

# Spatial Comparison and Quality Check of Farmer-recorded Daily Rainfall Data; A Case of Nyakach and Soin-sigowett, Kenya

**Mawora Thomas Mwakudisa, Edgar Ouko Otumba, Joyce Akinyi Otieno**

Department of Statistics and Actuarial Science, School of Mathematics, Statistics and Actuarial Science, Maseno University, Maseno, Kenya

**Email address:**

[tmawora@maseno.ac.ke](mailto:tmawora@maseno.ac.ke) (M. T. Mwakudisa)

**To cite this article:**

Mawora Thomas Mwakudisa, Edgar Ouko Otumba, Joyce Akinyi Otieno. Spatial Comparison and Quality Check of Farmer-recorded Daily Rainfall Data; A Case of Nyakach and Soin-sigowett, Kenya. *Science Journal of Applied Mathematics and Statistics*.

Vol. 8, No. 1, 2020, pp. 11-18. doi: 10.11648/j.sjams.20200801.12

**Received:** December 11, 2019; **Accepted:** January 2, 2020; **Published:** January 9, 2020

---

**Abstract:** Small scale farming is currently still heavily dependent on rainfall in developing nations. With the challenge of climate change, many innovations are proposed to help the farmers mitigate and adapt. The use of historical data provides a starting point in development of decision support tools for them. However, most climate data are not local, but far from the farmers. Thus, the challenge of representability of the data is questioned. In order to use the decision support tools with farmers at Nyakach and Soin-Sigowett, Kenya, historical data was used from a synoptic station 20 km away. The locals felt it was not representative enough, hence the need to look for more local data. In 2014, a CCAFS project empowered 100 farmers from the region with low cost rain gauges to collect and record their own data for use in decision support tools. In this paper, we look at the quality of the data comparing it to the KMS data. Line graphs were used to compare the total seasonal rain for more than 30 years with the farmers perception. In addition, pairwise t-tests have been used to compare difference in farmers recorded rain to the value at the synoptic station. Data from volunteer stations have also been used to confirm the validity of the spatial difference in the data. The results showed that quality of the farmers data is adequate for use. Further, data from farmers deviated from the main synoptic station half of the time. The results clearly show that there is need to allow locals collect their own data to help capture the spatial differences in climate. The farmers recorded data was good quality hence can be used in decision support tools to help them adapt to possible climate change.

**Keywords:** Rainfall, Quality, Farmers Data, T-test, Climate Change, Farmers Perception

---

## 1. Introduction

### 1.1. Background

In scientific studies, indigenous knowledge has been acknowledged and increasingly locals have been included in them [7, 15]. For instance, a number of agricultural experiments include on-farm experiments [12, 13]. These experiments could be either researcher-managed or farmer-managed [3, 11]. Another transition is to have completely farmer-managed and farmer-owned experiments. This was attempted in a Climate Change Agriculture and Food Security [5] study using farmers from Nyakach and Soin-Sigowett. The objective of the study was to increase the farmers informed decision making through experiments. The

study used several decision support tools, which included mathematical models to help simulate expected farm outputs in case of varied farm management decisions. One of the models, Agricultural Production System sIMulator [1], requires use of daily climatic data on rainfall, temperature and radiation. Complete long-term data was acquired from the Kenya Meteorological Services (KMS) station in Kisumu. The farmers felt KMS data was not representative enough for their region, hence need to help localize the data. This was done by empowering farmers to collect their own rainfall data for one season while conducting the farm experiment.

In Kenya, the Kenya Meteorological Services [9], is the main custodian of most weather information. To be able to effectively monitor weather over the country, KMS has data

collection stations which include among others thirty-six (36) synoptic surface weather observation stations for horizontal (spatial) atmospheric sampling of weather elements and more than one-thousand ( $> 1000$ ) rainfall stations. This translates to an average of one synoptic station in an area of 16,147  $\text{Km}^2$ , with more concentration in the wetter regions of the country. The synoptic weather stations collect daily data on rainfall, minimum and maximum temperatures, wind speed, wind direction, air pressure, soil temperature, solar radiation, sunshine duration, relative humidity, evaporation and cloud cover.

There are stations that collect data for personal use, and this increases the spatial coverage. There are other twenty-four Automatic Weather Stations (AWS) that record climate data and transmit it to receiving stations at KMS and 3000 volunteer observatory stations across the country. The AWS stations collect daily data on rainfall, temperature, wind speed, wind direction, air pressure, sunshine duration and relative humidity. Most volunteer stations collect rainfall data only.

Farmer-recorded rainfall is another useful resource and can be used for research such as to understand spatial differences in rainfall magnitude and trends. This is particularly important since rain-fed agriculture is the main source of livelihood to a vast proportion of the population in this region. To enhance food production in rain-fed agriculture, long-term rainfall records need to be recorded and analyzed to help inform farmers on the best Climate Smart Agricultural practices to either mitigate or adapt against rainfall extremities.

Researchers in East Africa have found that what farmers perceive as risks to agriculture do not compare to the evidence got from analyzing long term climatic records thus affecting their management response and hence recommend a participatory approach where farmers' get involved in discussions about climatic elements and their impact on food production.

### 1.2. Spatial Coverage of Climate Data

In order to use more spatial covered data, several tools have been developed so researchers can cover as much as possible. One tool is the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Climate Data Record (PERSIANN-CDR) that estimates daily rainfall for an area of 0.25 [2]. There also exists the Australian dataset for quality climate [8] and for North America [18] and a more recent German National Forest Inventory [6]. To assess quality, a spatial data can be used to look at questions like (1) Are there significant regional differences in accuracy among data sets? (2) How accurate are their mean values compared with extremes? (3) Does their accuracy depend on spatial resolution. Benkhe considered downscaled datasets from different stations and found considerable differences for the rainfall, but not so much with temperature [4]. In this study, there was no downscaled data for comparison, rather the available thus farmers contributed in the data collection process.

For the farmers in Soin-Sigowett, the main observatory sites are in Kisumu and Kericho. These are more cosmopolitan areas and the people who receive this actual rainfall are not the main users. The users are several kilometers away from the National Meteorological Service (NMS) stations. Nyakach is more than 20 kilometers by road from the Kisumu Observatory while Soin-Sigowett is around 30 kilometers.

Further, end users often perceive a need for temporal and spatial detail in the prediction, including exact rainfall amounts, onset and cession dates that currently cannot be achieved with much accuracy [14]. Hence, due to distances from the main observatory centers, farmers tend to believe that the rainfall experienced in the main observatory centers is not representative of the rainfall experienced in their regions.

### 1.3. Data for Informed Decision Making

Nyakach and Soin-Sigowett region of Kenya have two distinct wet seasons annually dictated by the Inter Tropical Convergence Zone. These are the “long” rain season which occurs in March-April-May (MAM) and the “short” rains in October-November-December (OND). Their crop and variety of crop they decide to plant depends on the season. The seasons have varying lengths, but the amount of rainfall is not distinctively different. Thus, having a local complete historical data for use can help researchers to work with them and make timely decisions on planting dates and the variety of crop.

In order to use climatic data for analysis, it should meet a threshold for which it can be considered as quality. For farmers in this study, completeness and accuracy in recording were important. The completeness was easily noticed by checking the recording sheet. Accuracy, however, is debatable given the spatial distance from the most trusted source, the KMS. In this paper we compare the most complete farmers data to KMS hence raise a discussion of their quality for use in decision support analysis, and models.

## 2. Methodology

This was principally a survey carried out by selected farmers over a period of eighteen (18) months. The farmers were provided with rain gauges which they used to record daily rainfall amounts. The rain gauges were manually drained whenever it rained. Focus group discussions was used to collect information on farmers' perception on previous rainfall information and the current rainfall events. To participate in the study, one-hundred (100) farmers were randomly selected from a list of one-thousand (1000) farmers who resided in the 10  $\text{km}^2$  region in Nyakach and Soin-Sigowett, where CCAFS operates. The one-hundred (100) farmers were chosen due to limited resources and were proportionally allocated according to the counties they resided, Kisumu or Kericho. Sixty (60) of the farmers were from Nyakach in Kisumu county while forty (40) from Soin-Sigowett in Kericho county.

## 2.1. Data

Three sets of rainfall data were used and collected in this study. These included long term daily rainfall, for period 1961 – 2004, provided by KMS for Kisumu and Kericho Meteorological Stations, daily rainfall data from thirty (30) volunteer stations in the area and farmer-recorded daily rainfall data for the period 2014 – 2015. The rainfall data from the volunteer stations varied in length for respective stations but were all within the period 1961 – 2015. Since farmers were not compelled to carry out the survey, there was selective attrition for some farmers due to various reasons which included theft and wear and tear of the rain gauges. Eventually, data was collected from forty-two (42) farmers who had recorded daily rainfall data for at least seventy-five percent (75%) of the period 2014 – 2015. In addition, qualitative data on farmers' recollection of previous rainfall events was collected through focus group discussions and plotted on charts. The objective of the analysis was to determine the quality of farmer-recorded rainfall data.

## 2.2. Data Analysis

The farmers' perception of rainfall events was discussed in terms of seasons, the "long" and "short" rains and compared with KMS data. Farmers recollection of rainfall received in the previous decade was scored as either "normal", "above normal" or "below normal". "Normal rain" was the farmers expectation of rainfall in a given season while "above normal" and "below normal" were scored when there were extreme wet and dry spells respectively. Using charts, seasons with "above normal" and "below normal" rainfall were marked with a cross, as shown in Figure 1 while normal rainfall was not marked. Farmers more often recalled extreme rainfall experiences and the more recent seasons. Rarely would they recall experiences more than thirty (30) years back. The farmers' recollection of rainfall events varied since some farmers considered the amount of rainfall in a

season while others considered the length of the season.

The KMS Kisumu rainfall data was summarized to seasonal totals and displayed using line plots since farmers in both Nyakach and Soin-Sigowett considered it more representative than KMS Kericho rainfall data. Kericho lies on a higher altitude and has significantly higher rainfall. A season whose rainfall total was more than the upper quartile amount was scored as "above normal" while seasons where the total rainfall amount was less than the lower quartile was scored as "below normal" otherwise they were scored as "normal". A rainy day is considered as one where the amount of rainfall received is more than 0.85 mm [16]. The second analysis considered the number of rainy days in a season and compared it with the farmers' recollection of extreme seasons. Finally, pairwise comparisons between farmer-recorded data and KMS daily rainfall data was done for a duration of eighteen (18) months, with an objective to validate the farmer-recorded data.

## 3. Results and Discussion

### 3.1. Farmer Perception of the Historical Climate Patterns

The first analysis considered the farmers perception of climate change and compared it to actual historical data. Climate change was considered as consistent change in weather pattern for at least thirty years. A participatory approach was used to collect their recollection for the past thirty years.

Farmers recalled their experienced over the past three decades and scored them on a chart as provided in Figure 1 and Figure 2. The figures have been divided into an upper part that gives their recollection for the long rains (MAM) and a lower part provides their recollection for the short rains (OND). They represented a case of above normal or below normal rainfall with cross (X). A season without an "X" was considered to be "normal".

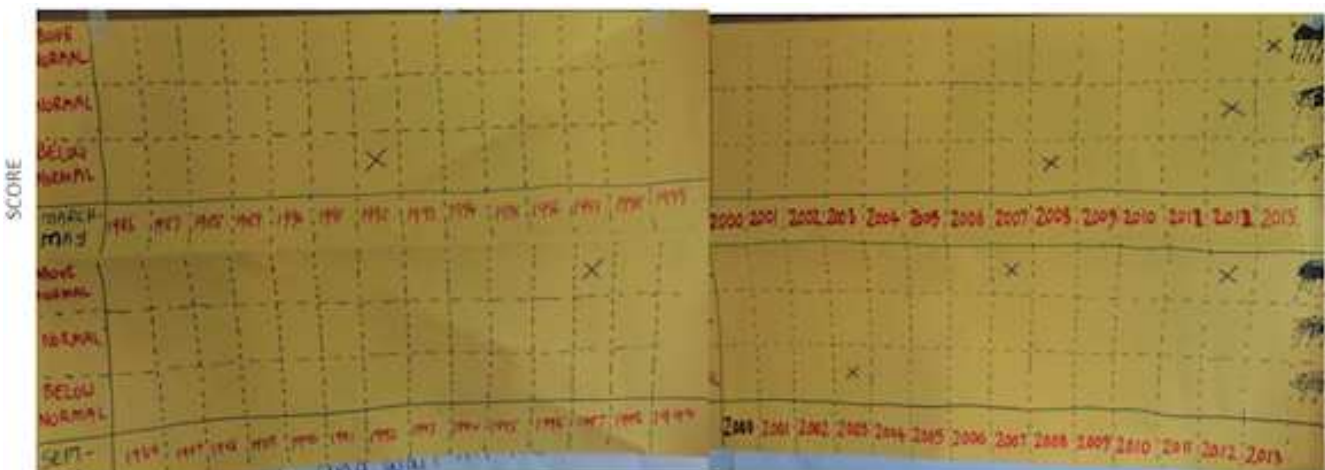


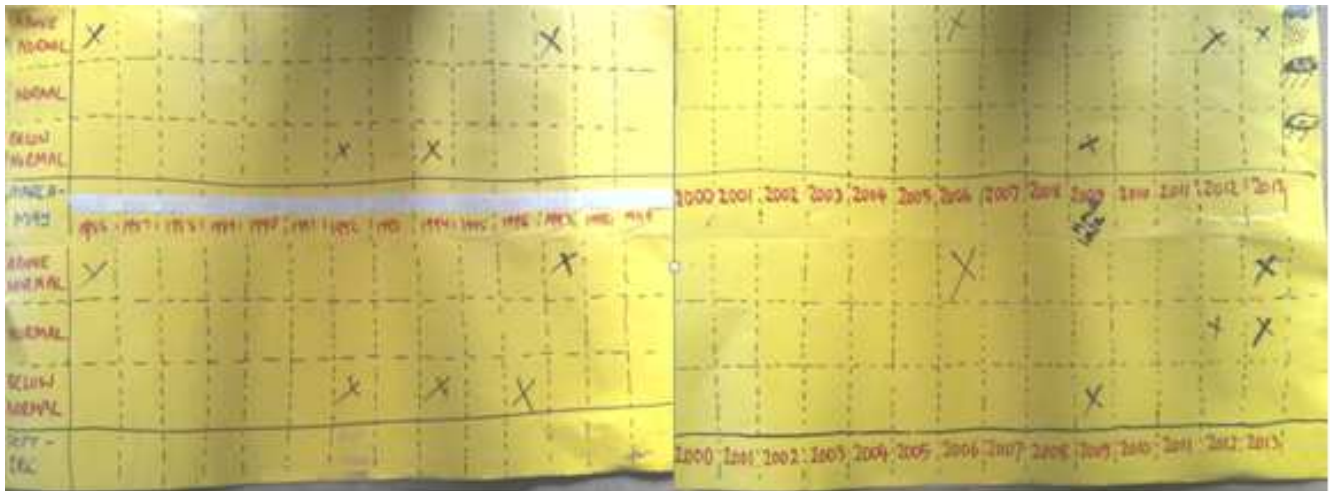
Figure 1. Farmer perception of rainfall occurrences over the last decade for Nyakach region.

Before the activity, most farmers believed that climate has changed, and there was generally less precipitation with incidences of extreme rainfall events. This

recollection helped them quantify and from Figure 1 and Figure 2 it was evident that going back, most of the seasons were within the normal range. Farmers from

Nyakach (Figure 1) recalled fewer extreme events unlike their counterparts in Soin-Sigowett (Figure 2). Farmers from Nyakach had more memories of extreme rainfall

compared to extreme drought. However, those from Soin-Sigowett had similar numbers of extreme rainfall and extreme drought.

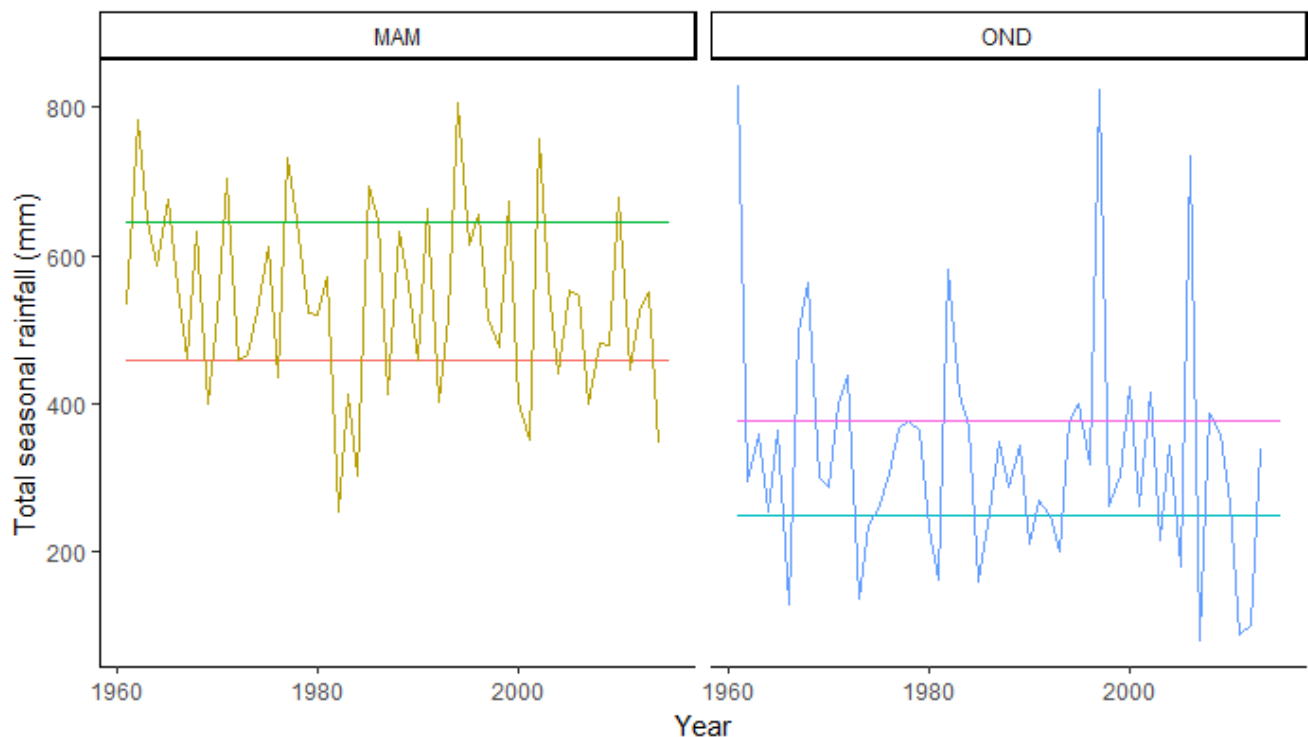


**Figure 2.** Farmer perception of rainfall occurrences over the last decade for Soin-Sigowett region.

### 3.2. Analysis of the KMS Kisumu Data

The recollection of farmers was interpreted by the amount of rainfall in a season, but for some it was length of the

season. Therefore, both total seasonal rainfall, and number of rainy days in a season were compared to the farmers' memories in Figure 1 and Figure 2.



**Figure 3.** Total seasonal rainfall for Kisumu long and short rains between 1961 and 2014.

The first step considered the total seasonal rainfall. Figure 3 shows line plots for cumulative rainfall for long and short rains. The horizontal lines are the upper and lower quartiles of the long-term seasonal rainfall for the two seasons. Any season whose total rainfall exceeded the upper quartile

amount was considered to be an extremely high rainfall event while the seasons whose total rainfall did not attain the lower quartile mark was classified as an extremely dry season.

Recollection of farmers from Nyakach matched two of the extreme events of the KMS Kisumu rainfall data; the El-Nino

rains experienced in the short rain season of 1997 and the drought experienced in the 2003 short rains. The data matched five extreme events that farmers from Soin-Sigowett had recalled.

Further, the number of rainy days in a season were compared to the farmers recollection. In Figure 4, the horizontal lines provide the 25<sup>th</sup> and the 75<sup>th</sup> percentiles of the number of rainy days using KMS data. The farmers from Nyakach had recalled the excess rain in the short

These were the high rainfall experienced in the short rain season of 1986, the long rain seasons of 2006 and 2007 and the dry spells experienced in the long and short rain seasons of 1992.

rains (OND) of 2007. The data showed that there was above average number of rainy days for 2006, but it dropped in the 2007 OND season. This shows that the farmers very well recollected the events albeit for slight differences in dates.

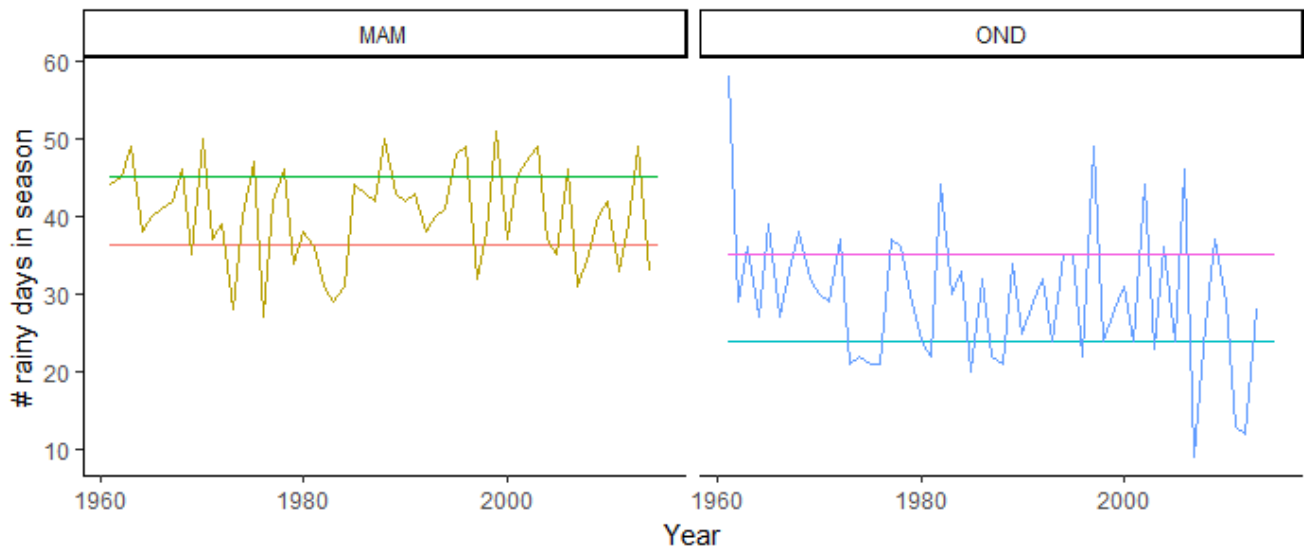


Figure 4. Number of rain days Kisumu long and short rains for fifty years.

Some farmers tend to link their lack of rainfall to the amount of crop yield. If not the yield, probable issues might have been the recollection of daily incidences with regard to a single day of extreme rainfall, as opposed to total seasonal rainfall.

### 3.3. Spatial Comparison of the Farmer Collected Rainfall Data to the KMS Kisumu Data

The above analysis compared the seasonal totals. Spatial differences might still be present in the daily data. It is evident that the farmers would record differing daily amount of rainfall in comparison to the Kisumu KMS station data. There was need to compare and see how different they are. The differences were analyzed in two-fold. The first was done by calculating the proportion of rain days over the eighteen-month period that had similar rainfall pairwise to the KMS Kisumu data. This was an exploratory procedure which was done for each farmer independently. The second method was done by conducting pairwise t-test to which farmers had statistically significantly different rainfall compared to the KMS Kisumu data.

For the first analysis, using the threshold of 0.85 mm as rainfall, any pairwise difference that fell within half of the threshold was considered to be similar. This daily difference in actual rainfall was calculated for comparison

on how much their local rain differed with KMS recorded rain. This was done for days that had recorded any amount of rainfall ( $> 0$  mm). The difference was calculated as

$$\text{diff} = \text{FarmerData} - \text{KMSData}.$$

The difference (diff) was categorized into three. This categorization considered half of the threshold for rainfall, that is, 0.425 mm. The three categories were

1. under-data: a case where the difference was less than "-0.425 mm". That is, the farmer experienced far less rainfall than the Kisumu KMS
2. over-data: a case where the difference exceeded "0.425 mm". That is, the farmer experienced far more rainfall than the Kisumu KMS, and
3. The rest were labelled near-data

Only 42 farmers' data was included for this comparison. They had shown a consistent collection of rainfall data. The proportions of rainy days under the three categories are summarized in Figure 5. In the figure, the farmers are divided according to a farmer group they belonged to. The groups FOKODEP and NECODEP were from Nyakach in Kisumu county. Farmers belonging to Kapsokale were from Soin-Sigowett in Kericho county.



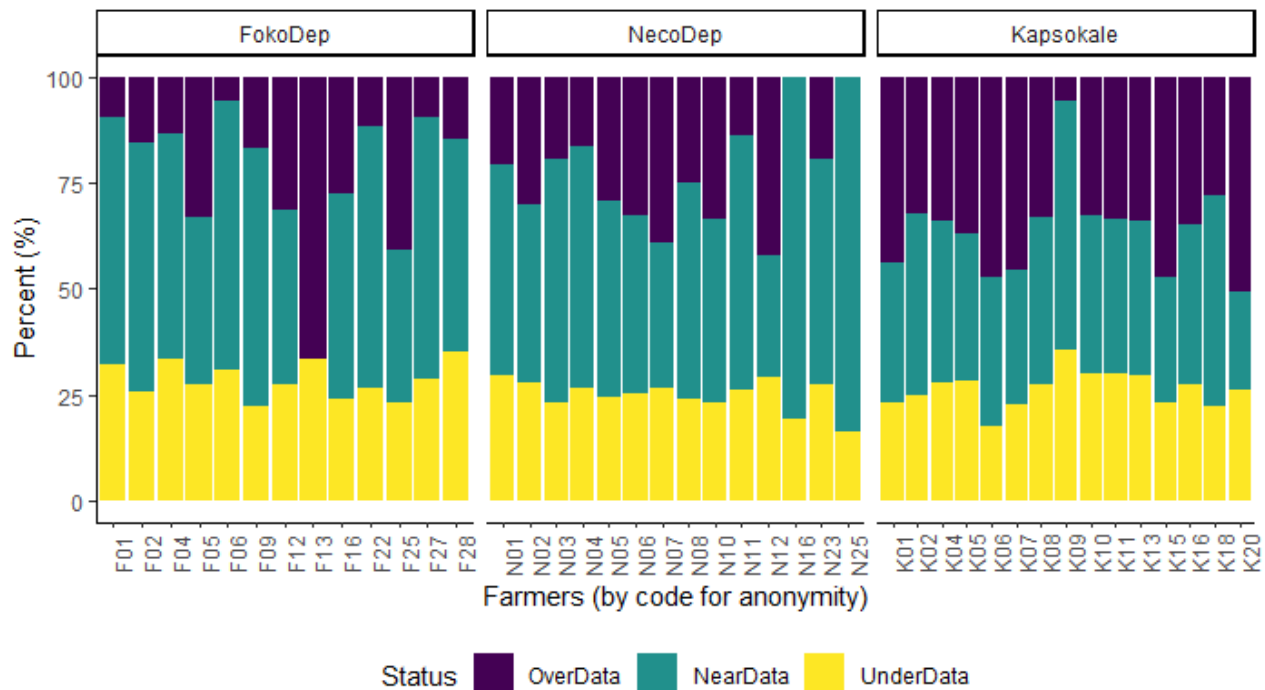


Figure 5. Comparing the daily rainfall data recorded by farmers to the KMS recorded data.

The analysis shows that in more than half the time, farmers' readings were close to that from KMS Kisumu. The farmers' data was less than the KMS data a quarter of the time and was greater than the KMS data in varying patterns for different farmers. Farmers from Kapsokale experienced slightly more proportions of higher rainfall than their counterparts from NECODEP and FOKODEP. This analysis shows that their local data was not too

different from the KMS data half of the time.

The region has volunteer stations whose data was available for analysis. The volunteer stations are on the outskirts of the farmers locations, but closer to them compared to Kisumu KMS. Data from thirty volunteer stations were available and used for the same analysis. Their results are provided in Figure 6.

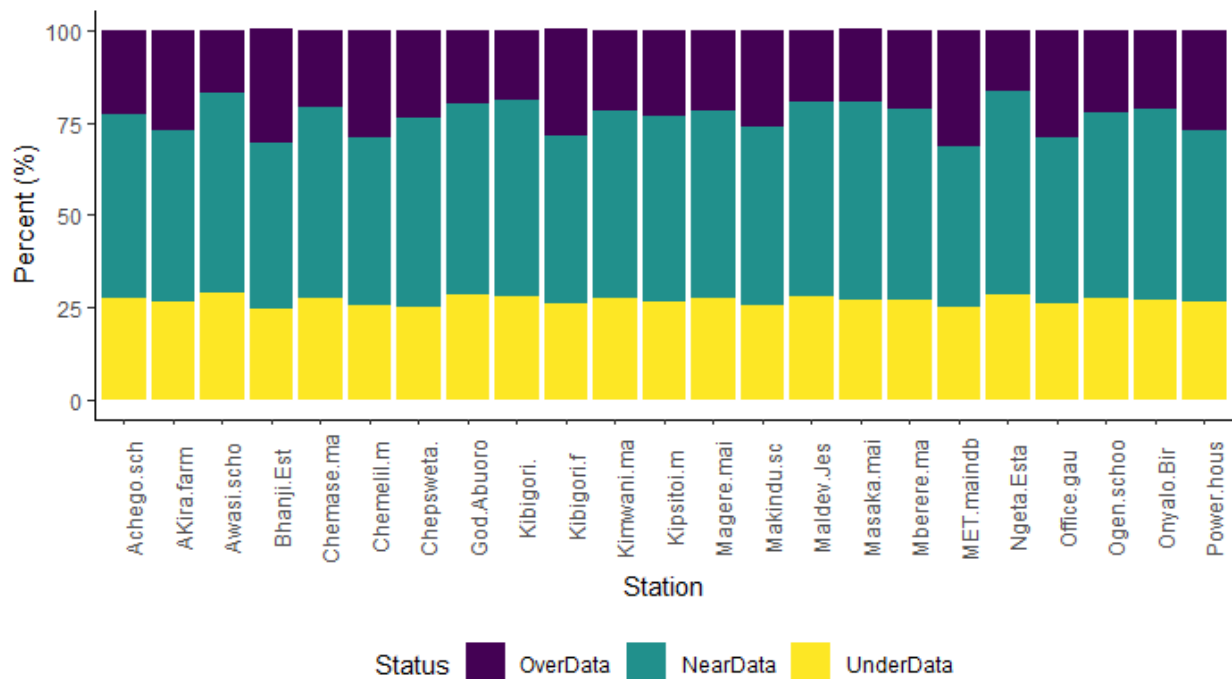


Figure 6. Comparing the daily rainfall data recorded by Volunteer stations to the KMS recorded data.

The volunteer stations had data for longer durations than the farmers data. Hence, was a better representation of long-term pattern of local rainfall for the region. Like with the farmers, rainfall recorded from volunteer stations was similar to that of Kisumu KMS. Long term, a quarter of the rainy days were having more rainfall and a quarter were having less rainfall. This goes to show that farmers' records well illustrated their local context. Micro-climatic differences were expected given the over twenty kilometers distance.

### 3.4. Comparing the Seasonal Farmer's Data to the KMS Data

Pairwise t-tests were conducted for each farmer's data to test for statistically significant differences in their records for each of the rainy days. For this, only days with positive rainfall were selected.

The test statistics for the paired sample t-test is  $t$  and is calculated as:

$$t = \frac{m}{s/\sqrt{n}}$$
 is the test statistic and

$$m = \frac{\sum_{i=1}^n d_i}{n},$$

and

$$s = \sqrt{\frac{\sum_{i=1}^n (d_i - m)^2}{n - 1}}$$

$n$  is the number of days used in the calculation with

$$df = n - 1,$$

and

$$d_i = \text{KMS Rain for Day } i - \text{Farmer Rain for Day } i$$

$$df = n - 1$$

One-tailed tests were used to test whether farmers were either getting more rainfall or they were getting less. The following results provides the farmers who had statistically significant test results. Only four out of the 42 farmers who had good enough data were seen to get significantly less rainfall compared to the Kisumu KMS. All the farmers were from the Nyakach region.

**Table 1.** Farmers with statistically significant lower rainfall than KMS rainfall pairwise.

farmer	df	p. val	diff	Se	LB	UB
F06	180	0.00	2.20	0.53	1.32	Inf
F22	180	0.01	1.32	0.60	0.34	Inf
N03	180	0.03	1.06	0.58	0.11	Inf
N11	149	0.03	1.24	0.63	0.20	Inf
N25	30	0.05	1.34	0.80	-0.01	Inf

A second series of one-tailed tests tested if the farmers received significantly higher rainfall than the Kisumu KMS. Again, only four farmers experienced significantly higher rainfall on a rainy day, compared to Kisumu KMS. Three of

the four farmers belonged to the higher altitude Kapsokale group.

**Table 2.** Farmers with statistically significant higher rainfall than KMS rainfall pairwise.

farmer	df	p. val	diff	Se	LB	UB
K01	120	0.01	-2.68	1.21	- Inf	-0.67
K06	90	0.01	-3.07	1.25	- Inf	-1.00
K07	109	0.00	-3.67	1.22	- Inf	-1.65
N07	119	0.02	-2.30	1.06	- Inf	-0.54

In summary, thirty-four out of the forty-two whose data was used for this test had similar rainfall to Kisumu rainfall. Only eight had rainy days that experienced significantly higher or lower rainfall compared to Kisumu KMS. Thus, Kisumu KMS data could be a good basis for discussing the climate in the Nyakach regions.

## 4. Conclusion

This study was an analysis of quality of farmers recorded daily rainfall data. Despite the challenges they faced, including theft scares, almost half of the farmers managed to record rainfall data that could be useful for analysis of local rainfall.

For most farmers, the data was similar to that of Kisumu KMS, with only eight having greater deviations in daily rainfall data. Thus, in case of lack of data, the Kisumu KMS data can be used, and it would represent most of the farmers experience, at least half of the time.

The data from farmers could be argued is representative of local rainfall. Further detailed analysis will be required before an individual farmer's data is considered to be a regional representative.

## Acknowledgements

The data used in this paper was collected under a CCAFS funded project in Nyando region. The project "Enhancing Climate Smart Agriculture in Nyando through innovative partnerships, Decision Support Tools and Climate Information Services" ran between 2014 and 2015. We are grateful to CCAFS for the support. We are also very grateful to the KMS Kisumu and Kericho offices who worked with us in analysis of their climate data to produce the outputs used in this paper.

## References

- [1] APSIM. (2019, 12 3). APSIM. Retrieved from APSIM: <https://www.apsim.info/>
- [2] Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Prat, O. P. (2015). *PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies*. Bulletin of the American Meteorological Society, 96 (1), 69-83.

- [3] Atta-Krah, A., & Francis, P. A. (1987). *The role of on-farm trails in the evaluation of composite technologies: The case of alley farming in Southern Nigeria*. Agricultural Systems, 23 (2), 133-152.
- [4] Behnke, R., Vavrus, S., Allstadt, A., Albright, T., Thogmartin, W. E., & Radeloff, V. C. (2016). *Evaluation of downscaled, gridded climate data for the conterminous United States*. Ecological Applications, 26 (5), 1338-1351.
- [5] CCAFS. (2019, 12 3). CCAFS. Retrieved from CCAFS: <https://ccafs.cgiar.org/regions/east-africa>
- [6] Dietrich, H., Wolf, T., Kawohl, T., Wehberg, J., Kandler, G., Mette, T.,... Bohner, J. (2019). *Temporal and spatial high-resolution climate data from 1961 to 2100 for the German National Forest Inventory (NFI)*. Annals of Forest Science, 76 (1), 6.
- [7] Gadgil, M., Berkes, F., & Folke, C. (1993). *Indigenous Knowledge for Biodiversity Conservation*. Ambio, 22 (no. 2/3), 151-56. Retrieved from [www.jstor.org/stable/4314060](http://www.jstor.org/stable/4314060)
- [8] Jones, D. A., & Wang, W. F. (2009). *High-quality spatial climate data-sets for Australia*. Australian Meteorological and Oceanographic Journal, 58 (4), 233.
- [9] KMS. (2019, 12 3). KMS. Retrieved from KMS: <http://www.meteo.go.ke/>
- [10] Kyei-Mensah, C., Kyerematen, R., & Adu-Acheampong, S. (2019). *Impact of Rainfall Variability on Crop Production within the Worobong Ecological Area of Fanteakwa District, Ghana*. Advances in Agriculture, 1-5.
- [11] Lightfoot, C., & Barker, R. (1988). *On-farm trials: a survey of methods*. Agricultural Administration and Extension, 30 (1), 15-23.
- [12] Luschei, E. C., Wychen, L. R., Maxwell, B. D., Bussan, A. J., Buschena, D., & Goodman, D. (2001). *Implementing and conducting on-farm weed research with the use of GPS*. Weed Science, 49 (4), 536-542.
- [13] Rockström, J., Kaumbutho, P., Mwalley, J., Nzabi, A. W., Temesgen, M., Mawanya, L., Damgaard-Larsen, S. (2009). *Conservation farming strategies in East and Southern Africa: yields and rain water productivity from on-farm action research*. Soil and Tillage Research, 103, 23-32.
- [14] Shilenje, Z. W., & Ogwang, B. A. (2015). *The role of Kenya meteorological service in weather early warning in Kenya*. International Journal of Atmospheric Sciences, 2015, 1-5.
- [15] Stabinsky, D., & Brush, S. B. (1996). *Valuing local knowledge: indigenous people and intellectual property rights*.
- [16] Stern, R., & Coe, R. (1984). *A model fitting analysis of daily rainfall data*. Journal of the Royal Statistical Society: Series A (General), 1-18.
- [17] Stern, R., & Cooper, P. (2011). *Assessing climate risk and climate change using rainfall data--a case study from Zambia*. Experimental Agriculture, 241-266.
- [18] Wang, T., Hamann, A., Spittlehouse, D. L., & Murdock, T. Q. (2012). *ClimateWNA—high-resolution spatial climate data for western North America*. Journal of Applied Meteorology and Climatology, 51 (1), 16-29.