’ for NPK3*15, but also not significantly different at a significance level of 0.05.

That we do not find a significant difference between the OCP-F1 an NPK-3*15 attainable yield is the consequence of the deviating results for the Katsina and Bauchi area. The trials in these areas are characterised by very low numbers of plants harvested and large variation in the number of plants harvested between the plots. It results in a large variation in the yield recorded and consequently a high SEM and CV. Based on the results from the IAR team (Niger-

Kaduna) and NAERLS teams (Nasarawa, Plateau and Taraba) we conclude that the OCP fertiliser treatments have an effect on the size and weight of the cobs.

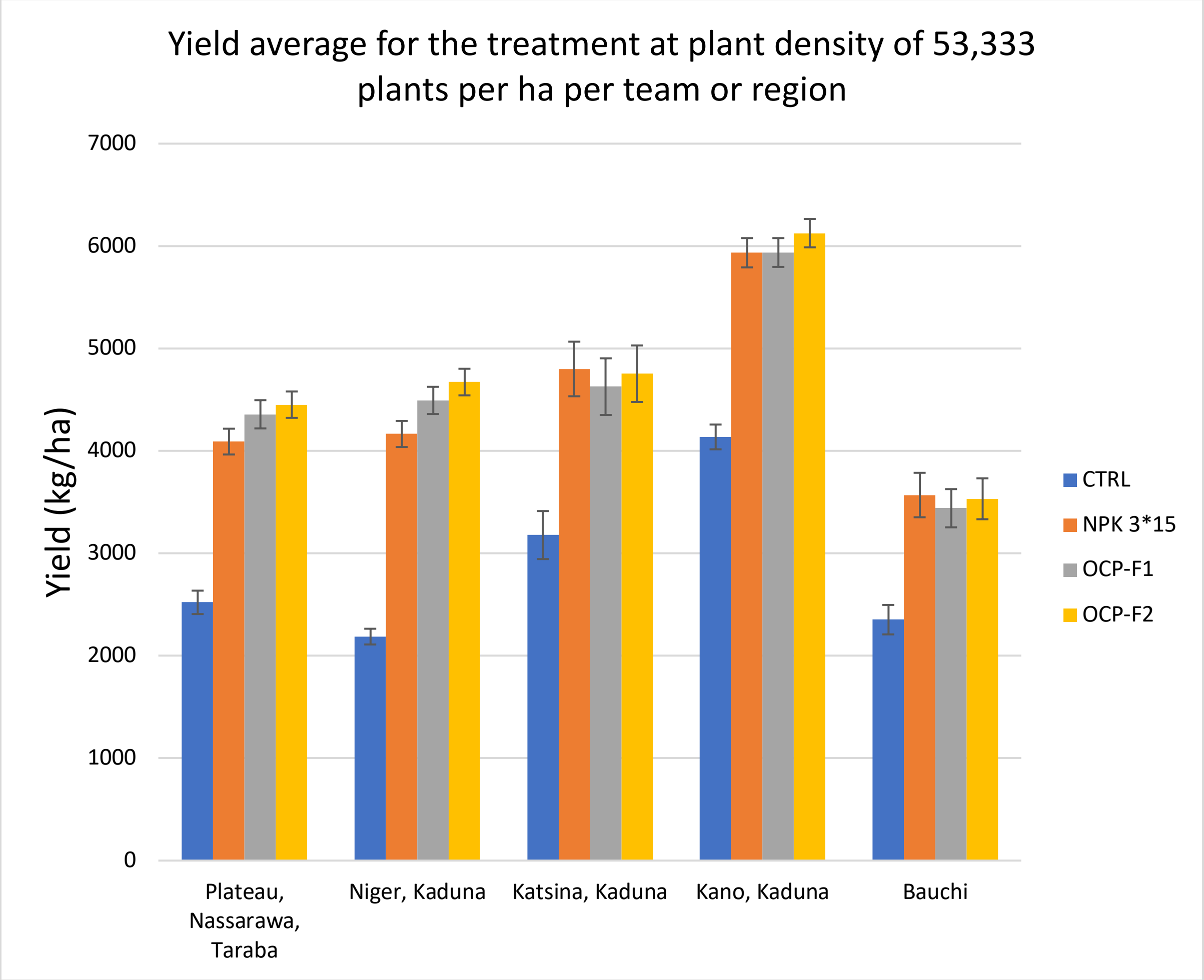


Figure 4 Yield average for the various treatments adjusted to a plant density of 53,333 plants per ha, representing full plant density. Error bars reflect the Standard Error of the Mean

The results for the 'attainable yield' need to be interpreted with some caution. The 'attainable yield' for the Control treatment overestimates the yield that can be attained in comparison with the fertiliser treatments because the cob:plant ratio for the control treatments (a mean of 0.8) is considerably lower than the same ratio for the fertiliser treatment (a mean of 0.9). In practice the difference between the fertiliser treatments and the control treatment will be bigger. Also, the large variation in the management of the trials and the often less than optimal crop management is not only reflected in the relatively low numbers of plants harvested but also in the relatively low average cob weight, which has an equalizing effect on the difference between the Control and fertilized plots. The average cob weight varies between 75g and 90g per cob for the fertiliser treatments of the trials run by the NAERLS and IAR teams, whereas for Kano-Kaduna region it is around 115g per cob. The maximum mean cob weight that is found for trials run by the NAERLS and IAR teams is around 150g, which would correspond to a maximum yield of the 8 t/ha. For the proper evaluation of the effect of the treatment on cob weight boundary analysis should be done, or analysis of the average cob weight for the trials that represent the upper 20% quantile of the yield frequency distribution.

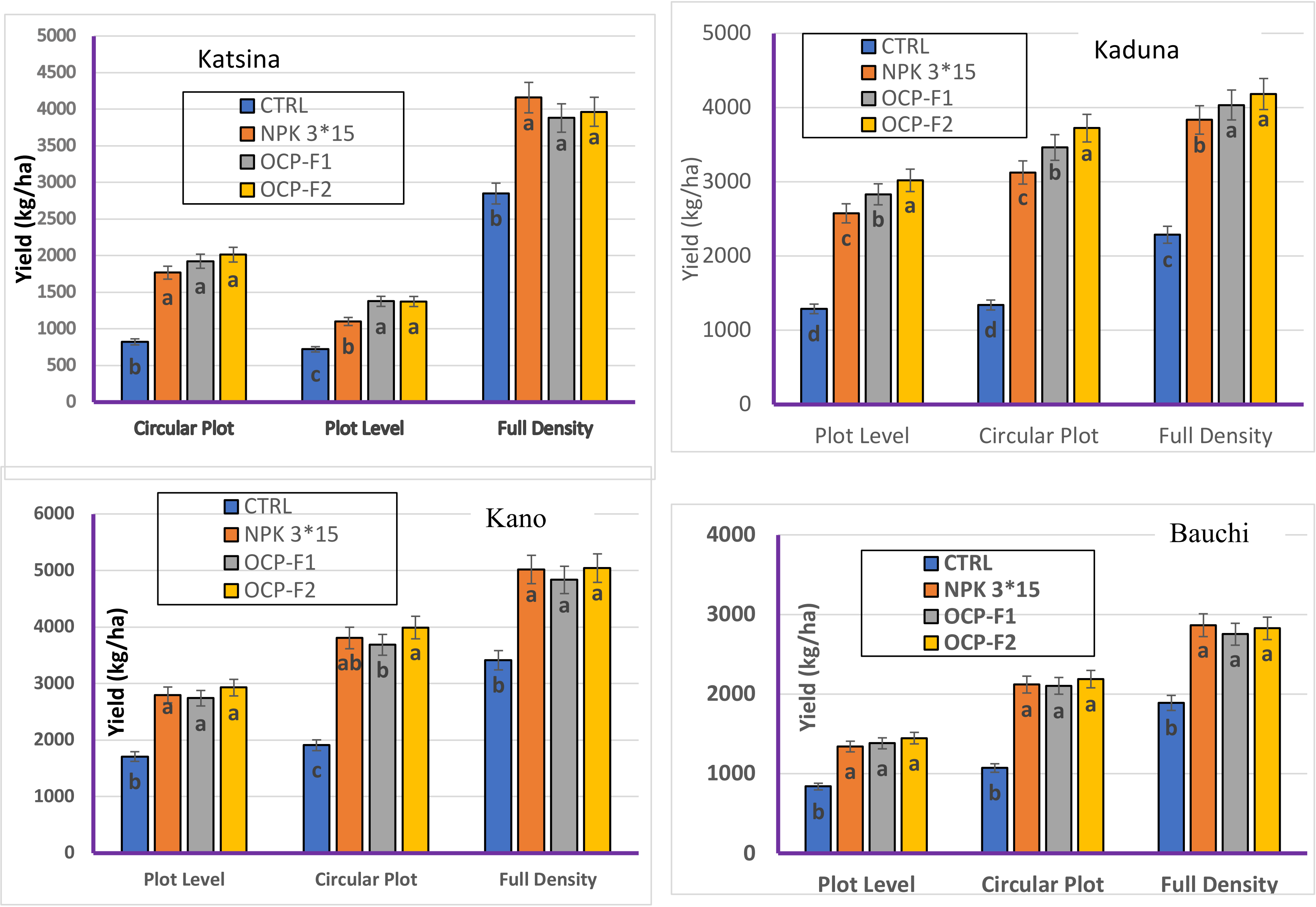


Figure 5.
Grain yield average for the various fertiliser treatments for the circular plot, the total plot and plot yield adjusted for plant density, broken down by state: Katsina, Kaduna, Kano and Bauchi.
(Note the difference in scale for the yield)

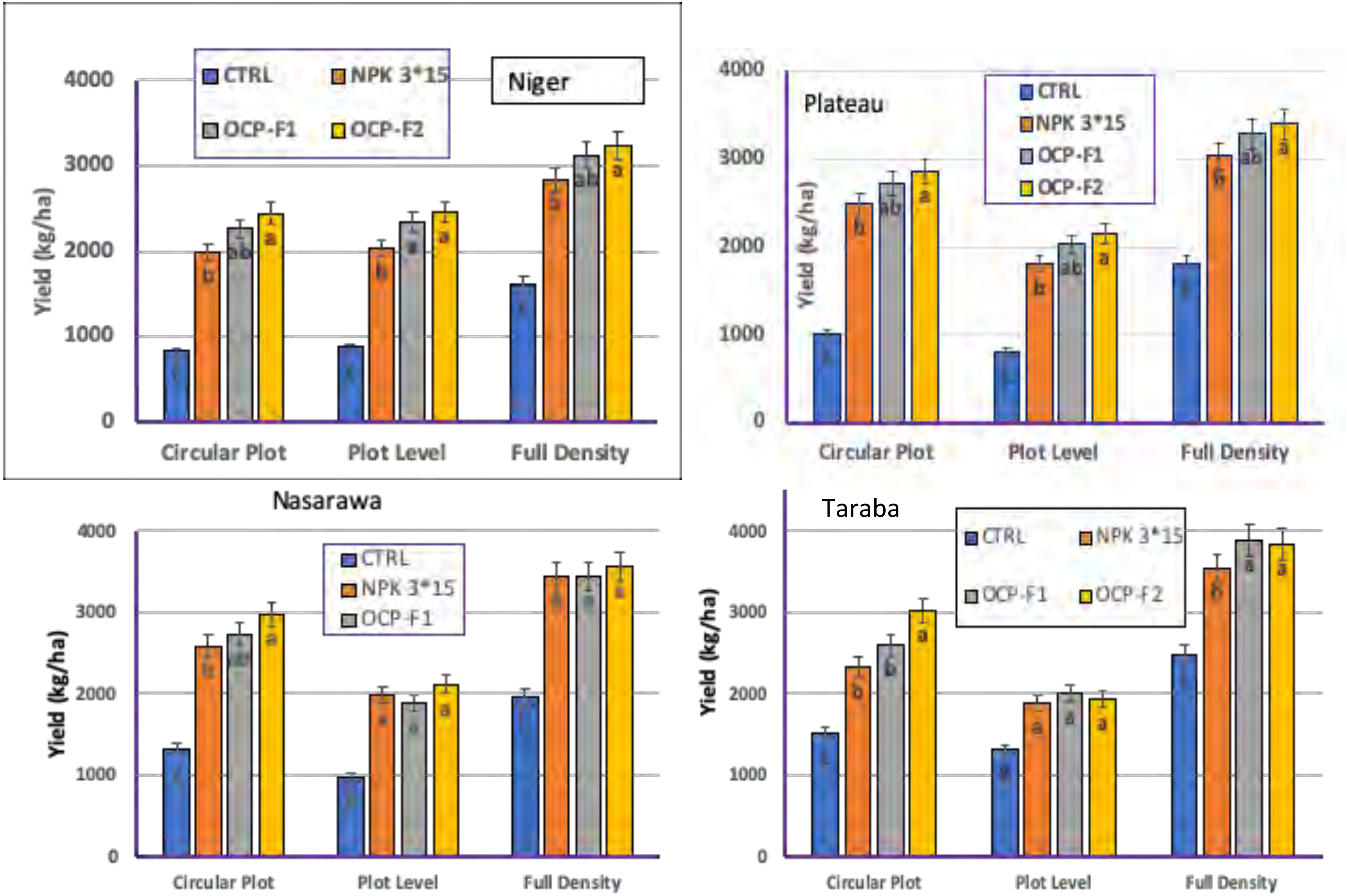


Figure 6.
Average grain yield for the various fertiliser treatments for the circular plot, the total plot and plot yield adjusted for plant density, broken down by state: Niger, Plateau, Nasarawa and Taraba.

4.4. Grain yield response to fertiliser treatment by state

On the previous pages the yield response to the different fertiliser treatment is presented for each individual state.

Grain yield is obtained by taking 80% of the 'yield at harvest'. The 80% is the average shelling rate that we determined. We find very little variation in the shelling rate between trials and therefore we are justified to apply the same shelling percentage to the average yield at harvest for each of the states. We present the 'circular plot' yield, the yield for the total plot ("plot level") and the plant density adjusted yield ("Full Density"). The letters in the bar indicate mean separation. The same letter indicates yield similarity at 5% statistical level of significance using Tukey HSD. The statistical analysis of the total data set already indicated significant differences between the OCP-F2 yield, the OCP-F1 yield and the NPK-3*15 yield. This is not reflected the same in the results for the individual states, because the numbers of trials per state is too small to get the same significance levels. Moreover, for 'Kaduna' we have contributions from various teams which adds to the variability in the yield response.

Comparing yield levels between 'states' is not very meaningful, and we also explained that because of the relatively low number of plants harvested in the trials for 'Bauchi' and 'Katsina' the result for these two 'states' is not very reliable. But otherwise we can look at the difference in trends as far as the response to the different fertiliser formulations is concerned. For 'Kaduna', 'Niger', 'Plateau', 'Nasarawa' we find the same trend in that OCP-F2 performs best, followed by OCP-F1 and subsequently NPK-3*15. For 'Taraba' the trend is different in that OCP-F1 performs best (even though results for the C-plot and the T-plot contradict each other). For 'Kano state' we see that OCP-F2 and NPK-3*15 have similar results, but that OCP-F1 tends to show the lower yields even then NPK-3*15. The 'Kano area' represented a fairly small area.

The trend for the adjusted yield is the same as for the yield measured for the circular plot and total plot giving credibility to the results obtained. Looking at the best performing trials we observe that the attainable yield levels with the application of 150 kg/ha of basal application and 100 kg/ha top dressing with urea is around 6.5 t/ha.

4.5 Evaluating the response to OCP-F1 and F2 new fertiliser formulations

Figure 7 presents the average yield increase for OCP-F1 and OCP-F2 with respect to NPK triple 15, broken down for the different teams, q.q. regions. When considering the average yield increase, we see that OCP-F2 performs significantly better than OCP-F1 (apart from Bauchi where this is not significant). The average yield increase for OCP-F2 for the Niger/Kaduna, Katsina/Kaduna, Kano/Kaduna area of observations ranges from 400 kg/ha to 520 kg/ha approximately, and these are in themselves not significantly different. The average yield increase for OCP-F2 for the Plateau, Nasarawa and Taraba, is 238 kg/ha. There does not seem to be a direct relationship between the general yield level and the expected yield increase over NPK triple 15, which suggest that these effects are region specific. The trend for mean yield difference for OCP-F1 – 'NPK triple 15' is the same as for OCP-F2.

However, these results are obtained considering all trials, while knowing that not all locations are equally suitable for both new formulations. We find contrasting results for the performance of the OCP-F1 and OCP-F2 fertilisers. For example, for the Niger-Kaduna 'area of observation' we find a difference of 565 kg/ha on average between the yield obtained with OCP-F1 and OCP-F2 fertiliser, confirming that both fertilisers respond differently to the condition in the terrain. When making a distinction between the trials and selecting only those trials where the one OCP fertiliser performs better than the other, and calculating the yield difference for either OCP-F1 or OCP-F2 with NPK triple 15, we get much better and more realistic result for the yield increase expected with the application of either one or the other OCP fertiliser over the ‘NPK triple 15 fertiliser’. The results are presented in Figure 8.

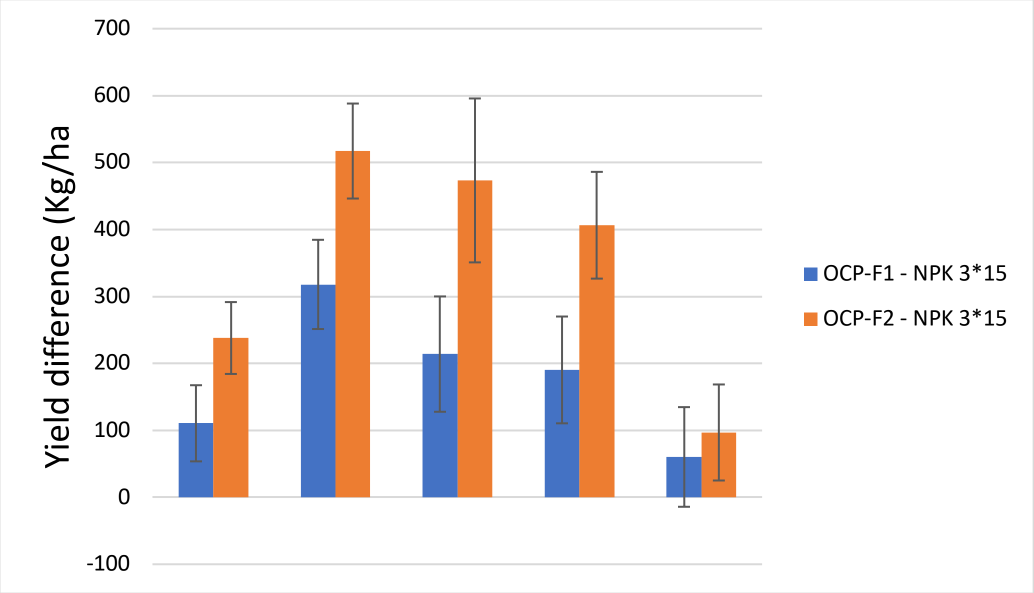


Figure 7 Mean difference in yield for both OCP-F1 and OCP-F2 fertiliser with NPK triple 15, broken down by for the various teams, q.q. regions. Error bars indicated the standard error of the mean.

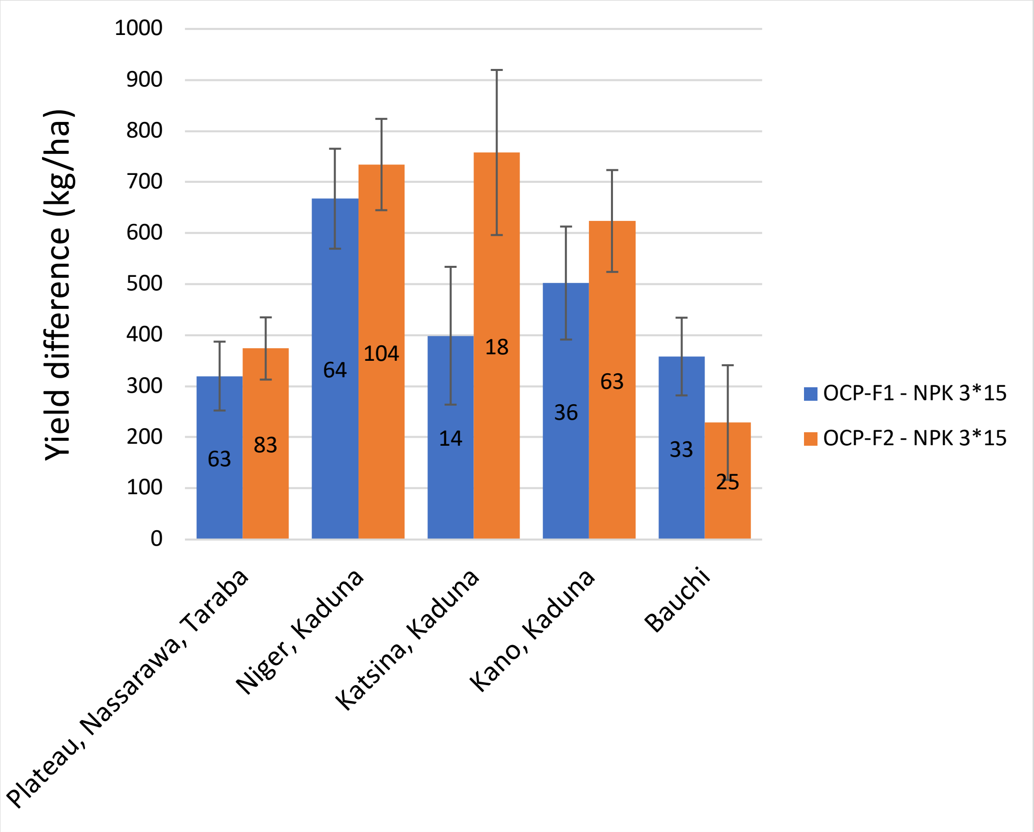


Figure 8 Mean yield increase for either OCP-F1 or OCP-F2 treatment with respect to the NPK triple 15 treatment, broken down by team, q.q. region, considering only the trials where that particular OCP fertiliser performs best of the two OCP fertilisers. Error bars indicate SEM

The mean yield increase for OCP-F2 ranges from 374 kg/ha for the Plateau, Nasarawa & Taraba AoO to 758 kg/ha for the Katsina-Kaduna AoO (excluding the results for Bauchi). For OCP-F1 the yield increase ranges from 320 kg/ha for the Plateau, Nasarawa & Taraba AoO to 667 kg/ha for the Niger-Kaduna AoO. In this case the difference in yield increase for OCP-F1 and OCP-F2 are much less, apart from the Kaduna-Katsina AoO. The gains in yield increase are highest for OCP-F1 indicating that OCP-F1 is suitable for more specific or restrictive conditions.

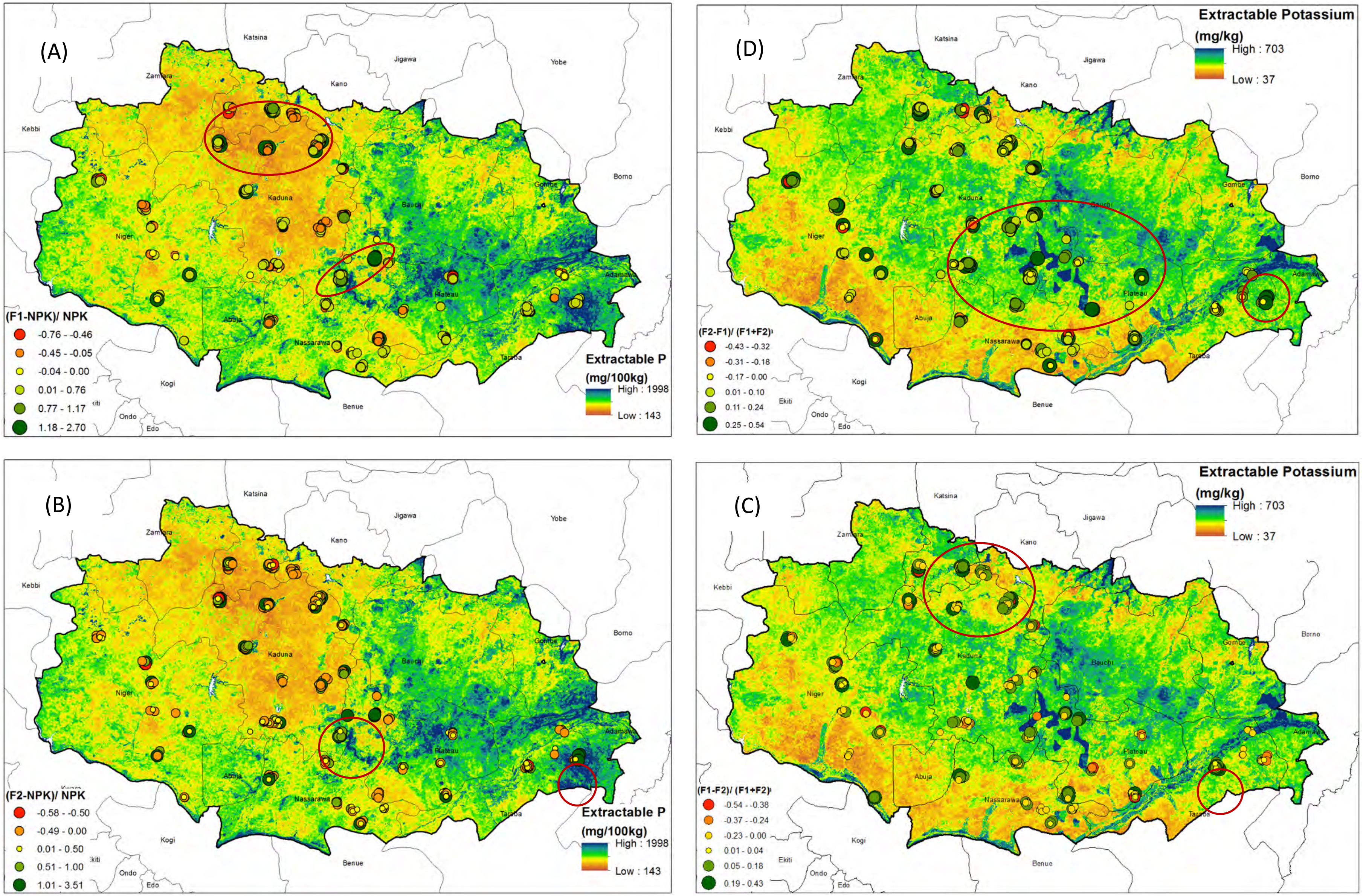
The yield increase in terms of percentage of the yield for 'NPK triple 15' treatment is presented in Table 6. For OCP-F1. The mean yield increase ranges from a 4% for Bauchi to 12% yield for the Katsina-Kaduna 'area of observation', in case no distinction is made between the trials. For OCP-F2 the yield increase ranges from 10% for the Plateau, Nasarawa & Taraba AoO to 27% for the Katsina-Kaduna AoO. However, as mentioned, the OCP-F1 nor the ICP-F2 fertiliser is intended to be used as single alternative for 'NPK triple 15' fertiliser under al conditions, the percentages presented in the last two columns give a better indication of the expected yield increase in case the application can be targeted to the right conditions. The percentage increase in mean yield compared to 'NPK triple 15' ranges from 16% for Plateau, Nasarawa & Taraba AoO to 44% for the Katsina-Kaduna AoO for OCP-F2. For OCP-F1 it ranges from 14% to 23% yield increase also for Plateau, Nasarawa & Taraba AoO and Katsina-Kaduna AoO respectively.

Table 6 Mean percent yield increase for OCP-F1 and OCP-F2 compared to NPK triple 15, considering all trials (overall) and considering only those trials for which either OCP-F1 or OCP-F2 is the best performing of the two OCP fertilisers, broken down by team, q.q. regions.

Regions serviced by the various teams	OCP-F1 overall (%)	OCP-F2 overall (%)	OCP-F1 targeted (%)	OCP-F2 targeted (%)
Plateau, Nasarawa, Taraba	5	10	14	16
Niger, (Kaduna)	11	17	22	24
Katsina, (Kaduna)	12	27	23	44
Kano, (Kaduna)	5	12	14	18
Bauchi	4	6	21	14

4.6 Spatial distribution and factors that determine response to fertiliser application

Even though we see difference in the magnitude of the relative yield increase between states, the differential response between OCP-F1 and OCP-F2 is not explained by the 'state' or region. Rather we see a lot of variation within the 'state' or region and even within the cluster we generally do not see a consistent pattern in the response to either OCP-F1 and OCP-F2. To investigate the spatial variation, we have plotted the relative yield difference for OCP-F1 with NPK 3*15 (Map 10A) and for OCP-F2 with NPK 3*15 (Map 10B) for each trial. We have done the same for the relative yield difference between the OCP-F1 and OCP-F2 treatments (Map 10C for F2-F1 and Map 10D for F1-F2).



Map 12. (A) Relative yield increase of OCP-F1 over NPK triple 15 and (B) of OCP-F2 over NPK triple 15 plotted for each trial with red colour indicating a decrease and green colours indicating increase in yield and with the size of the circle indicating the magnitude of the increase or decrease against backdrop of map of extractable P; same for (C) relative difference of F1-F2 and (D) F2-F1 against backdrop of map of extractable potassium.

The yield increase for F2 compared to NPK-3*15 seems evenly distributed, though we find consistent positive outliers for one cluster in the Taraba region and for a few clusters in the central region of our ROI (the northern part of Plateau bordering with Kaduna). For OCP-F1, on the other hand, we find consistently high values for relative yield increase for another cluster in the Taraba region, for some clusters in the central part of the ROI and for some clusters in the Katsina and northern Kaduna region.

We would be interested to identify the areas where OCP-F1 is performing better than OCP-F2 and vice versa, such that we would be able to target the application of either the OCP-F1 or the OCP-F2fertiliser to that area. The map 10D does not show a clear spatial pattern in where OCP-F1 is performing better than OCP-F2. There is one cluster in Taraba state where find OCP-F1 consistently performing better (See Map10D) in contrast to another cluster in Taraba where OCP-F2 is consistently and considerably performing better. Otherwise there seems to be the area of Katsina, Kano and northern Kaduna encircled in red on Map10D where the chances for OCP-F1 performing better are relatively high (still also with trials where OCP-F2 is performing best.

The maps also do not seem to show a clear relationship in response patterns for either OCP-F1 or OCP-F2 in relation to either extractable P and extractable K. This is not surprising as the factors that determine the response to a particular nutrient application are quite complex. A low available K level does not automatically mean a good response to a fertiliser that contains higher K_2O content, certainly if it is not the only limiting nutrient. We already indicated that P and K as well as some micronutrients are low to critically low throughout the ROI. For K, like for other nutrients, other factors come into play, like the capacity of the soil to hold the nutrient and making it available for plant uptake, for example.

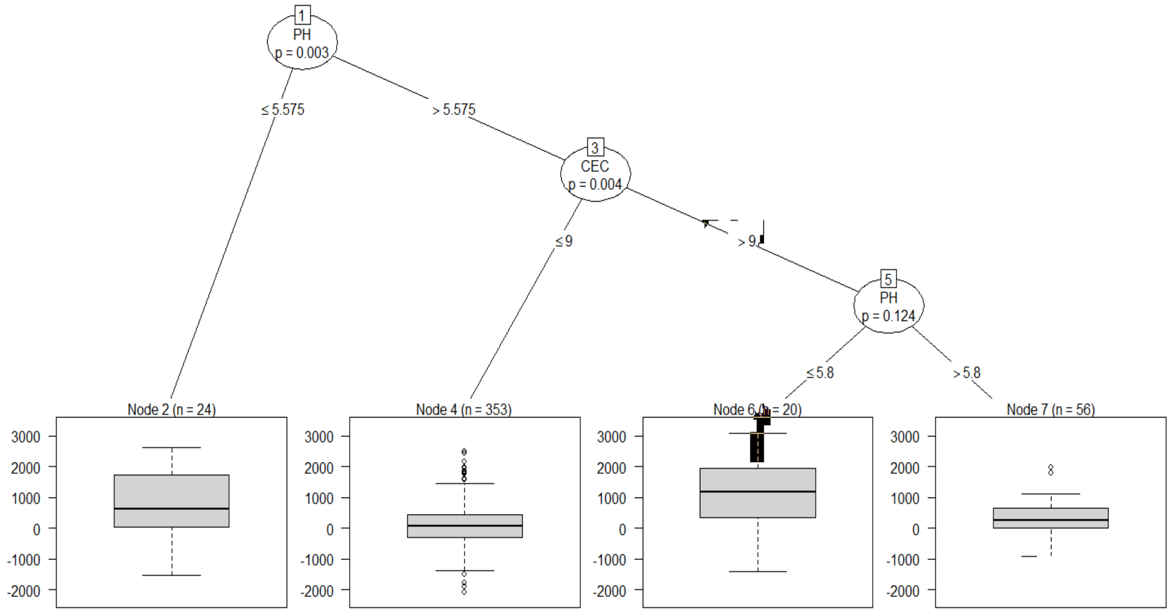


*Figure 9 Result of the regression tree analyses of the adjusted yield difference for the OCP-F1 and NPK-3*15 treatment. A pH of less than 5.6 especially in combination with a relatively high CEC of more than 9 cmol/kg gives considerable yield increase, but the number of observations (n) remain small.*

Regression tree analysis revealed that the general yield response to fertiliser application is determined by texture (a higher percentage of sand reduces the yield), by SOC if the texture is very sandy, by pH and altitude mainly. When looking at the yield difference of OCP-F1 with NPK triple 15, we find a consistent pattern for C-plot, T-plot and adjusted yield levels, in which lower pH (less than 5.6) in combination with a relatively high CEC (more then 9 cmol/kg) gives a significant higher yield for the OCP-F1 treatment. For the C-plot data, a SOC of above 0.75% has a positive effect in response to the OCP fertiliser. For the yield difference of OCP-F2 with 'NPK triple 15' we find the same pattern unfortunately, with less restrictive cut-off values, however. The cut-off value for pH is at 5.85, which explain the overall higher incidence of cases in which OCP-F2 performs better than OCP-F1. When doing the regression analysis for the yield difference between OCP-F1 and OCP-F2, we do not find any factor that has a significant effect on the result.

The data suggests that the variation in responses, especially to OCP-F1 is determined by specific locally varying conditions, with a patchy distribution in the landscape. This was also found in earlier studies done in Western Kenya. Further research would be needed to investigate the local spatial patterns in the response to improved fertilisers and the combination of factors that determine that response. It would require other methods, amongst others for a more precise assessment of the soil properties. But this is mainly for theoretical purposes. A more practical approach would be for the farmer to try out the fertiliser that is the best recommended (either OCP-F1 or OCP-F2) for that specific location based on the information on the soil condition (whether existing or newly acquired) and get confirmation of the recommended fertiliser, or decide on the alternative fertiliser, based on the evaluation of the response.

5. General conclusions and recommendations

The conclusion from this study is that both new OCP fertiliser formulations perform significantly, and considerably better than the commonly used 'NPK triple 15' fertiliser. This yield increase is realized under less than optimum management conditions and conditions that are experienced on farmers' fields. That is, a considerable yield increase can be expected with common agricultural practices and generally accepted fertiliser application rates (three bags of an NPK fertiliser and 2 bags of urea). With 'NPK triple 15' yield levels of less than 3 t/ha this means a considerable increase in fertiliser use efficiency that is required for the farmer to get an acceptable rate of return on the fertiliser investment.

The two new formulation respond differently to the conditions in the field, which is concluded from the contrasting results (yield difference) for both fertilisers at level of the individual field trials. In case the right fertiliser is selected for the prevailing conditions, an average yield increase ranging from 374 kg/ha to around 750 kg/ha can be expected with the use of OCP-F2. For OCP-F1 the expected average yield increase ranges from 320 kg/ha to 667 kg/ha. The results differ from one region to the other also in relative terms. For OCP-F2 the yield increase in percentage of the yield of the NPK triple 15 treatment ranges from 16% to 44% (the latter for the Katsina-Kaduna region of interest). For OCP-F1 the yield increase in percentage of the NPK triple 15 yield ranges from 14% to 23%.

The results confirm the decision of OCP to develop the two alternative formulations that are contrasting in K_2O. The results for having the two alternative fertilisers are much better than when only one of the two would have been offered as an alternative fertiliser. There does not seem to be much reason to believe that a possible third alternative fertiliser would improve the situation much, since there is no reason to assume that there are specific regions with a constraint envelope that we have not yet observed. Rather, the whole ROI seems to be characterized by very low levels of available P and extractable K and the observed micronutrient levels.

For as yet, it is rather difficult to predict the response to either OCP-F1 and OCP-F2 or to make an informed choice for either OCP fertiliser based on information on the soil properties, even though results suggest that pH and CEC play an important role in the selection of the most suitable fertiliser. This is partly due to the large variation we observe in the yield for the various treatments and partly due to the level of confidence in the prediction of soil properties that is inherent to the methods used. Another aspect is that the general conditions do not seem to vary much across the ROI. There is little contrast between the regions, with P and K (and S, B and Zn) being limiting throughout and making the response to either OCP-F1 or OCP-F2 dependent on specific local conditions and therefor difficult to predict.

Appendix 1. Project reports and documents

[1] Ado Yusuf (2016). Final Phase 1 report: training soil sampling, soil sampling and sampling in Kaduna and Niger states. Internal technical Report IITA-OCP S.A. (May 2016), IAR, Zaria.

[2] Samndi Musa Ayuba (2016). Report on the training for soil sampling, Project report, BUK, Kano.

[3] AfSIS (2016). Soil nutrient status of the croplands of central Nigeria, 2016. Project report, Africa Soil Information Service, Arusha, Version: December 4, 2016

[4] Jeroen Huising (2016). OCP Nigeria Maize Validation Trials - Protocol, Updated and final version 15 Dec 2016, Project document, Ibadan

[5] IITA (2017). Report on the training of the supervisors, team leaders and facilitators on the establishment of the validation trials for the OCP-Nigeria project for the testing of new maize fertiliser formulations, conducted at the institute of National Agricultural Extension Research and Liaison Services (NAERLS), ZARIA, 9th and 10th of May 2017. Includes the protocol for the VT, the field book for recording of data and the report status on the verification and identification of locations for the VTs by E. Jeroen Huising, Ibadan.

[6] Jeroen Huising (2017). Report on OCP training of trainers on the protocol for harvesting the validation trials and use of the ODK form for data collection on harvested trials, conducted 19-21 Sept 2017. Includes reports on sensitization and establishment of the validation trials by BUK, IAR and NAERLS, and includes the standard operating procedure for yield assessment of the OCP VTs, collecting grain samples and soil samples at time of harvesting. Internal project document, Ibadan.

[7] IITA (2018). Data Editing Workshop – Quality control and editing of data collected on the maize harvested from the validation trials, conducted 20-22 Feb 2018, Internal project document, Kano

[8] E. Jeroen Huising, Kamaludeen Tijjani (2018). Results of the validation of OCP newly developed maize fertiliser formulations for the maize growing belt of Nigeria, Summary report of report no. 3 of the cooperation agreement IITA-OCP S.A., Ibadan.

[9] E. Jeroen Huising, Kamaludeen Tijjani (2018). Results of the validation of OCP newly developed maize fertiliser formulations for the maize growing belt of Nigeria - Statistical analysis of the yield at harvest and grain yield. Report no. 3 of the cooperation agreement IITA-OCP, Ibadan

[10] J. Huising (2018) Bullet points on data quality control for the OCP validation trials data. Internal report for instruction on the data quality control procedures

[11] Kamaludeen Tijjani (2018). Performance/validation report of OCP fertiliser formulations in maize belt area of Nigeria: yield comparisons at state level. Internal report on the analysis of treatment effect at state level, Kano.

Appendix 2. Composition of the two OCP fertiliser formulations OCP-F1 and OCP-F2

Fertiliser formulation %	N	P2O5	K2O	S	Zn	B2O3
OCP F1	11	22	21	5	1	1
OCP F2	14	31	0	9	1	1

Appendix 3 Data table with the results of the statistical analysis underlying the various graphs presented in this report

Table A1.1 Descriptive statistics of the yield data for the fertiliser treatments for the different data set provided by the various teams for the different areas of observation.

NAERLS	AVR	N	STDEV	SE	MAX	MIN	CV
Average control yield cplot	1.573	157	0.921	0.073	4.864	0.130	0.59
Average yield Triple 15 cplot	3.052	166	1.375	0.107	7.610	1.043	0.45
Average yield OCP Formula 1 cplot	3.354	168	1.457	0.112	8.032	0.320	0.43
Average yield OCP F2 cplot	3.645	167	1.734	0.134	8.990	0.542	0.48
Average control yield	1235	151	803	65	4900	67	0.65
Average yield NPK triple 15	2351	155	1180	95	5591	438	0.50
Average yield OCP F1	2466	151	1294	105	6006	546	0.52
Average yield OCPF2	2597	156	1267	101	5923	344	0.49
Average yield control corrected for plant density	2521	144	1372	114	8134	286	0.54
Average yield NPK triple 15 at full PD	4091	156	1570	126	10723	1186	0.38
Average yield OCP F1 at full PD	4358	155	1717	138	11169	1089	0.39
Average yield OCP F2 at full PD	4451	155	1608	129	11298	862	0.36
Average difference yield OCP F1 - NPK triple 15	111	142	678	57	3103	-2165	
Average difference yield OCP F2 - NPK triple 15	238	148	653	54	3296	-1277	
Average diff yield OCP-F1 - NPK 3*15 (F1>F2)	320	63	536	67	2182	-918	
Average difference yield OCP-F2 - NPK 3*15 (F2 > F1)	374	83	555	61	1750	-981	
IAR							
Yield at harvest Cplot_CNTR trt average	1.318	174	0.928	0.070	6.830	0.115	0.70
Yield at harvest Cplot_NPK-15-15-15	3.316	175	1.769	0.134	9.045	0.369	0.53
Yield at harvest Cplot OCP-F1	3.693	174	1.842	0.140	9.445	0.550	0.50
Yield at harvest Cplot_OCP-F2	3.966	174	1.931	0.146	10.004	0.705	0.49
Stats for plot yield at harvest							
Yield at harvest CNTRL_avr	1263	165	693	54	4034	166	0.55

Yield at harvest NPK triple 15_avr	3009	172	1447	110	7440	394	0.48
Yield at harvest OCP-F1_avr	3324	172	1657	126	8111	387	0.50
Yield at harvest OCP-F2_avr	3524	172	1648	126	8862	494	0.47
Yield at harvest corr for PD_CNTR	2187	158	965	77	5116	589	0.44
Yield at harvest corr for PD_NPK 3*15	4165	170	1663	128	9949	1304	0.40
Yield at harvest corr for PD_OCP-F1	4493	171	1738	133	10261	1343	0.39
Yield a harvest corr for PD_OCP-F2	4672	171	1695	130	10336	1253	0.36
Yield diff OCP-F1 NPK trip15_average	318	168	864	67	2911	-1790	
Yield diff OCP-F2 – NPK trip15_average	517	168	920	71	3464	-1365	
Yield diff OCP-F1–NPK 3*15 (IF F1>F2)	667	64	784	98	2621	-744	
Yield diff OCP-F1–NPK 3*15 (IF F2>F1)	734	104	912	89	3464	-1365	

BUK team 1

Yield_at_harvest_cplot_NPK triple15	2.443	59	1.197	0.156	5.809	0.308	0.49
Yield_at_harvest_cplot_CTRL	1.030	54	0.791	0.108	3.292	0.200	0.77
Yield_harvest_cplot_OCP-F1	2.519	57	1.109	0.147	5.756	0.647	0.44
Yield_harvest_cplot_OCP-F2	2.703	57	1.217	0.161	5.147	0.372	0.45
Yield_harvest_plot NPK 3*15	1738	40	868	137	3887	86	0.50
Yield_harvest_plot CTRL	1019	30	584	107	2631	326	0.57
Yield_harvest_plot OCP-F1	1963	35	776	131	3226	447	0.40
Yield_harvest_plot OCP-F2	2142	36	1159	193	5366	228	0.54
Yield_harvest_NPK 3-15_corr PD	4800	39	1664	267	9843	1907	0.35
Yield_harvest_corr PD_CTRL	3178	29	1263	234	6343	1141	0.40
Yield_harvest_corr PD_OCP-F1	4627	34	1616	277	9184	2302	0.35
Yeld_harvest_corr PD_OCP-F2	4754	34	1610	276	8698	1766	0.34
Yield_diff_OCP-F1–NPK 3-15	214	34	502	86	1371	-874	
Yield_diff_OCP-F2–NPK 3-15	473	35	724	122	2607	-1110	
Yield diff_F1–NPK3-15 (F1>F2)	399	14	505	135	1371	-378	
Yield diff_F1–NPK3-15 (F2>F1)	758	18	686	162	2607	-214	

BUK team 2

C-plot data							
Average yield ctrl	2.365	104	1.060	0.104	5.807	0.523	0.45
Avr yield NPK 3*15	4.689	105	1.393	0.136	7.770	1.710	0.30
Avr yield OCP F1	5.006	106	1.515	0.147	8.393	1.148	0.30
Avr yield OCP F2	5.361	106	1.554	0.151	9.240	1.795	0.29
Plot data							
Avr Yield at Harvest_CTRL plot	2206	102	1040	103	4882	428	0.47
Avr Yield at Harvest_NPK 3*15	3465	104	1367	134	6811	731	0.39
Avr Yield at Harvest OCP-F1	3680	105	1444	141	7742	611	0.39
Avr Yield at Harvest OCP-F2	3888	106	1482	144	7599	806	0.38
Plot data at full PD							
Avr Yield CTRL plot at full PD	4137	102	1222	121	6650	1152	0.30
Avr Yield NPK triple 15 at full PD	5935	104	1456	143	9109	2215	0.25
Avr Yield OCP-F1 at full PD	5937	104	1436	141	11049	2244	0.24
Avr Yield OCP-F2 at full PD	6126	105	1415	138	11996	2055	0.23
Avr yield diff OCP-F1━NPK 3-15	191	100	798	80	1985	-2077	
Avr yield diff OCP-F2━NPK 3-15	407	101	801	80	2800	-1268	
Avr yield diff OCP-F1 - triple 15 (F1>F2)	502	36	664	111	1985	-418	
Avr yield diff OCP-F1 - triple 15 (F2>F1)	624	63	792	100	2800	-891	

BUK team 3

C-Plot data							
Average yield ctrl	1.341	72	1.043	0.123	4.975	0.180	0.78
Avr yield NPK 3*15	2.673	74	1.646	0.191	7.975	0.273	0.62
Avr yield OCP F1	2.621	73	1.527	0.179	7.196	0.290	0.58
Avr yield OCP F2	2.727	71	1.623	0.193	6.780	0.355	0.60
Plot data							
Avr Yield at Harvest_CTRL plot	1039	63	754	95	3733	110	0.73
Avr Yield at Harvest_NPK 3*15	1679	64	1029	129	4983	135	0.61
Avr Yield at Harvest OCP-F1	1727	62	996	127	4278	204	0.58
Avr Yield at Harvest OCP-F2	1825	61	1181	151	5641	269	0.65
Yield data corrected for plant density							
Avr Yield CTRL plot at full PD	2352	67	1177	144	5886	615	0.50
Avr Yield NPK triple 15 at full PD	3569	70	1814	217	9533	1051	0.51
Avr Yield OCP-F1 at full PD	3441	70	1560	186	8945	657	0.45
Avr Yield OCP-F2 at full PD	3533	69	1661	200	7638	683	0.47
Avr yield diff OCP-F1━NPK 3-15	61	59	572	74	1218	-1902	

Avr yield diff OCP-F2 – NPK 3-15	97	57	542	72	1711	-1272
Avr yield diff OCP-F1 - triple 15 (F1>F2)	358	33	437	76	1218	-440
Avr yield diff OCP-F1 - triple 15 (F2>F1)	229	25	562	112	1711	-736

Appendix 4 Data table with the average grain yield response to fertiliser treatments per state using different approaches for estimating grain yield

Table A2.1 Grain yield response to fertiliser treatments using the circular plot-based data and plot-based data and for simulated full plant density, arranged by State

	Treatment			
Katsina	CTRL (kg/ha)	NPK 3*15 (kg/ha)	OCP-F1 (kg/ha)	OCP-F2 (kg/ha)
Circular Plot	823	1768	1924	2014
Plot Level	722	1102	1376	1374
Full Density	2848	4159	3880	3965
Nasarawa				
Circular Plot	1326	2584	2736	2965
Plot Level	959	1982	1889	2119
Full Density	1953	3435	3438	3569
Niger				
Circular Plot	825	1992	2262	2446
Plot Level	871	2034	2335	2452
Full Density	1618	2826	3121	3243
Plateau				
Circular Plot	1004	2488	2710	2858
Plot Level	801	1812	2032	2147
Full Density	1814	3029	3279	3401
Taraba				
Circular Plot	1518	2326	2592	3012
Plot Level	1305	1889	2013	1940
Full Density	2482	3533	3893	3838
Kano				
Plot Level	1709	2799	2741	2928
Circular Plot	1909	3806	3685	3991
Full Density	3412	5018	4833	5042
Kaduna				
Plot Level	1288	2575	2831	3019
Circular Plot	1340	3125	3462	3722
Full Density	2286	3833	4035	4182
Bauchi				
Plot Level	838	1341	1381	1447
Circular Plot	1071	2119	2103	2188
Full Density	1889	2865	2752	2825

Appendix 5 Soil characterization and validation trials database

Report no 4 of the cooperation agreement between *OCP-Africa* S.A. and the International Institute of Tropical Agriculture (IITA)

Project: Developing efficient and affordable fertiliser products for increased and sustained yields in the maize belt of Nigeria

Project code: PJ 02375

Client: *OCP-Africa*

Authors: E. Jeroen Huising

Date: 12 June 2019

1. Introduction and purpose of the OCP-IITA-AFSIS Soil Properties and Validation Trials database

In its endeavor to develop new affordable and efficient maize fertilisers for Nigeria this project, that is funded by *OCP-Africa*, has collected a lot of data, both in relation to the soil characterization that was undertaken to get insight in the nutrient limitations based on which the new fertiliser formulations were designed as well as in relation to the validation trials that were conducted to validate and test the new formulations in a large number of locations. The aim was to develop a few fertilisers that were best suited for the different regions of the maize belt of Nigeria considering the varying agroclimatic and soil conditions within the area. The approach was to develop the new formulations based on best-bets informed by information on the possible limitations in nutrient availability across the maize belt of Nigeria and then subsequently test these new formulation widely, rather than doing elaborate nutrient response test on a large number of locations to determine the best formulation for these specific locations. This requires insight in the spatial variability of these soil conditions, and it requires insight in the spatial variability of the crop response to the various fertilisers to be validated. To this end the soil was sampled in nearly 3000 locations distributed across the 225,000 km^2 maize belt of Nigeria and the fertilisers were tested in a large number of validation trials likewise distributed across the maize belt in Nigeria. Two new fertiliser formulations were tested in comparison to the much used NPK triple 15 fertiliser in each of the trials. The project intended to have validations trials conducted on 1500 locations. However, in the end the project was able to establish these validation trials in 866 locations. The project was implemented by IITA in collaboration with its national partners BUK, NAERLS and IAR, who each took responsibility for establishing and managing the trials in different regions. The soil characterization was carried in collaboration with AFSIS.

The Project reported on the soil characterization work done, and likewise did the Project report on the results from the validation trials. The Project now also wants to make the data available and report on the data that was collected during the project. The Project considers this to be good practice. To provide access to the data and insight in how the data was collected and generated helps to understand how the results were obtained and helps in the interpretation of the results. It also becomes possible to validation the results as presented in the various reports.
Second, the Project herewith wants to facilitate further analysis of the data. This is a very rich data set that can be used for other purposes than it was intended for in this project. Also, further analysis can possibly be done beyond what we have been able to do and had time for. Further investigation into the factors that determine the response, or the conditions under which one or the other fertiliser formulation performs best would help to better target the application of either of the two formulations and improve the recommendations for the use of these fertilisers. For example, from our preliminary investigation it shows that the pH and the CEC are important factors in determining the response to OCP-F1 and OCP-F2 compared NPK triple 15; a relatively lower pH and relatively higher CEC favors the use of both the OCP-F1 and OCP-F2 fertilisers. We also find that soil texture is an important factor in determining the response to the NPK fertilisers in general (whether NPK triple 15, OCP-F1 or OCP-F2). However, we have not been able to determine the factors that determine the response to either the OCP-F1 or OCP-F2 (that is the

response to one of the OCP fertilisers compared to the other) and whether clear spatial patterns can be discerned in the response to the one or the other fertiliser.
Third, the data set can be linked or added to other data sets, and as such contribute to further analysis and improvement of the results. For example, the spectral data together with the results of the wet chemistry may contribute to the improvement of the calibration models of the spectral data and improvements of the prediction models for the soil properties if it were only for Nigeria. Also, worth mentioning is the TAMASA project conducted in Nigeria by CIMMYT and IITA, in which also al lot of agronomic trials have been conducted with maize as test crop. Combined use of the data from both projects may help to improve the models used for maize fertiliser recommendations.

All the further use of the data is served by the proper documentation of the data. It is only then that all the requirements for the use of the data can be observed: That is the data should be available, accessible and applicable. The availability of the data depends on when the data will be released by OCP-Africa, which will be half-a-year after closure of the project. The data will be accessible through the open data repository of IITA. The applicability of the data for the various purposes can be determined based on the documentation of the data and that is what this report is about. The data set will be available with the full meta-data standard adhered to.

2. Overview of the OCP soil properties and fertiliser validation trial database for central region of Nigeria

The full set of data pertaining to soil characterization and validation trials consists of soil data referring to soil properties as determined in the lab using different techniques of analysis and referring to soil properties as predicted from the MIR spectra, and it consist of the data in reference to the validation trials. The soil data also includes the MIR spectra itself. The data related to the validation trials may refer to data related to the establishment and the management of the trials as well as to the harvesting of the trials. During the harvesting of the trials a sample of the cobs from each plot harvested was taken for analysis in the lab and all the data of the measurements in the lab is also included. For the harvest data we do not only present the yield data as it is derived from the measurements in the field, but we also present the data recorded in the field that is subjected to quality control and data editing procedures as annotated data files in which notes and comments are included for data items that have been edited, rejected or simply highlighted because the values are suspicious. It also includes the formulas used for the yield calculations, such that these can be verified.

Even though this is not a formal database, it is still important to describe the relationship between the tables and to identify the key attributes through which that relationship is established. Also, we identify the key attributes that uniquely identify any record or row of data inside a table. We do not present a formal entity relationship diagram, but the basic entities that apply in this case refer to the Field, the Validation Trial, the Plot and the Experiment, and then we have the Cob Sample. Formally, we have also the individual Cob as an entity, because we have recorded

data for each of these individual cobs. However, the data for each cob only has statistical relevance and there is no need for the identification and 'characterization' of each of cob. That is why the data is presented but not a unique identifier for each record is presented.
In relation to the soil data the basic entity is the Soil Sample, since that is what all the data presented relates to. We can identify the cluster of observation and sample collection points and the 'sentinel site' or 'block' as an entity, but this is not relevant in this case because we are not presenting any specific data related to those entities.

Field: The field refers to the particular location where a trial is conducted, and in our case mostly refers to farmer's fields. Each field has its own characteristics, like location, farmer's name, topographic positions, and one can think of land use history, etc. However, we have not collected elaborate information on the characteristics, because this information was not considered relevant for the purpose of the trials.

Validation Trial (VT): The validation trial refers to the one particular instance of the experiment, which means that one particular trial is implemented on a particular field at a particular time (particular cropping season or year) following a specific or particular design that is characteristic of the total experiment.

Plot: The plot is a well-defined part of the field or trial to which a particular treatment is assigned. The specific plot where a particular treatment is implemented is not consequential for the outcome of the trials, because we assume that the soil characteristics are the same for each plot in the field/trial, for example. The most consequential is the treatment and because we do not have replicated treatments in the trial, the plot can be and is identified by the trial identifier (index number or barcode) and the treatment code. All observations related to the harvest are done at plot level and the plot form the smallest unit of observations.

Maize Cob Sample: Refers to the sample of cobs taken for each plot to be analyzed in the lab for moisture content, grain weight, 100-grain weight, etc. The cob sample is identified by the QR code and the QR code also presents the link the plot from where the sample was taken.

Soil Sample: This is the sample of soil taken at particular location and time, and at a particular depth. The soil sample is shipped to the soil lab for analysis. The soil sample is identified by a unique QR code that also serves for the registration of the sample in the lab. The link to other observations in the field is through the location, and that includes link to the validation trials.

There is information stored in the protocols, and in the standard operation procedures that is important for the evaluation and interpretation of the results, both in as far as results from the soil analysis as well as results from the validation trials are concerned. The documents are part of the database even though we have tried to capture the information as much as possible in the meta-data that accompanies the data files. If this data is to be part of a larger database and the data to be used in combination with or in comparison to data from other experiments, than this information would have to be made explicit and stored in a structured manner. We would then have to define entities like ‘experiment’ to describe the design of the experiment, ‘treatment’ to formally describe the treatments as part of the experimental design, ‘fertiliser’ to describe the specific characteristics of the fertilisers used in the trials, ‘crop’ and ‘crop variety’, to describe the specific characteristics of the test crop used in the experiment.

The data in our database is generally not (or only partly) the primary or original data as it is collected in the field. In a project like ours with so many trials implemented by so many different teams, the implementation of the validation trials may differ on details, and the collection of the data may not have been done in a consistent manner, despite the protocols and standard operating procedures having been defined. Often unit conversions have to be carried out, typing errors have to be corrected and outliers have to be detected and rejected. Calculation of the effective plots size, important for the conversion from yield per plot to yield per hectare, depends on the orientation of the plot, the number of plant rows harvested the number of plants within the row harvested for example. This is all part of the quality control procedure and the data is thus processed to the first level.

In our case editing and corrections have been made to the original data, but these corrections can also be made through adjustment in equations used for calculating the yield per hectare. Rather than presenting the level 1 processed data, or the final calculated yield data we present the data files in which the corrections are shown and in which the formulas used for the calculation of the yield are made explicit, such that is fully transparent how quality control has been performed and what corrections have been made and to show that the results obtained are reliable.

The data is recorded by entering data in electronic form (ODK forms) and subsequently uploaded to an ODK server. We have made use of different servers for the soil data and for the data related to the validation trials. The original data is still available at those different sites and can be exported and downloaded from there, though the access may be restricted. We have the latest version of the files downloaded, and we will keep these for possible future reference, but these are not part of the data that we want to share.

The data is not stored using a formal database structure and a database management system (DBMS), but rather as a set of files, either as EXCEL (.xlsx) or as comma separated values files (.csv). All relevant metadata is stored as separate data sheets within the workbooks. The metadata refers to general project information, general information on the specific data set (e.g. filename, general content, origin, etc.) and specific information on the attributes, or the data that is contained in the various columns in the data sheet. The description of the fields in the tables is not included in this report, but rather we refer to the data files itself.

The data will be transferred to GCSPACE using CKAN as data repository, once of the requirements for the metadata have been fulfilled, following the Open Data Policy of IITA and the CGIAR. We will then also protect the data files, such that no changes can be made by unauthorized persons.

3. Description of the tables

Below the data files are described, rather than the tables per se. Some files contain more than one data sheet, apart from the sheets that are used for the recording of the metadata. This relates to the different sections in the ODK form for which data is recorded, but it refers to the same object or entity, like the measurement on the sample of cobs and the measurement on the individual cobs in the sample.
We have stored the files in separate folders. All files related to observations on the validation trials are stored in one folder and all files related to the soil characterization are stored in different folder where distinction is made between the soil analyses data originating from the ICRAF lab in Nairobi and the soil analyses data originating from the lab at IITA in Ibadan.
As mentioned further information on the data is stored in the data files itself; that is the description of the attributes (i.e. referring to the data contained in the various columns in the data set) as well as meta-data like the geographical scope, owner or author of the data etc. The metadata is provided especially for the files related to the Validation Trials and less so for the soil data. If further information is required, like on the specific methods used for the soil analyses or the unit of measurement in which value as given, in as far as not indicated, this can always be provided.

Table 1. Description of the data files related to the validation trials

Validation Trials (VT)	**Folder: Harvest&Yield**data						
file name	**extension/ file type**	**creation date**	**key attribute / identifier**	**description / content**	**origin**	**#records**	**remarks / methods**
OCP_Yld-data&covariates_complete	xlsx	25/05/2019	sin (index number for validation trial) and treatment name	Yield at harvest data for the various treatment plots of the validation trials, including coordinate data of the trial site locations, and including the covariate data for those locations obtained from different data sources, including some soil properties.		2399	Data set compiled from the different source files; full set of metadata included as separate sheets in the workbook.

BUK_T3_VT_yieldatharvest_final	xlsx	24/03/2019	parent_index (index number of the validation trial) and treatment name	Contains yield at harvest data calculated based on the circular plot, the plot and for full plant density, for the VT under responsibility of the BUK team 3	BUK team 3_VT_harvest_corr&process	312	Final data set after data curation and quality control; data set is not complete in that we have empty cells for missing data (missing in original data set or rejected during the quality control process); contains full meta-data set
BUK-T2_VT_yieldatharvest_summ	xlsx	24/03/2019	parent_index (index number of the validation trial) and treatment name	Contains yield at harvest data, calculated based on the circular plot, the plot and for full plant density, for the VT under responsibility of the BUK team 2	BUK team 2_VT_harvest data_corr&process.xlsx	424	Final data set resulting from data curation and quality control; data set is not complete in that we have empty cells for missing data (missing from the beginning or rejected during the quality control process); with full meta-data set
BUK_T1_VT_yield at harvest_summ	xlsx	24/03/2019	parent_index (index number of the validation trial) with the treatment name	Contains yield at harvest data, calculated based on the circular plot, the plot and for full plant density, for the VT under responsibility of the BUK team 1	BUK team 1_VT_harvestdat_corr&process	264	Final data set resulting from data curation and quality control; data set is not complete in that we have empty cells for missing data (missing from the beginning or rejected during the quality control process
IAR_VT_yield harvest_summ.xlsx	xlsx	25/04/2019	parent_index (index number of the validation trial) with the treatment name	Contains yield at harvest data calculated based on the circular plot data, plot data and as attainable yield (assumed full plant density)	IAR_VT_harvestdat_clean&process	708	Contains empty cells where data was missing or when data was rejected during quality control check
NAERLS_VT_yield_summ	xlsx	03/09/2018	parent_index (index number of the validation trial) with the treatment name	Contains yield at harvest data calculated based on the circular plot data, plot data	NAERLS_harvest data_clean_	692	Contains empty cells where data was missing or when data was rejected during quality control check

				and as attainable yield (assumed full plant density)	final		
BUK team 1_VT_harvestdat_corr&process	xlsx	24/09/2018	_index; sequential number assigned to each record in the original data file	Harvest and yield data recorded in the field from the Validation Trials related to the plot/ treatment, controlled for quality, corrected and edited, including yield calculations retaining the equations used; validation trials conducted under responsibility of BUK team 1	Data file ODK form "OCP_VT_Harvest (Field)", exported from https://odk.ona.io/cimmyt_gcap/	272	Includes parameters used for quality control and data annotations regarding suspicious values and corrections made; meta-data and data dictionary included as separate data sheets in the workbook.
BUK team 2_VT_harvest data_corr&process.xlsx	xlsx	17/07/2018	parent_index (index number for the validation trial) with the treatment name	Yield at harvest data for the plot/treatment of each Validation, controlled for quality, corrected and edited, including yield calculations retaining the equations used; validation trials conducted under responsibility of BUK team 2	Origin: Data file with data recorded using the ODK Form "OCP_VT_Harvest (Field)", exported from https://odk.ona.io/cimmyt_gcap/	432	Includes full metadata set in separate data sheets in the workbook; includes all kinds of parameters used for quality control and summary statistics for the various treatments; includes calculation of the yield at harvest
BUK team 3_VT_harvest_corr&process	xlsx	29/09/2018	parent_index (index number for the validation trial) with the treatment name	Data recorded in relation to yield at harvest for the plot/treatment for each Validation Trial, controlled for quality, corrected and edited,	Data file ODK Form "OCP_VT_Harvest (Field)",	416	Includes full metadata set in separate data sheets in the workbook; includes all kinds of parameters used for quality control and summary statistics

				including yield calculations retaining the equations used; validation trials conducted under responsibility of BUK team 3	exported from https://odk.ona.io/cimmyt_gcap/		for the various treatments; includes calculation of the yield at harvest
IAR_VTharvestdat_clean&process	xlsx	12/10/2018	parent_index (index number for the validation trial) and treatment name	Data recorded in relation to yield at harvest for the plot/treatment for each Validation Trial, controlled for quality, corrected and edited, including yield calculations retaining the equations used; validation trials conducted under responsibility of the IAR team	Origin: Data file with data recorded using the ODK Form "OCP_VT_Harvest (Field)", exported from https://odk.ona.io/cimmyt_gcap/	804	Includes full metadata set in separate data sheets in the workbook; includes all kinds of parameters used for quality control and summary statistics for the various treatments; includes calculation of the yield at harvest
NAERLS_harvestdata_clean_final	xlsx	03/09/2018	parent_index (index number for the validation trial) and treatment name	Data recorded in relation to yield at harvest for the plot/treatment for each Validation Trial, controlled for quality, corrected and edited, including yield calculations retaining the equations used; validation trials conducted under responsibility of the NAERLS team	Data file ODK Form "OCP_VT_Harvest (Field)", exported from https://odk.ona.io/cimmyt_gcap/	900	Includes full metadata set in separate data sheets in the workbook; includes all kinds of parameters used for quality control and summary statistics for the various treatments; includes calculation of the yield at harvest

OCP_VT_CobsData_2018_02_13_EJH_v1	xlsx	07/08/2018	Barcode (QR code) for the cob sample (or recording instance identifier) and/or record index number (_index or -parent _index in the second sheet)	Contains data from the laboratory measurements on the maize cobs sampled from the field. Some data, like the moisture content relates to the whole sample, whereas there is also data on the individual cobs in the sample (length, weight, number of kernels, etc.). The data is captured in two data sheets	Data file ODK Form "OCP_VT_Cobsmeasurement", exported from https://odk.ona.io/cimmyt_gcap/	2587/11828	Data captured in two data sheets; one is related to measurements related to the full sample of cobs and the second to the individual cobs in the sample. We have calculated and added the average cob weight for each sample
OCP_VT_harvest_location	xlsx	07/08/2018	_index, is the index number for the records/validation trials corresponding with the index number in all other files related to the harvest data collection	Data from section A from the OCP_Harvest form that contains data recorded during the harvest of the field and related to general characteristics of the field, viz. like location, enumerator, etc..	ODK Form "OCP_VT_Harvest (Field)", exported from https://odk.ona.io/cimmyt_gcap/	705	Contains the full metadata set in separate data sheets

VT_Form_OCP1_2018_02_19_21_58_55	xlsx	02/06/2019	_uuid is the instance identifier (there is also an record identifier (_id) and the index number (_index) can be used as identifier). The identifier of the field trial is given by the QR code for the field trial that is listed in the column "ssid"; the VT_ID that is the given name for the trial that can be used for reference	Data related to the establishment of the field trials, e.g. date of land preparation, planting date, form or mode of land preparation etc.	ODK form "VT_Form_OCP1”; Data exported from the ODK server https://mobilsurvey.qed.ai/nisis, on 19 Feb 2018	866	There is one file downloaded from the ODK server in June 2018, which contains data from one field trial more, but this data does not seem to be very reliable. The number of records indicate the number of trials that have been established
VT_Form_OCP2_2018_02_19_21_58_55.xls	xlsx	19/02/2018	The instance identifier (_uuid) can be considered as the key attribute (see also the record identifier (_id) and the index number (_index). The field trial to which the observation refers is identified by ‘VT_ID’ that is the given name for the trial	Data related to the management of the field trials (activities) like date of fertiliser application, weeding, gapping etc.	Data from ODK form ‘VT_Form_OCP2’; Data exported from the ODK server https://mobilsurvey.qed.ai/nisis, on 19 Feb 2018	332	We do not have the data on the dates of the field//crop management activities for all the trials

For the soil data we have not included all the data files in the folder. For the prediction of the soil properties from the MIR spectra a model (using Markov Chain Monte Carlo simulation model and Random Forest regression model) is used to generate a large number of predicted values for each of the soil properties that are stored in separate files (50 runs in our case). From the distribution of the values the median value is chosen and reported as the predicted value. These original files that also give the confidence region of the prediction have not been included in this listing, but they can be made available on request.

Table 2 Description of data files related to the soil characterization, compiled and not specifically related to the soil analyses data from one or the other lab

Soil Characterization **phase _ OCP**	**Folder name: Soil data**						
file name	**extension /file type**	**creation date**	**key attribute / identifier**	**description / content**	**origin**	**#records**	**remarks / methods**
Soil_chemdata_pred &georeference_ OCPproj	xlsx	14/10/2016	Soil sample ID (or the QR code, attribute name:"Barcode"); there is also the IITA AS Lab registration code (attribute name: "SSN")	Soil chemical properties (e.g. SOC, soil nutrients) predicted from MIR spectral analysis results, of soil samples collected from croplands of central Nigeria during soil characterisation phase; contains coordinates of the sample locations	"OCP_ Nigeria_ MIR"	2900	We assume these are the predicted soil chemical properties based on the spectral data from the ICRAF lab in Nairobi, though the number of records raises some questions
soils_top_maize	csv	21/06/2016	recording Instance identifier (_uuid) and soil sample ID (SSID, QR code reading)	Contains data from 'Crop Scout', evidence of livestock and type livestock, and identifies crops cultivated at the specified location, includes location coordinates, photo reference number	Data from Crop Scout ODK form; data exported from https://mobilsurevey.qed.ai/nisis	727	We assume the barcodes (QR codes refer to the soil sample reference code, since these observations are done during the soil sampling exercise in the field

Table 3. Data files related to the analysis of the soil samples from the IITA lab in ibadan

Data files for the soil characterisation - from Analytical Services Lab IITA

ALL OCP Received SOIL_IITA	xls	23/09/2016	Barcode soil sample (QR code) and IITA code soil sample registration IITA AS lab	This file is only to link the QR code of the soil samples collected in the field to the soils lab registration code of the same sample	Analytical services lab, IITA	5765	Final list of the all the samples that have been registered at the lab with the barcode and the IITA code and sample number and sequential number (count); there is a separate sheet with the duplicate codes; approximately half of the number of samples (the top soil samples) have been analysed
IITA_codebook	csv	20/06/2016	IITA-AS lab registration code ("SSN") and soil sample ID ("ssid" which is the QR code)	This file is only to link the QR code of the soil samples collected in the field to the soils lab registration code of the same sample	Analytical services lab, IITA	3433	There are some records with lab registration numbers but no 'ssid' (barcodes) and there also some records that have the barcode (ssid) but no lab registration number
OCP_SSID	csv	12/06/2016	SSID (QR code of the soil sample	Latitude and longitude of the sample location, identification whether it is a top soil or subsoil sample, soil depth restriction (y/n), date of sample collection and SSID	Data recorded using the ODK form and uploaded to https://mobilsurevey.qed.ai/nisis (ODK server)	5807	This is the registration of all the samples collected, both top and subsoils samples. Not all subsoil samples have been analyses; there are some duplicate values for the SSID
OCP Predresult Aug 2016_IITA	xls	23/09/2016	SSN (iita code sample registration) and Barcode (QR code of soil sample)	Prediction of pH Carbon, N, P, Ca, Mg, K, Na, Zn, Mn, Fe, Cu, S, Hp, Ec, B, Al, (plus barcode for SSID and iita code)	Analytical services lab; compiled from the files with the simulated and predicted data for each element separately	3638	

OCP-Topsoil Predictions Dec 2016	xls	30/09/2017	iita code (sample registration number for the AS lab) and barcode (QR code soil sample	Prediction of pH Carbon, N, P, Ca, Mg, K, Na, Zn, Mn, Fe, Cu, S, Hp, Ec, B, Al, (plus barcode for SSID and iita code) for top soil samples only	AS lab; extracted from file with predicted results for all scanned samples	2804	
OCP-WETCHEM Result IITA	xls	23/09/2016		Contains results from the wet chemistry analysis from IITA AS lab: ppm Ca, ppm Mg, ppm K, ppm Mn, ppm Fe, ppm Cu, ppm Zn, ppm B, ppm P, ppm Al, ppm Mo, ppm S	Analytical services lab, IITA	282	

Table 4 Data files of soil analysis results from the ICRAF lab in Nairobi

Data files from ICRAF Nairobi lab	**folder name: OCP_Nigeria_data_ICRAF lab**						
file name	**extension / file type**	**creation date**	**key attribute / identifier**	**description / content**	**origin**	**#records**	**remarks / methods**
OCP_Nigeria_cn	csv	15/03/2017	Code ICRAF lab Nairobi; icr and sequential number (e.g. icr168896)	Contains data on the soil organic carbon and the nitrogen and 'acidified carbon' and 'acidified nitrogen'	ICRAF lab	281	All samples from the OCP project have been sent to ICRAF lab for cross reference; Methods are no specified but could be obtained from the soil lab at ICRAF (assumed CHN analyser used); 281 soil samples are the soil samples analysed for wet chemistry for calibration of the MIR spectra
OCP_Nigeria_ldpsa	csv	15/03/2017	Sample registration code for ICRAF lab	Texture information: Sand, clay and silt fraction in %	ICRAF soils lab	282	methods laser diffraction particle size analysis

			Nairobi (e.g. icr168896)				
OCP_Nigeria_ mineralogy	csv	15/03/2017	ICRAF lab sample registration code (e.g. icr168896)	Gives distribution in percentage for the following minerals: Albite Chrysotile Diopside Epidote Gibbsite Hematite Hornblende Ilmenite Kaolinite Microcline Muscovite Orthoclase Pyrophyllite Quartz Vermiculite	ICRAF soils lab	285	Not specified
OCP_Nigeria_mir	csv	15/03/2017	ICRAF lab sample registration code (e.g. icr168896)	Gives the MIR spectral reflectance data; that is reflectance values for a large number of wavelengths specified in nanometres	ICRAF soils lab	3042	MIR spectroscopy measured with the Bruker Alpha station
OCP_Nigeria_pxrf	csv	15/03/2017	ICRAF lab sample registration code (e.g. icr168896)	Gives results for the total element analysis using the X-Ray fluorescence spectroscopy	ICRAF soils lab	281	TXRF - units are not specified in the file
OCP_Nigeria_sample codes	csv	15/03/2017	ICRAF lab sample registration code (e.g. icr168896)	Gives relation between the icr number (ICRAF sample registration code), the iit number (IITA ASlab sample registration code) and the QR code of the soil sample (QR code assigned during the soil characterisation exercise)	ICRAF soils lab	3042	Sample registration code is the same for the MIR and the wet chemistry analyses - same code refers to the same sample
OCP_Nigeria_wet chemistry	csv	15/03/2017	ICRAF lab sample registration code (e.g. icr168896)	Gives data from the wet chemistry analysis: pH, m3. Al, m3.B, m3.Cu, m3.Fe, m3.Mn, m3.P , m3.S, m3.Zn,	ICRAF soils lab	282	Extraction method: Mehlich3; Analyses are done using ICP-OES

		PSI, ExNa, ExCa, ExMg, ExK, ExBas, Ecd, ExAc	

Appendix 6. Pictorial

Picture 1 Mrs Bola Awotide (OCP) inspecting the fertilisers at customs at Lagos

Picture 2 Soil samples from the OCP project at IITA Ibadan Analytical Services lab

Picture 3 Preparation of soil samples for analysis in the lab

Picture 4 Soils samples are identified by their QR code which is also used for registration in the lab

Picture 5 Group photo of participants of the training workshop on the establishment of the validation trials, Zaria

Picture 2 Dr Adamu Yakubu (NAERLS) inspecting one of the proposed trial site locations in Plateau; example of PTS being rejected

Picture 7 Land preparation for one of the VTs by Plateau 1 team (adjustments made upon inspection of the field

Picture 8 Preparing for distribution of inputs and establishment of the trials

Picture 9 All facilitators were provide a motorbike for transport and transportation in the field

Picture 10 Crossing the bridge

Picture 11 Crossing without a bridge

Picture 12 Land preparation - measuring distance between rows

Picture 13 Ploughing by animal-drawn plough

Picture 14 Planting at planting distance of 25cm within the row

Picture 15 Self-made device for soil sampling

Picture 3 Top dressing with urea

Picture 17 Combatting the Fall Army Worm

Picture 18 Farmer/Facilitator in between two plots (NPK and Control)

Picture 4 Example of picture that Facilitators share through the WhatsApp group of how the crop is doing in a particular trial

Picture 20 Training in the field on the harvesting of the plots and data collection

Picture 21 Weighing the cobs in the field (in this case of the cobs sampled from the circular plot)

Picture 22 Tunrayo Alabi (IITA) and Alex Verlinden (AfSIS) during the training on soil sample collection. In remembrance of Alex Verlinden

IITA

Transforming African Agriculture

OCP

AFRICA

AfSIS