Uncertainties in the Estimation of Global Observational Network Datasets of Precipitation over West Africa

Matthew OJ, Abiye OE, Summonu LA, Ayoola MA and Oluyede OT

1Institute of Ecology and Environmental Studies, Obafemi Awolowo University, Ile-Ife, Nigeria
2Centre for Energy Research and Development (CERD), Obafemi Awolowo University, Ile-Ife, Nigeria
3Department of Physics and Engineering Physics, Obafemi Awolowo University, Ile-Ife, Nigeria

Abstract

This study assessed the uncertainty in estimating long-term (1971-2010) mean precipitation, its inter-annual variability, and linear trend of three network observation datasets over West Africa. A reference data, defined as a multi-dataset ensemble of precipitation observations of the Climate Research Unit (CRU) of the University of East Anglia, the Global Precipitation Climatology Centre (GPCC) and the University of Delaware (UDEL), all at horizontal resolutions of 0.5° by 0.5° were obtained and used in this study. Uncertainties in these climatological parameters of precipitation at both annual and seasonal time scales were examined in terms of inter-dataset variability using signal-to-noise ratio (SNR), correlation, root-mean-square errors and the normalised standard deviation. Results showed that the mean, inter-annual variability and trends climatology varied for different datasets. The three datasets had good agreement (SNR > 5) in terms of the annual mean precipitation and its inter-annual variability in most parts of West Africa. However, the agreement between the datasets was poor in the very dry Sahel parts of northern Niger, Mali, and Mauritania (SNR ≤ 1) due to very little precipitation and possibility of relatively low station density in these regions of complex terrain. In terms of correlation (0.89 ≤ r ≤ 0.98), and normalised standard deviation, NSD (0.8 ≤ NSD ≤ 1.7), the uncertainties in the spatial variations in linear trend were larger than mean precipitation and their inter-annual variability for both annual and seasonal scales. The long-term annual precipitation trend in the region is highly uncertain except in a few small areas.

Keywords: Global network datasets; Uncertainties; Reference data; SNR; West Africa

Introduction

All measurements have some degree of uncertainty that may come from a variety of sources. The term "uncertainty" is used to refer to a possible value that an error in a measured dataset may have [1,2]. In science, true measurement value doesn’t exist; what we usually have is an estimate. An error in the estimate is the amount of inaccuracy in the estimate compared with a known standard value [3-5]. Sources of uncertainty in observational measurements are often broadly categorized as (statistical) random or systematic errors [6,7]. Random observational errors are associated with the observed frequency distributions in the primary data. They occur for many reasons ranging from misreading of the pointer instrument, rounding errors, the difficulty of reading the instrument to a precision higher than the smallest marked gradation, incorrectly recorded values, errors in transcription from written to digital sources, and sensor noise among others. The contribution of random independent errors to the uncertainty on the average/ensemble of a number of datasets is much smaller than the contribution of random error to the uncertainty on a single observation, even in the most sparsely observed years [6]. Nonetheless, where observations are few, random observational errors can be an important component of the total uncertainty and can, in principle, be made arbitrarily small by repeated measurement or large enough sample size [7].

Uncertainties due to systematic effects include calibration error, poorly-sited instrument, and background uncertainties and almost everything else that might bias a measurement, which are caused by a lack of knowledge or uncertainty in the measurement model, such as the reading error of an instrument as well as uncertainties in the interpretation of a measurement which cannot be made arbitrarily small by simply getting more data. Kennedy [6] submitted that systematic observational errors are much more problematic, because their effects become relatively more pronounced as greater numbers of observations are aggregated. In this case, the author believed that averaging observations together from many different instruments/sources would tend to reduce the contribution of systematic observational errors to the uncertainty of the average.

A number of researchers and institutions have developed observation-based gridded analysis datasets of global or regional coverage with fine spatial resolutions [8-14]. These network of observation datasets provide precipitation and/or surface air temperatures over extended periods of multiple decades at spatial resolutions of 0.5° or finer. This is, of course, a substantial improvement over previous generation data sets that are typically at much coarser (e.g. 2.5°) horizontal resolutions [15]. These recent fine-scale datasets allow us to better examine the regional precipitation and temperature climatology and to perform more reliable evaluations of today’s high-resolution climate simulations, especially over the regions of complex terrain, that are important for climate-change impact assessments and climate model evaluations [16].

*Corresponding author: Matthew OJ, Institute of Ecology and Environmental Studies, Obafemi Awolowo University, Ile-Ife, Nigeria, Tel: +234 (0)703 3312 738; E-mail: abefematt@yahoo.com

Received August 04, 2017; Accepted August 28, 2017; Published August 31, 2017


Copyright: © 2017 Matthew OJ, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
The commonest two forms of uncertainties in these observation-based gridded analysis datasets are those that arise from the estimation of area-averages from a finite number of noisy and often sparsely distributed observations [17]. The grid-box sampling uncertainty refers to the uncertainty accruing from the estimation of an area-average data anomaly within a grid box from a finite, and often small, number of observations. Large-scale sampling uncertainty refers to the uncertainty arising from estimating an area-average for a larger area that encompasses many grid boxes that do not contain observations. Although these two uncertainties are closely related, it is often easier to estimate the grid box sampling uncertainty (variability within a grid box), than the large-scale sampling uncertainty (consideration of the rich spectrum of variability at a global scale). There are many scientifically defensible ways to produce a large scale observational data set. For example, one might choose to fill gaps in the data by projecting a set of Empirical Orthogonal Functions (EOFs) onto the available data or fill the data using simple optimal interpolation [6]. The choice of methods used, inputs and the foundational assumptions made during data set creation, however, produced a wide spread of results and different structural uncertainties as detailed in Thorne et al. [18].

Dunn et al. [19] in their assessment of the effects of different methodological choices made during the construction of HadEX2 gridded data sets of climate extremes showed that the number of input stations networks (required) in or around a grid box and the gridding method used had significant effects on individual grid boxes. Robust correlations on grid box scales were obtained in areas which have high station density. They also submitted that precipitation indices, being less spatially correlated, can be more susceptible to methodological choices, but coherent changes are still clear in regions of high station density. Thus, understanding the uncertainties present within a data set can enable better decision making, as well as enhancing further research applications [20].

Recent researches, however, suggested that there exist substantial amounts of differences amongst today's gridded precipitation datasets (resulting in uncertainties in the calculated precipitation climatology) and the uncertainty and the spread amongst multiple data sets vary according to regions as well as seasons [15,21,22]. The studies further revealed that uncertainties in the calculated precipitation climatology defined relative to their climatological means are generally larger in the dry regions and/or local dry seasons. In addition, uncertainty due to the differences between various datasets needs to be examined and quantified in all climate studies because the absolute accuracy of individual dataset cannot be quantified in practice. However, analysis of uncertainty in the previously used global network datasets for climate studies and assessment of inter-dataset differences have not been extensively documented so far over the West African region. In this study, we investigate the uncertainty in calculating fundamental properties of regional climate characteristics of precipitation for three global observational network datasets (which have been commonly used in recent past for validation of regional model simulations) between 1971 and 2010 over West Africa. It examined, for the first time, the uncertainty in assessing the key precipitation characteristics of the Climate Research Unit of the University of East Anglia, the Global Precipitation Climatology Centre and the University of Delaware. This was with a view to providing useful information for better interpretation and analysis of future precipitation projections over the region.

Materials and Methods

The study area

Figure 1 is a sketch map of West Africa showing the geographical locations of the member countries.

The network observation datasets

The University of Delaware (UDEL) version 3.01 observation dataset consists of monthly climatology of precipitation at 0.5° x 0.5° latitude/longitude grid, at a time series spanning 1900 to 2010. The data were put together by Matsuura and Willmott [23] of the University of Delaware and updated in National Centre for Atmospheric Research, NCAR

Figure 1: A sketch map of West Africa showing the geographical locations of the member countries.
The ensemble mean precipitation climatology

Figure 2 presents the annual mean, inter-annual variability, and trends of the ensemble mean precipitation of CRU, GPCC and UDEL datasets over West Africa for the 1971-2010 period. The spatial distribution of the mean annual precipitation in the region is characterized by the wetter regions in south and a decrease towards the north. For example, annual total precipitation of about 3000 mm were recorded in the coastal areas of the Guinea Republic, Guinea (Bissau), Sierra Leone and Liberia as well as some locations in south-southwest Nigeria (Figure 2a). The annual precipitation decreased from about 1400 mm in most central parts of Nigeria, Benin, Togo, Ghana and Senegal to about 850 mm in the north. The driest region in West Africa with precipitation of about 50 mm or less per annum covers interior northern parts of Niger, Mali, and Mauritania. The inter-annual variability of the annual precipitation also shows similar distribution as the annual means with the highest variations (300 mm or more) in the south/coastal zones and the least about 50 mm in the north (Figure 2b). Linear trend of the annual precipitation varies substantially across the region. The most notable features include the positive trend (10 mm/year) in most parts of the region particularly in the driest region, including Niger, Mali, and Mauritania, and about 20 mm/year in a few locations in Nigeria and Liberia (Figure 2c). Maximum negative trends of about 20 mm/year were obtained over most parts of Sierra Leone and along the wet coastal region of Guinea, Côte d’Ivoire, Ghana and Nigeria.

The seasonal cycle of ensemble rainfall climatology (Figure 3) resembles the annual mean climatology but the magnitudes varied across the months with a peak in August. The rainfall (>100 mm/month)
starts from the coast in March, moves to the north progressively to reach its northern limit in the Sahel in the month of August (maximum precipitation of about 600 mm in the coastal locations and about 100 mm or less in the Sahel) and then retreats to the south. Some coast regions such as Guinea, Côte d’Ivoire, Liberia, Ghana and Nigeria had two peaks; one around June/July and a second in September, with a ‘little dry season’ depicted by reduced precipitation in the month of August. The dry season occurs between November and February in most parts of the region. Figure 4 shows the spatial distribution of the inter-annual variability of the monthly precipitation during the study period. The variations were about 150 mm or more in the wet season and less than 20 mm in the dry season. Linear trends of the monthly precipitation as depicted in Figure 5 varied substantially over the region with both positive and negative trends across the months. Notable changes are the positive trends (25-50 mm/year) and the negative trends (-25 to -50 mm/year) in monthly precipitation over most parts of West Africa in the months of August and May respectively.

**Uncertainties in precipitation climatology**

The signal-to-noise-ratios (SNRs) for estimating the linear trend of the annual mean precipitation, for the individual datasets, over the 40-year period are presented in Figure 6. The results showed that SNR exceeded 5 in most of the study domain for the estimation of the annual mean (Figure 6a). Only very few locations in the northern parts of Mali, Niger and Mauritania had small but reliable SNR (1<SNR<5) values. The SNR for the inter-annual variability over the region was also larger than 5 but generally smaller than that for calculation of the mean (Figure 6b). For the estimation of trend, however, SNR values were generally too low (≤2) (Figure 6c). Figures 7-9 show the SNR for the seasonal mean precipitation of the three climatological properties (mean, inter-annual variability and the trend). The values for the mean property are somewhat smaller than those for the annual precipitation climatology and they exceed 5 in about the same region as for the annual precipitation, particularly during the wet season (April/May-October) (Figure 7). For the inter-annual variability (Figure 8), the spatial distribution of SNRs is similar to those obtained for the mean climatology but lower in magnitudes. However, the SNRs for estimation of seasonal linear trend were generally too low (0<SNR ≤1) (Figure 9).

The statistical significance of trends as described by the estimated p-values from each datasets in calculating the linear trend of the annual mean precipitation are shown in Figure 10. The p-values were large (>0.5) in most parts of West Africa for the three datasets. It was also observed that these regions of large p-values correspond to the regions of large SNR values. The patterns of spatial variations were very similar for both GPCC and UDEL observations with only few places, particularly in the northern Niger, Mali, and Mauritania having low p-values (<0.1) (Figures 10b-10c). However, GPCC dataset appear to have more areas with high p-values than UDEL observations. The CRU observation had more locations with low p-values (<0.1) than the other two. The CRU observations had low p-values in most parts of the country like Niger, Mali, Mauritania, Nigeria, Guinea, Côte d’Ivoire, Liberia and Sierra Leone (Figure 10a). Furthermore, Figures 11-13 illustrate statistical significance p-values in calculating the linear trend of the seasonal mean precipitation for CRU, GPCC and UDEL respectively. The results indicate that the CRU observations had large p-values in most parts of West African region during the wet season only (Figure 11). However, the p-values were large over the region (and beyond) during all months (Figures 12 and 13).

The Taylor diagrams in the Figure 14 depict the spatial variations in the mean, inter-annual variability and the trend of the three observational datasets (using the simple multi-dataset ensemble as the

---

**Figure 3:** The mean of seasonal precipitation (of the ensemble mean of CRU, GPCC and UDEL) for the period 1971-2010 over West Africa.
Figure 4: The inter-annual variability of seasonal precipitation (of the ensemble mean of CRU, GPCC and UDEL) for the period 1971-2010 over West Africa.

Figure 5: The time trend of seasonal precipitation (of the ensemble mean of CRU, GPCC and UDEL) for the period 1971-2010 over West Africa.
Figure 6: The signal-to-noise ratio (SNR) in estimating the (a) mean, (b) inter-annual variability and (c) the trend of annual precipitation of the ensemble mean climatology.

Figure 7: The signal-to-noise ratio (SNR) for the mean seasonal precipitation of the ensemble mean climatology.
Figure 8: The signal-to-noise ratio (SNR) for the inter-annual variability in seasonal precipitation of the ensemble mean climatology.

Figure 9: The signal-to-noise ratio (SNR) for the linear time trend in seasonal precipitation of the ensemble mean climatology.
Figure 10: The $p$-values in estimating the linear trend in annual mean precipitation in (a) CRU, (b) GPCC and (c) UDEL datasets.

Figure 11: The $p$-values in estimating the linear trend in seasonal mean precipitation for CRU dataset.
Figure 12: The p-values in estimating the linear trend in seasonal mean precipitation for GPCC dataset.

Figure 13: The p-values in estimating the linear trend in seasonal mean precipitation for UDEL dataset.
The areas encompassed by the green polylines showed the levels of spread or the range of uncertainty in terms of both standardized deviation and correlation in the datasets. Larger areas were obtained for the spatial variations in linear trend (Figures 14e,14f) than mean precipitation and their inter-annual variability for both the annual and 'wet season' (Figures 14a-d). However, the areas are small and very similar to those for their inter-annual variability (Figures 14c,14d).

Discussion

Results indicated that the mean and inter-annual variability in annual distribution of precipitation in West Africa is generally zonal and the amount per year generally decreases inland from the coast. These results are in agreement with the findings of previous independent works over the region [33-37]. The increase in precipitation from the Sahara to the humid equatorial zone was assumed to relate to a rapidly increasing depth of the moist layer equator-ward from the Intertropical Convergence Zone, ITCZ [33]. The fact that the coastline areas remain under the influence of Tropical Maritime (mT) accounts for why locations like Conakry (Guinea) Lagos (Nigeria), Freetown (Sierra Leone), and Monrovia (Liberia) experience the highest annual mean rainfall amounts [38-42]. Conversely, very little rainfall in the north could be attributed to the influence of the dominant warm, dry Tropical Continental (cT) air.

The seasonal cycle of ensemble precipitation climatology suggested progressive movement of rainfall from the coast in March to reach its northern limit in the Sahel in the month of August and then retreats to the south in October. These patterns follow the position and migration of the ITCZ and its associated air masses which bring rainfall from tropical moist oceanic air that moves from the Atlantic Oceans inland. The observed reduction in rainfall in the month of August particularly in some coast regions such as Guinea, Côte d’Ivoire, Liberia, Ghana and Nigeria is likely connected with the relative stability that exists over the coastal area as a result of lower sea surface temperature and divergence of specific humidity [43-46]. The inter-annual variability of the monthly precipitation during the study period was higher in the wet season (April-October). This shows that the precipitation climatology over the West Africa region is primarily determined by the ‘summer’ (May-October) rainfall. It is so because West Africa’s precipitation regime is characterized by the influence of the mT air mass which is dominant over most parts of the region during this period of wet season [33].

Linear trend of the annual precipitation varied across the region with prominent positive trend in most parts of the region particularly in the driest region and a few wet coastal locations in Sierra Leone, Guinea, Côte d’Ivoire, Ghana and Nigeria. The observed trends and changes in the precipitation mean interannual variability could be attributed to changes in the main tropical circulation features associated with the West African monsoon such as the African easterlies, tropical easterlies, African westerly (over the continent) and the West African westerly jet (over the Atlantic) as reported by Nicholson [33].

Furthermore, results showed that the mean, inter-annual variability and trends climatology varied for different data sets. This is inevitable because each data set utilizes different raw data, data quality control, and analysis methodology [15,47-49]. For example, the SNRs for the annual mean precipitation and its inter-annual variability exceeded 5 in most parts of the study domain. Hence, the three datasets examined in this study agree well in terms of the annual mean precipitation and its inter-annual variability in the West African region. However, the...
SNRs were generally lower for the inter-annual variability than for the mean. This suggests that uncertainty in calculating the inter-annual variability is larger than in calculating the mean climatology. The agreement between the datasets are poor in the regions such as very dry Sahel parts of northern Niger, Mali, and Mauritania with low SNRs. These low values of SNR could be attributed to the so little precipitation recorded over the regions which made it very difficult to obtain reliable trends. According to Kim et al. [22], precipitation measurements in these regions require high-density networks. Thus, possibility of relatively low station density in these regions of dry climate and complex terrain could also be another reason for the poor agreement. Conversely, the SNR for the linear tendency of the annual precipitation were generally too low in most regions. This is an indication that long-term annual precipitation trend in the region is highly uncertain except in a few small areas.

The dependability of the three characteristics of the seasonal precipitation (i.e. mean, inter-annual variability and trend) calculated from these datasets are similar to that of the annual precipitation. The results showed that the three datasets agree more closely for the seasonal mean values than for the inter-annual variability. In all, the regions of large SNR correspond to the regions of small p-values in calculating the linear trend. This suggests that some of the uncertainty in the multi-dataset ensemble may be inherited from the uncertainty in calculating the trend from individual data sets. Similarly, a significant portion of the region of small p-values shows small SNR values. This also suggested that inter-dataset differences could be the main cause of the uncertainty in calculating long-term trends. Unlike the CRU monthly mean precipitation which was only significant in the raining season, GPCC and UDEL showed high significant p-values in all months.

The results showed that the uncertainties in the spatial variability occurred in both the spatial pattern and the magnitude for the annual and wet season mean. The uncertainties in the spatial variations in linear trend were larger than mean precipitation and their inter-annual variability for both the annual and raining season. However, there was similar spread in the inter-annual variability of the annual and wet season precipitation climatology in the region. In terms of the correlation, which ranged between 0.89 and 0.98, the results showed consistency in the spatial pattern between individual datasets and the reference data measured. The ranges of spatial variability measured in terms of the normalised standard deviation were higher for the linear trend than the range of the mean and the inter-annual variability. These results are indications that these datasets are affected by some common factors in determining their mean and inter-annual variability. Some of these factors include, variations in data analysis and interpolation schemes employed [21,47]. Similarly, high errors in the use of long-term trends of these datasets must be noted in future climate studies over the region.

Summary and Conclusion
The uncertainties in estimating both annual and seasonal mean climatology, standard deviation and trend of precipitation for three commonly used network observation datasets for regional model validations (i.e., CRU, GPCC and UDEL) has been examined over West African region for 1971-2010 period. To achieve this objective, an ensemble of these three gridded analysis datasets was obtained. The statistical indices used include signal-to-noise ratio (SNR), correlation, root-mean-square error and standardised standard deviation. Results showed that the inter-annual variability of the monthly precipitation during the study period were higher in the wet season (April/May-October), indicating that the precipitation climatology over the West African region is primarily determined by the wet season rainfall. This is an indication of significant influence of the movement of ITCZ on the occurrence of rainfall in West Africa and the dominance of Tropical maritime air mass over most parts of the region during the wet season. The linear trend of the annual precipitation varied across the region with prominent positive trend in most parts of the region particularly in the driest region and a few wet coastal locations in Sierra Leone, Guinea, Côte d’Ivoire, Ghana and Nigeria. This was attributed to the changes in the main tropical circulation features associated with the West African monsoon. Furthermore, the mean climatology of the annual and seasonal mean precipitation values and their inter-annual variability are highly reliable in most parts of West Africa except in the very dry Sahel parts of northern Niger, Mali, and Mauritania. This was evident from the facts that SNR values exceeded 5 in large parts of the region. In addition, some of the regions of large SNR had small p-values in calculating the linear trend and a significant portion of the region of small p-values showed small SNR values. These are indication that some of the uncertainty in the multi-dataset ensemble may be inherited from the uncertainty in calculating the trend from individual data sets and those inter-dataset differences could be the main cause of the uncertainty in calculating long-term trends. Finally, the uncertainty characteristics varied according to the climatological properties. Except for the linear trends, the climatological mean values and inter-annual variability estimated over the 40-year period are highly reliable in most areas in the region.

Acknowledgements
The authors are grateful to the Climate Research Unit of the University of East Anglia, the Global Precipitation Climatology Centre and the University of Delaware for providing the fine resolution observational network datasets used in this study.

References