Response Report to the ALU – An agricultural survery for more than 9,500 African housholds

SUMMARY

In looking at the 'Scientific Data survey done on 9,500 African households" that this must have taking a lot of time to compile the data. I believe that this survey must have taking a lot of effort from mania people, organizations, and universities to gather and to put together such a document. As this document points out that many people have done their theses and dr. on this data set, that many organizations have concluded their own insights from the data and that even I can draw my own insights from this data.

Coming from an engineering background I appreciate factual untampered data. Raw data from calibrated instrumentation with specific time, date stamps from which I can come for anomalies and apply learnt mathematical/statistical models too. From this these results I feel comfortable then to draw my own insights.

Looking at how these researchers set out with data capturers to gather the data used is in my mind a huge task. I believe that the questionnaire had good intensions but looking at the geographical area covered (all the countries surveyed), at the language barriers they data capturers must had challenges with and to get to the remote random selected places I assume that this would affect the data and its results.

The approach of "Multi-stage stratified random sampling" still could be affected by human error.

If I were asked to look at gathering such data I would ask the following two questions:

1. Has global warming affected our commercial farms crop yields with regards to temperature and rainfall?
    a. Has the farmer had to use different watering techniques to increase the yield – not just relay on rainfall
2. Has our global growth that has contributed to global warming placed a higher demand on the produce produced from the farms?

How I would gather this data to build my data set:

- weather data for the last 20 years for Africa
    - temperature
    - humidity
    - rainfall
    - windspeed
- gather as many areal and satellite photos of the targeted geographic regions – use image recognition software to determine
    - crop growth over a period
        - to sublimit this with the use of newer drone footage for closed higher resolution pictures
- Data from the seed supplier's tons sold vs grain gathered – this would be a bonus
    - From the seeds that was planted to the actual crop gathered to what was lost while loading trucks to rotting in silos and theft

I would assume that with this approach I can eliminate the amount of errors in my data as I am ruling out human factor of data collection.

I don't think that this would be an easy study to compile nor would it be any less expensive but once these models have been refined for the data inputs and image recognition software becomes better this system would easily be repeatable and useful for future use.

Dieter Gebert

15 February 2018