# Mining for gold: large scale data management and analysis of on-farm legume trials

Joost van Heerwaarden[1], Marcel Lubbers[1], Samuel Adjei-Nsiah[2], Freddy Bayukija[2], Peter Ebanyat[2], Nkeki Kamai[2], Emmanuel Sangodele[2], Endalkachew Wolde-meskel[3], Birhan Abdulkadir[3], Arnold Amppiah[2], Anthony Epel[2], Abubakari Mzandah[2], Shehu Muhammad[2], Bassey Ukem[2], Justine Onyinge[2], Aron Othieno[2], John Sekham[2], Theresa Ampadu-Boakye[2], Linus Franke[4]

## Introduction

N2Africa is a large, multi-partner "research-in-development" project that aims to improve legume productivity for small scale farmers.
One of the core principles of the project is that data from on-farm trials is analysed quickly to provide feedback for subsequent seasons.
Obtaining, handling and analysing information across large numbers of trials can be challenging and requires a well designed data management system.
We have developed and implemented an informatics pipeline in which data is collected on electronic tablets, organised and stored in a central database and made available for analysis by on- and offline tools (Figure 1).
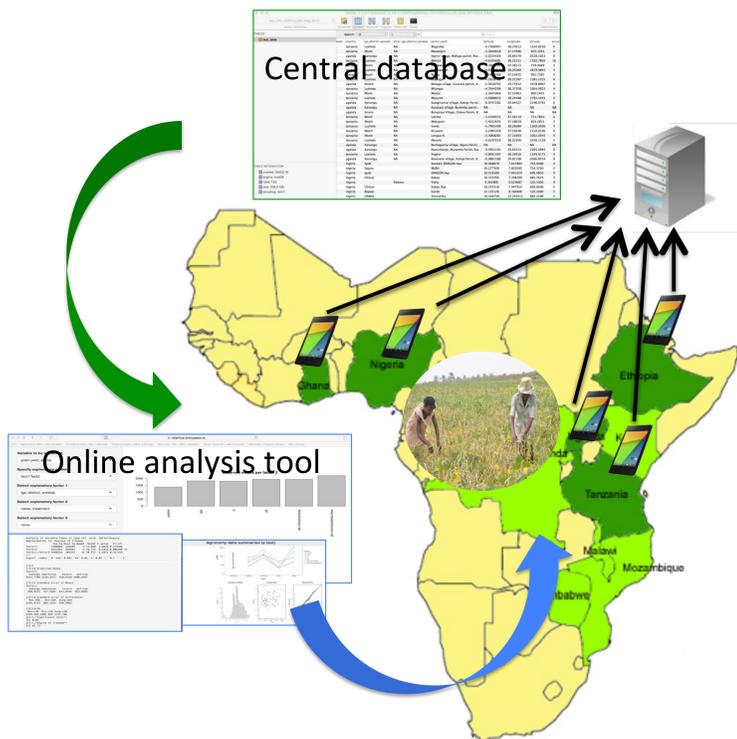


Figure 1. Outline of N2Africa's data management system

## Analysis of yields and input responses across countries and crops

In 2015, we obtained yield data from 356 on-farm trials across 5 countries (Ethiopia, Ghana, Nigeria, Tanzania and Uganda). Figure 2. presents a summary of the yields obtained for unfertilized and fertilized (typically phosphorus at < 30kg/ha) plots for six different legumes. Yield gains varied between crops but averaged 206 kg/ha.



Figure 2. Mean yields for control (red) and fertilized (green) plots for six legumes.

## Data on stresses and their effects on yields

The electronic forms contain questions related to to the performance of the trials, such as severity of different stresses. This allows us to determine important causes for low trial yields (Figure 3).
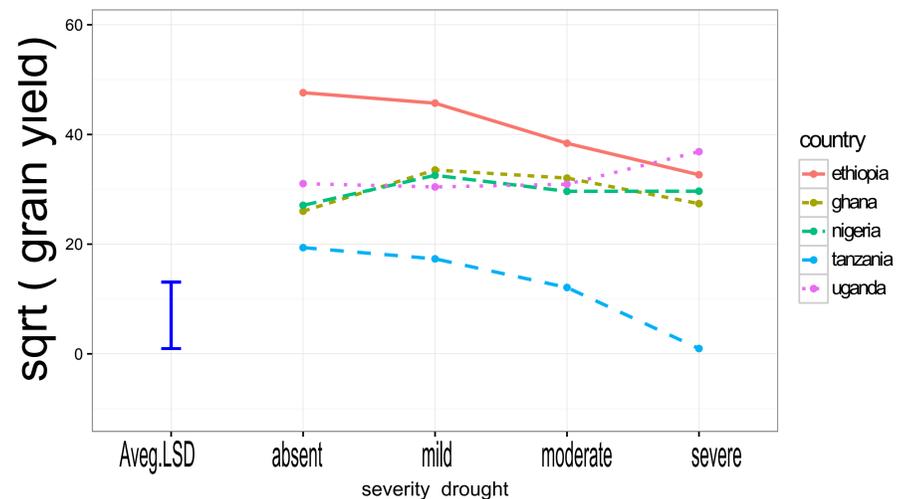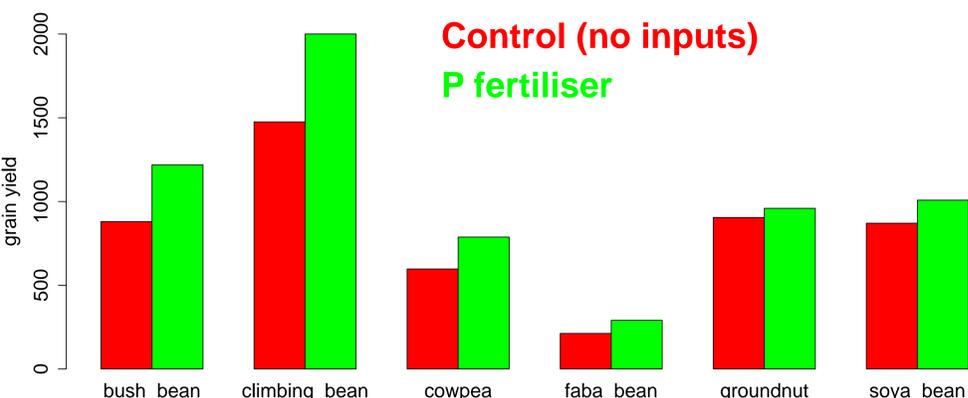


Figure 3. The effect of drought severity on grain yield (square root transformed) for the five countries.

## Associated geographic and image data

Another advantage of the use of electronic forms is the ability to record reliable GPS data using the device's native positioning system, as well as to collect media such as photos of the trial plots (Figure 4.)



Figure 4. Map of trial site based on GPS data (left) and images of two plots mid-season (right)

## Conclusions

- We built an effective pipeline for handling electronic data collected on-farm
- The speed and uniformity of data collection achieved allows for timely analysis of trial outcomes
- Compiling a rich set of agronomic data and associated information offers scope for large-scale and in-depth analysis of determinants of success and failure of technology interventions on farmers' fields

## Acknowledgements

[1]Plant Production Systems, Wageningen University
Address: Droevendaalsesteeg 1,
6708PB, Wageningen, The Netherlands
E-mail: joost.vanheerwaarden@wur.nl,
marcel.lubbers@wur.nl

[2]I International Institute of Tropical Agriculture (IITA)
PMB 5320, Oyo Road, Ibadan 200001,
Oyo State, Nigeria.

[3]I International Livestock Research Institute (ILRI)
P.O. Box 5689
Addis Ababa, Ethiopia

[4]Soil, Crop and Climate Sciences, University of the Free State,
Nelson Mandela Rylaan,
Bloemfontein 9300, South Africa

WAGENINGEN UR
For quality of life