RMetS
Royal Meteorological Society

Combined use of satellite estimates and rain gauge observations to generate high-quality historical rainfall time series over Ethiopia

Tufa Dinku,[a]* Kinfe Hailemariam,[b] Ross Maidment,[c] Elena Tarnavsky[c] and Stephen Connor[d]

[a] *International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, NY, 10964-8000, USA*
[b] *National Meteorological Agency, Addis Ababa, Ethiopia*
[c] *Department of Meteorology, University of Reading, UK*
[d] *Honorary Fellow, School of Environmental Sciences, University of Liverpool, UK*

ABSTRACT: Climate data are used in a number of applications including climate risk management and adaptation to climate change. However, the availability of climate data, particularly throughout rural Africa, is very limited. Available weather stations are unevenly distributed and mainly located along main roads in cities and towns. This imposes severe limitations to the availability of climate information and services for the rural community where, arguably, these services are needed most. Weather station data also suffer from gaps in the time series. Satellite proxies, particularly satellite rainfall estimate, have been used as alternatives because of their availability even over remote parts of the world. However, satellite rainfall estimates also suffer from a number of critical shortcomings that include heterogeneous time series, short time period of observation, and poor accuracy particularly at higher temporal and spatial resolutions. An attempt is made here to alleviate these problems by combining station measurements with the complete spatial coverage of satellite rainfall estimates. Rain gauge observations are merged with a locally calibrated version of the TAMSAT satellite rainfall estimates to produce over 30-years (1983-todate) of rainfall estimates over Ethiopia at a spatial resolution of 10 km and a ten-daily time scale. This involves quality control of rain gauge data, generating locally calibrated version of the TAMSAT rainfall estimates, and combining these with rain gauge observations from national station network. The infrared-only satellite rainfall estimates produced using a relatively simple TAMSAT algorithm performed as good as or even better than other satellite rainfall products that use passive microwave inputs and more sophisticated algorithms. There is no substantial difference between the gridded-gauge and combined gauge-satellite products over the test area in Ethiopia having a dense station network; however, the combined product exhibits better quality over parts of the country where stations are sparsely distributed.

KEY WORDS   climate; data; Ethiopia; satellite; rainfall; merging; interpolation

Received 20 August 2012; Revised 8 September 2013; Accepted 30 September 2013

1. Introduction

There is a growing understanding that climate variability and change pose serious challenges to development, particularly in Africa. The availability of decision-relevant climate information, at national and community level, is critical in assessing climate threats and opportunities and for integrating climate issues into development practices. Long-term, temporally homogeneous time series of climate data with good spatial coverage are therefore needed in many applications including climate risk management and adaptation to climate change. However, availability of reliable climate information has been one

of the major constraints to the use of climate information in Africa (Sperling, 2003; IRI *et al.*, 2006). The conventional source of climate data has been measured at weather stations. The number and quality of weather stations, however, have been declining over many parts of Africa (Washington *et al.*, 2006). The available stations are unevenly distributed and mainly located along main roads in cities and towns.

To overcome the lack of rain gauge measurements over many parts of Africa, many users resort to satellite rainfall estimates. Different techniques exist to retrieve rainfall from satellite thermal infrared (TIR) and passive microwave (PMW) sensors and combination of the two (e.g. Kidd and Huffman, 2011; Kidd and Levizzani, 2011). In terms of studying climate, currently available data sets suffer from a number of shortcomings that include short time series, coarse spatial and temporal resolutions, temporal inhomogeneity and sometimes of poor quality. Many of the most successful algorithms, such as those given in Section 2, combine data from

* Correspondence to: T. Dinku, International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, NY 10964-8000, USA. E-mail: tufa@iri.columbia.edu

different sources with different weightings for each year depending on data availability. While this approach may provide the most accurate estimates for any 1 year, the inter-annual variations may be influenced as much by the different mix of inputs as by actual changes in rainfall amount so that trends and variability statistics will be inaccurately portrayed.

A temporally homogeneous rainfall data set covering over 30-years time period is generated for Africa from the archive of Meteosat TIR imagery calibrated against local gauge data by the TAMSAT (Tropical Applications of Meteorology using SATellite data and ground-based observations) research group. Here, a case study over Ethiopia is presented where this approach is locally calibrated and the resulting dekadal (ten-daily) rainfall estimates over Ethiopia are combined with observations from rain gauges. Many satellite products already include rain gauge measurements for removing biases and merging the gauge observations with the satellite estimates. However, for most of Africa only the synoptic observations obtained through the Global Telecommunications System (GTS) are generally available for these adjustments. In practice, many synoptic stations are silent or report only sporadically (Washington *et al.*, 2006). For instance, only about 20 stations are accessible from Ethiopia through the GTS. However, as with many other countries, Ethiopia has a much greater density of gauges available within the country. In this work, data from a network of over 500 gauges have been used to produce a gridded rain gauge data set, to calibrate the TAMSAT rainfall estimation algorithm based on only TIR data from Meteosat and using the gauge observations, and to generate a bias adjusted product. The final output is a historical time series of ten-daily rainfall products at spatial resolution of 10 km. An added advantage of these gridded products is that they can easily be combined with other relevant data and analysed in a Geographic Information System (GIS) environment.

This work has been a collaborative effort of the International Research Institute for Climate and Society (IRI) at Columbia University with the National Meteorological Agency (NMA) of Ethiopia and the TAMSAT group at the University of Reading, UK. NMA staff members were trained in data quality control, gridding, satellite rainfall calibration and most of the work has been done at NMA headquarters. This aims to ensure the capacity to sustain the work beyond the life of the initial project.

As mentioned earlier, access to available climate data is a big constraint to using climate information in Africa. However, producing quality climate information alone may not lead to the use of the data. The data or derived information products should be made available to the wider user community so they can assess its practical and economic worth. The Ethiopian work includes this provision; NMA's web page has been improved to add tools for viewing and basic analyses of the data (http://www.ethiomet.gov.et/). Virtual 'map rooms' have been created for climate analysis, climate monitoring and climate forecast and are available to the public. The Ethiopian project may serve as a template to improve the availability and use of climate information in Africa. This article describes the methodology used for combining station data with locally calibrated version of the TAMSAT satellite-based rainfall estimates. Readers interested in the overall description of the project may refer to Dinku *et al.* (2011), and for further information on the TAMSAT approach for rainfall estimation Grimes *et al.* (1999).

2. Study region and data

2.1. Study region

This work is centred over Ethiopia, part of the Horn of Africa. Ethiopia has the most complex topography on the continent. Elevation ranges from below sea level in the northeastern part of the country to 4620 m just a few hundred kilometers to the west (Figure 1). The East African Rift Valley together with the surrounding plateaus and mountain chains form the main topographic features. The major rain-producing feature is the Inter Tropical Convergence Zone (ITCZ); however, topography plays a significant role by creating diverse microclimates ranging from hot deserts over the low lands to very cold mountain ranges. The main rainy season for most of the country is from June to September, while southern and southeastern parts of the country receive rainfall from March to May and from September to November. The southwest, which is the wettest part of the country, receives rain for almost 10 months of the year. The relationship between rainfall and elevation is not straightforward. There are parts of the country where rainfall increases with elevation, but rainfall could decrease with elevation over some other areas (Dinku *et al.* 2008a).

2.2. Data

2.2.1. *Rain gauge data*

The rain gauge data used here are part of the proprietary archives of the Ethiopian NMA. Over 500 stations (Figure 1) are used to calibrate the TAMSAT rainfall estimation algorithm, generate gridded rain gauge time series, and combine these with the satellite-based rainfall estimates. The use of such large number of rain gauges was possible because the work was done at NMA with NMA personnel. The IRI and the TAMSAT group at the University of Reading provided onsite training and technical support.

The data set used in this article covers the period from 1983 to 2010, but the data available at NMA are continually updated to the latest dekad (10-day period). Although the number of gauges used is relatively high, Figure 1 shows that the distribution of the stations is not uniform. For example, many stations are located along main roads. A positive aspect of the station distribution over Ethiopia is that most of the stations are located over the highland regions where the spatial variability of rainfall is greatest. Station distribution over the lowland areas is extremely sparse with no stations over the border
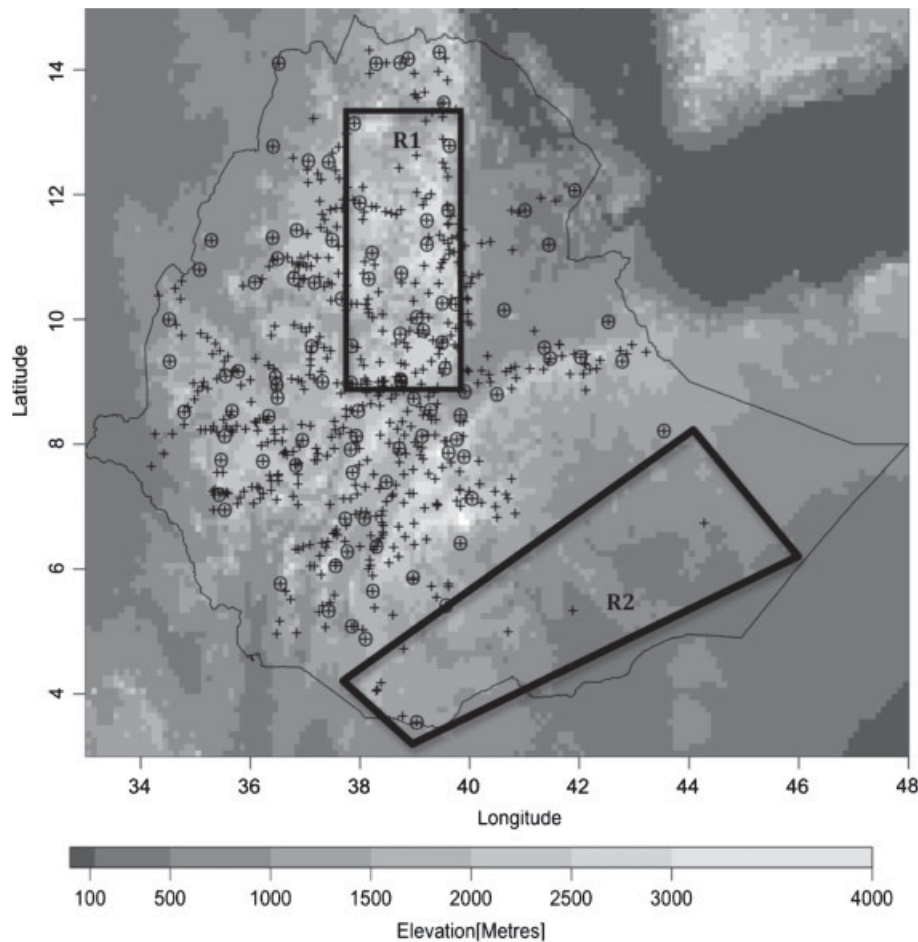
Figure 1. Study area and distribution of rain gauge stations. The circled stations are the operational stations that report on daily basis. The background is elevation. The two rectangular boxes are selected for evaluation of the different products over relatively data-rich (R1) and data-poor (R2) parts of the country.

regions. The temporal coverage of the rain gauge data is not uniform either (Figure 2). There were relatively few stations in the early 1980s, followed by a steady increase starting in the late 1980s. The number of reporting stations decreased again during 1991 and 1992. This is a result of disruptions in observations during government transition.

2.2.2. Satellite data

The aim of the project is to generate a homogeneous time series of dekadal rainfall data at a spatial resolution of 10 km going back in time as far as possible. There are many satellite estimates available with reasonable accuracy and good temporal and spatial resolutions. The TRMM Multi-satellite Precipitation Analysis (TMPA, Huffman *et al* 2007), the Climate Prediction Center (CPC) morphed product (CMORPH, Joyce *et al.*, 2004) and the Global Satellite Mapping of Precipitation (GSMAP, Okamoto *et al.*, 2007) are some examples. However, these products do not go back far enough for climatological applications and do not satisfy the homogeneity criteria as they use different sources of TIR and PMW data over the years. There are satellite products that go back over 30 years. These include the Global

Precipitation Climatology Project (GPCP, Adler *et al.*, 2003), and the CPC merged analysis (CMAP, Xie and Arkin, 1997). These products also suffer from inhomogeneity in addition to coarse spatial ($2.5° \times 2.5°$ latitude–longitude) and temporal (monthly) resolutions. To overcome these problems, a satellite rainfall time series has been generated from the Meteosat geostationary satellite TIR data. The time step of the TIR imagery is 30 min for the earlier satellites, decreasing to 15 min for Meteosat Second Generation and thus takes into account the strong diurnal cycle of rainfall over Africa. Raw METEOSAT data were obtained from the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). This task was carried out by the TAMSAT group over all Africa and the study here uses a subset of the TIR data over Ethiopia and TAMSAT software packages for calibration of the rainfall estimation and deriving the rainfall estimates.

Missing image slots is a major challenge, particularly during the early 1980s (Figure 3). This may not mean that the data do not exist, but only that the available data within a dekad were insufficient to calculate dekadal totals. Dekadal totals were not computed if gaps in the TIR archive exceeded 6 h for any day in a dekad.
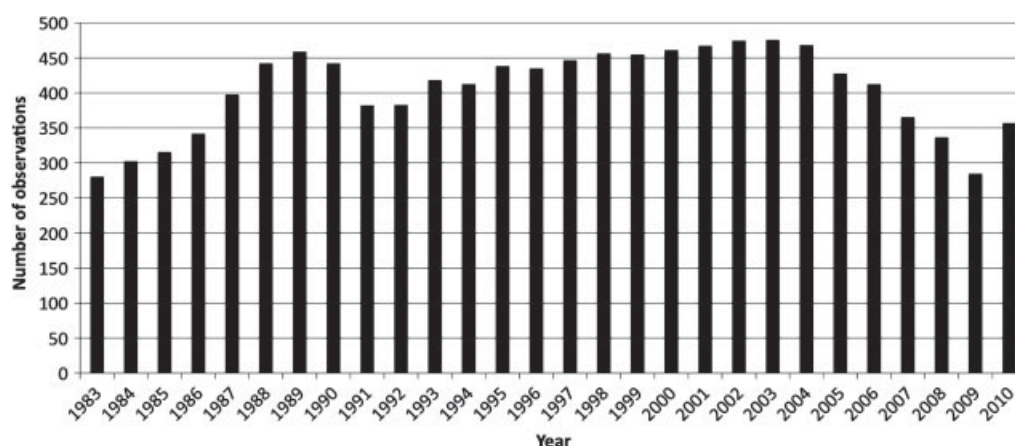
Figure 2. Changes in the number of available rain gauges stations over the years (1983–2010). The number of stations shown here represents only stations used for this study; which uses stations with at least 10-years of data.
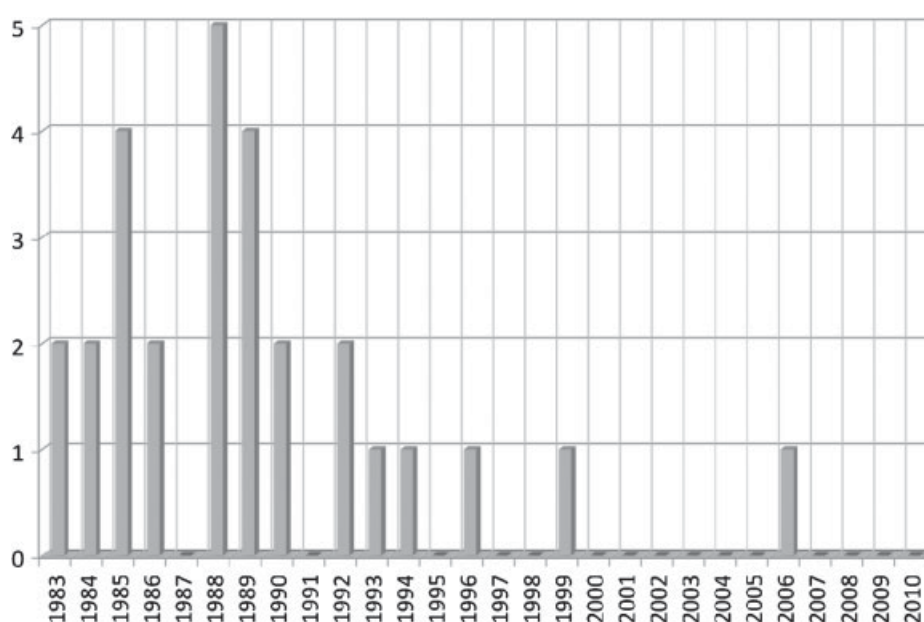

Figure 3. Number of missing dekads for the satellite data. Dekadal totals are not computed if gaps in the thermal infrared archive exceeded 6 h for any day in a dekad.

In these cases, gaps have been filled by interpolating from the closest available dekads. The distances between the missing dekad and the two closest dekads have been used to calculate weights. The interpolated values may not be realistic when the available dekads are far from the missing dekad, however, adjustment with rain gauge observations helps to overcome part of this problem.

3. Methodology

The overall process involves three major tasks such as (1) quality control of gauge data, (2) calibration of the TAMSAT approach and generation of dekadal satellite rainfall estimates, and (3) merging satellite rainfall estimates and rain gauge data. This section provides a brief description of these tasks.

3.1. Data quality control

The Ethiopian NMA performs routine quality checks on data in their archives. However, additional quality checks were performed, including checking station coordinates and identifying suspicious rainfall values. The most challenging and time-consuming task was ensuring the accuracy of station coordinates. The NMA has recently updated its station coordinates using GPS. However, the reporting of GPS readings has not been consistent across regions in terms of format (i.e. both degree-minute and decimal degree formats were used). Some the GPS measurements had to be retaken.

Spatial and temporal checks were carried out on the dekadal rainfall values. The spatial quality check procedure compares values of a given station with corresponding values of the nearby stations (Dinku *et al.*, 2008b). The temporal check compares the consistency
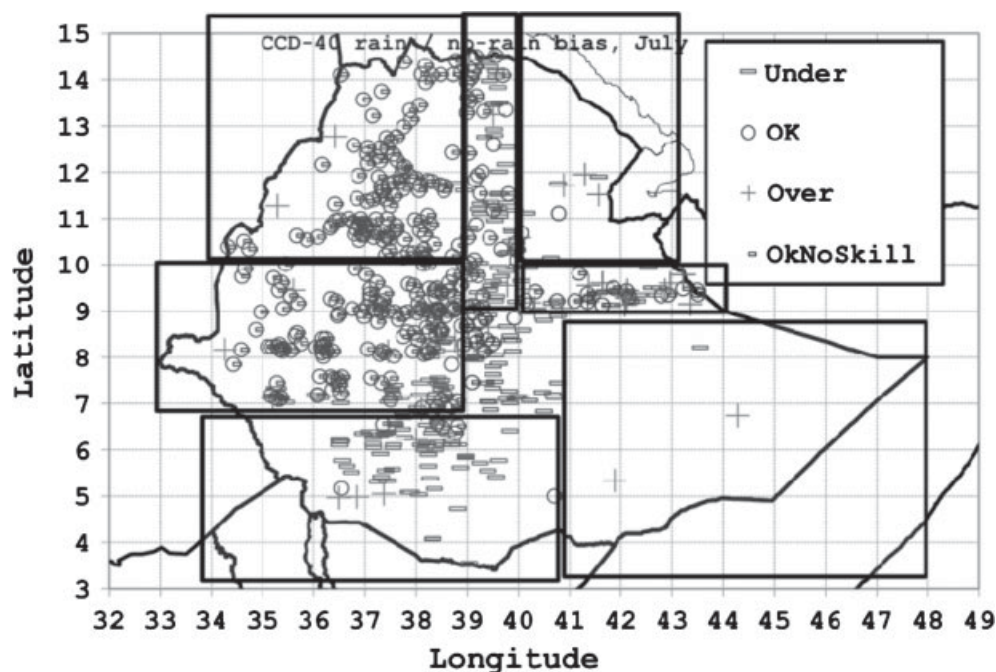
Figure 4. An example of TAMSAT calibration zones for the month of July. The different boxes represent different calibration zones, which are delineated based on comparison between estimated and observed rainfall values at each station location. This figure shows the result for CCD threshold temperature of $-40°$ C. Threshold temperature is adjusted within each zone to optimize agreement with gauge data.

of a given dekadal value with values for the same dekad of other years at the same station (Eischeid *et al.*, 1995). The temporal quality control used here checks just for extremely high values; thus, zero values were not included. Values that were identified as outliers according to the criteria set for the spatial and temporal checks were flagged as suspect. Only about 1% of the data points were flagged as suspect. NMA personnel examined suspect records to determine if these values should be discarded. Here, a choice was made to rely on the experience of the NMA experts rather than discarding data automatically. Values that were deemed suspicious were then recorded as missing data in the quality-controlled gauge archive.

3.2. Generation of satellite rainfall estimates

The TAMSAT method (Grimes *et al.*, 1999; Thorne *et al.*, 2001) was selected for the retrieval of satellite rainfall estimates because it is simple to implement and has been shown to be as good as the more sophisticated methods in Ethiopia (Dinku *et al*, 2007). The TAMSAT algorithm uses the Cold Cloud Duration (CCD) as an intermediate product to estimate rainfall from TIR data. The CCD value for a given satellite pixel is defined as the length of time the pixel is colder than a specified threshold temperature T_c (Grimes *et al.*, 1999; Thorne *et al.*, 2001). The algorithm was calibrated over Ethiopia using over 600 stations. The calibration period was from 1994 to 2005. Calibration of the algorithm is carried out as a two-stage process. The optimum T_c is found in the first stage, while the second stage involves fitting a linear regression model

to CCD and rainfall values. Temperature thresholds and calibration parameters are known to vary in space and time. To account for this, separate calibrations are carried out for each calendar month. Then, homogeneous calibration regions are determined within each calendar month on the basis of groupings of gauges with similar behaviour (an example for July is presented in Figure 4). These homogenous calibration zones reflect the effect of the complex topography on CCD–rainfall relationships. Delineation of the different calibration zones was possible because of the availability of observations from the large number of rain gauges referred to earlier. From here on, these locally calibrated rainfall estimates over Ethiopia are referred to as ETHTAM.

Rainfall estimates were initially produced at a spatial resolution of $0.0375°$ latitude–longitude. These have been resampled at $0.1°$ latitude–longitude grid. These resampled rainfall estimates will be treated as point values for the remaining part of the analysis. However, it is to be expected that the areal mismatch of gauge and satellite data will contribute to apparent satellite underestimation of high rainfall amounts. To some extent, this deficiency is overcome by the bias removal procedure described in the next section.

3.3. Merging satellite rainfall estimates and rain gauge data

Different approaches have been tested to merge satellite-based rainfall estimates and rain gauge observations. These include simple bias adjustment and regression kriging (RK) (Odeh *et al.*, 1995; Hengl *et al.*, 2004; Hengl *et al.*, 2007). Here, we test both bias adjustment

Table 1. Error statistics comparing simple bias adjustment (ADJ) with regression kriging (RK).

	CC	Eff	Bias	MAE
ADJ	0.84	0.70	1.0	13.0
RK	0.86	0.73	0.98	12.1

The statistics are computed using cross-validation with operational stations.

and regression kriging. The bias adjustment involves the following steps:

i. Extract satellite rainfall estimates at rain gauge locations;
ii. Compute the difference between the satellite estimate and rain gauge values at each station location;
iii. Interpolate these differences to each grid point (same as satellite pixel centres) using inverse distance weighting; and
iv. Add the interpolated differences back to the satellite estimate.

RK models the value of a variable at a desired location as the sum of the deterministic and stochastic components (Hengl *et al.*, 2004; Hengl *et al.*, 2007). The deterministic component is obtained through linear regression on an auxiliary variable and the stochastic components are interpolated residuals. In this study, the satellite rainfall estimates are used as the auxiliary variable. The open-source R language (http://www.r-project.org) with the gstat (http://www.gstat.org) package was used to implement RK.

Table 1 and Figure 5 compare the two approaches. Table 1 presents the results from a cross-validation using about 100 stations. Figure 5 shows an example of the adjusted satellite and combined satellite-gauge using RK. There is no substantial difference between the two approaches. For this case, it seems that the distribution of the rain gauges is more important than the methodology used. The other reason why the simplest approach produces the same result as the more sophisticated algorithm might be the complex topography and associated climate. This makes fitting a single robust
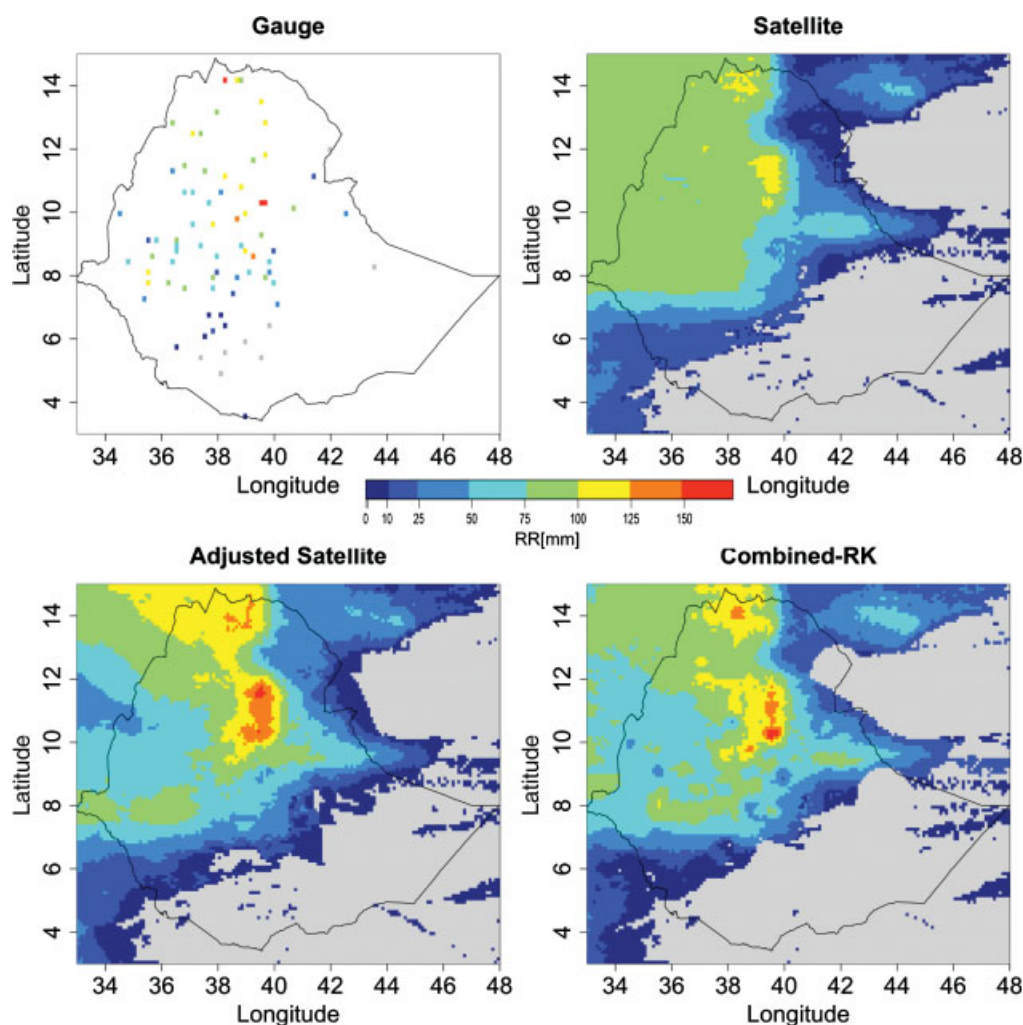

Figure 5. Comparison of bias adjustment and regression kriging for combining station measurements and satellite estimate. The four maps are station observation (top left), satellite estimate (top right), bias-adjusted satellite (bottom left), and station-satellite combined using regression kriging (bottom right).
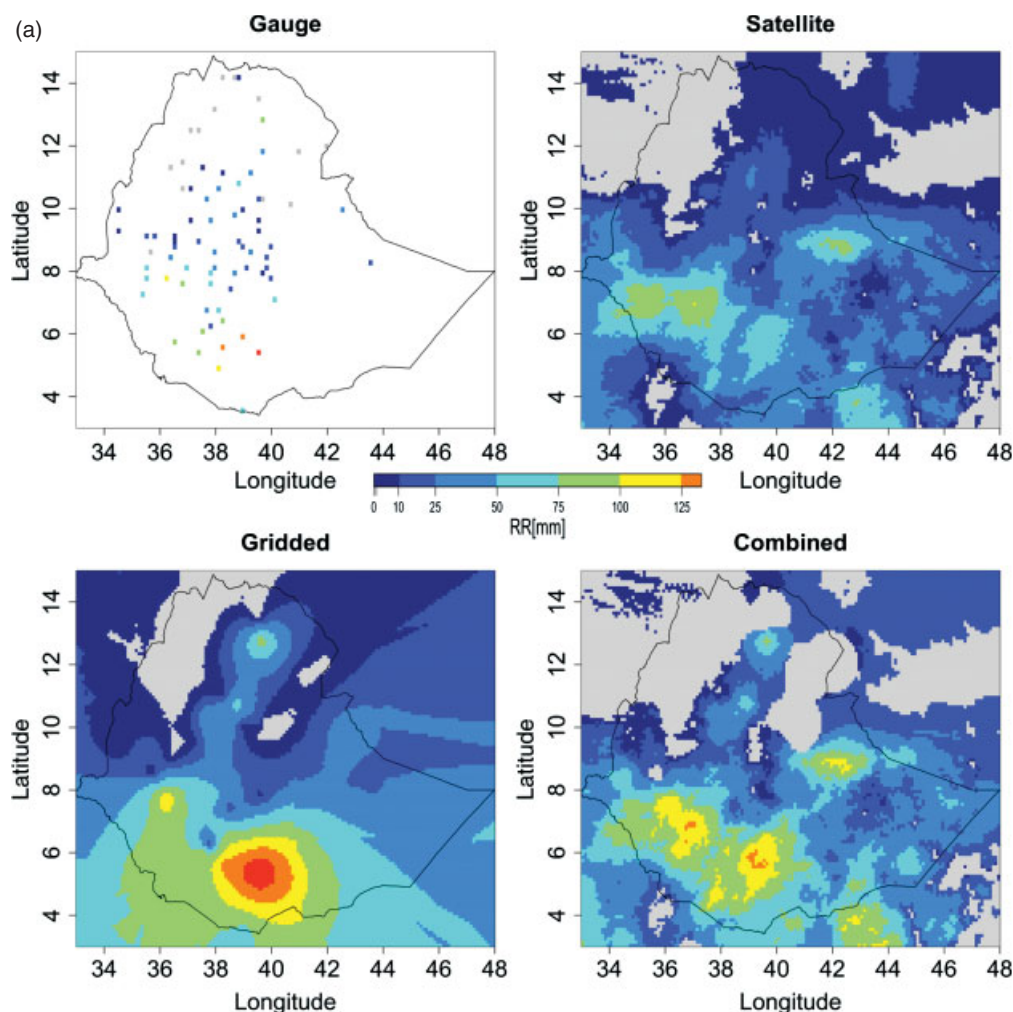
Figure 6. (a) Sample output products for the second dekad of April 1996. The top left panel is the rain gauge data for that dekad from the operational stations, while the top right panel shows the satellite estimate. The lower left panel is interpolated rain gauge values, and the lower right panel is combined product generated using satellite and rain gauge data from operational stations. (b) The same as (a), but the gauge, gridded and combined product use all available stations.

variogram that can be applied across the country a challenge. As a result, the bias adjustment approach has been adopted here because of its simplicity. The bias adjustment method implemented here uses different numbers of stations at different search radius depending on station density. This is to avoid using stations from distant locations and at the same time ensure that enough numbers of stations are used. Up to a maximum of seven stations were used if found within 100 km of the satellite pixel. Between 100 and 200 km, the maximum number of stations used was reduced to five. If less than three stations were found within 200 km, the radius was increased to 300 km. If a minimum of three stations were not found within 300 km, the average correlation distance, then no adjustments were made to the satellite data at that pixel position. These distances and number of stations are set after experimenting with different radius and number of stations.

Merging has been done for three data sets separately. The first data set uses all available stations, while the second uses those with at least 60% of time series complete. The first should produce best product at a given time, while the second for generating a more temporally consistent time series for climate analyses. The third data set is limited to 'operational' stations. There are about 80 operational stations (Figure 1), which report on daily basis. Data from other stations arrive at NMA with a delay ranging from a couple of months to up to a year. The NMA uses the operational stations for monitoring activities, daily and dekadal weather reports, and preparation of dekadal agrometeorological bulletins. Thus, merging these stations with satellite estimates is for improving NMA's monitoring capabilities.

4. Products and their evaluations

4.1. Description of products

Three products are generated: locally calibrated satellite-based rainfall estimates (ETHTAM), gridded rain gauge – only product, and the combined satellite-gauge rainfall product. Both the satellite and the combined products are available from 1983 to current. Figure 6(a)
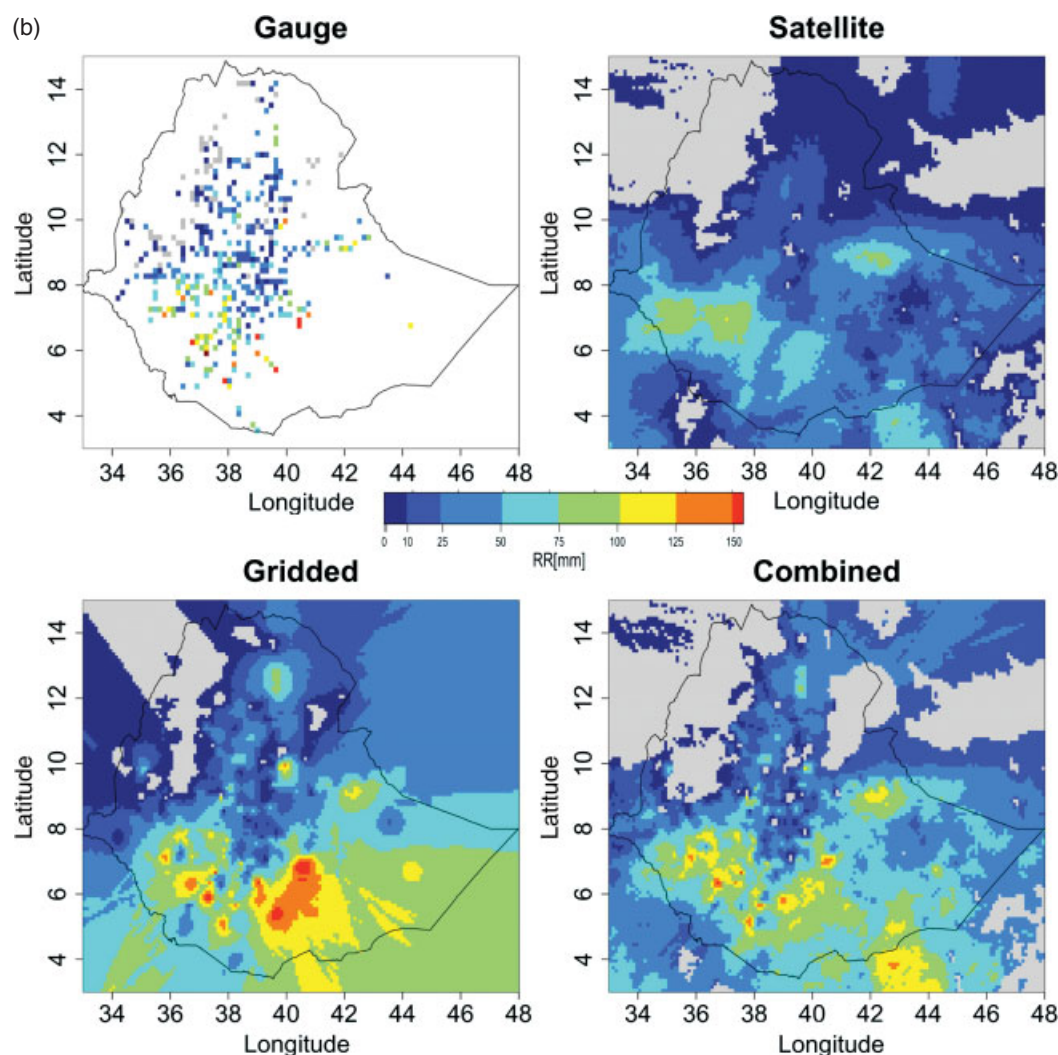
Figure 6. (*Continued*)

and (b) presents the examples of these products for the second dekad of April 1996. The top left panel represents station measurements for that dekad. The top right, lower left and lower right panels are satellite estimate, gridded rain gauge data, and combined gauge-satellite, respectively. Only operational stations are used in Figure 6(a), while all available stations are used in Figure 6(b). The rain gauge data shows the overall rainfall pattern. The main weakness of the rain gauge map in both figures is the lack of stations over the lowlands, particularly over the southeastern part of the country. The satellite product depicts the overall spatial structure of the rainfall field reasonably well. However, the satellite product underestimates rainfall amounts over most parts of the country. Underestimation is more severe over areas of high rainfall amounts. Underestimation of high rainfall values is a known weakness of TIR-based retrieval algorithms (Kidd and Huffman, 2011; Kidd and Levizzani, 2011). However, part of this discrepancy could also be due to the mismatch between point gauge data and pixel-average satellite estimates. The gridded gauge field also shows the overall

spatial structure of rainfall as shown by the gauge data; however, it also exhibits significant smoothing. This is somewhat expected because of the averaging involved. The other problem with the gridded data is the values over lowland areas. Gridded values over these areas are estimated from observations over the nearby highlands where rainfall values are much higher. This has resulted in unrealistically high values over the lowlands (Figure 6(a)). The interpolated rain gauge product in Figure 6(a) also misses the high rainfall area over southwestern part of the country, which is seen in the satellite estimate. This is because there are no rain gauges over that area. These problems overcome to some extent in Figure 6(b), which uses more stations (about 80 *vs* over 500 stations). There is less smoothing in Figure 6(b), but the problem over the lowlands still persists. The combined product overcomes, to some degree, both the lack of stations over the lowlands and the underestimation by the satellite product. The merged product combines the spatial information from the satellite estimates with the point measurements at gauge locations. The combined and gridded products are similar over station-rich parts
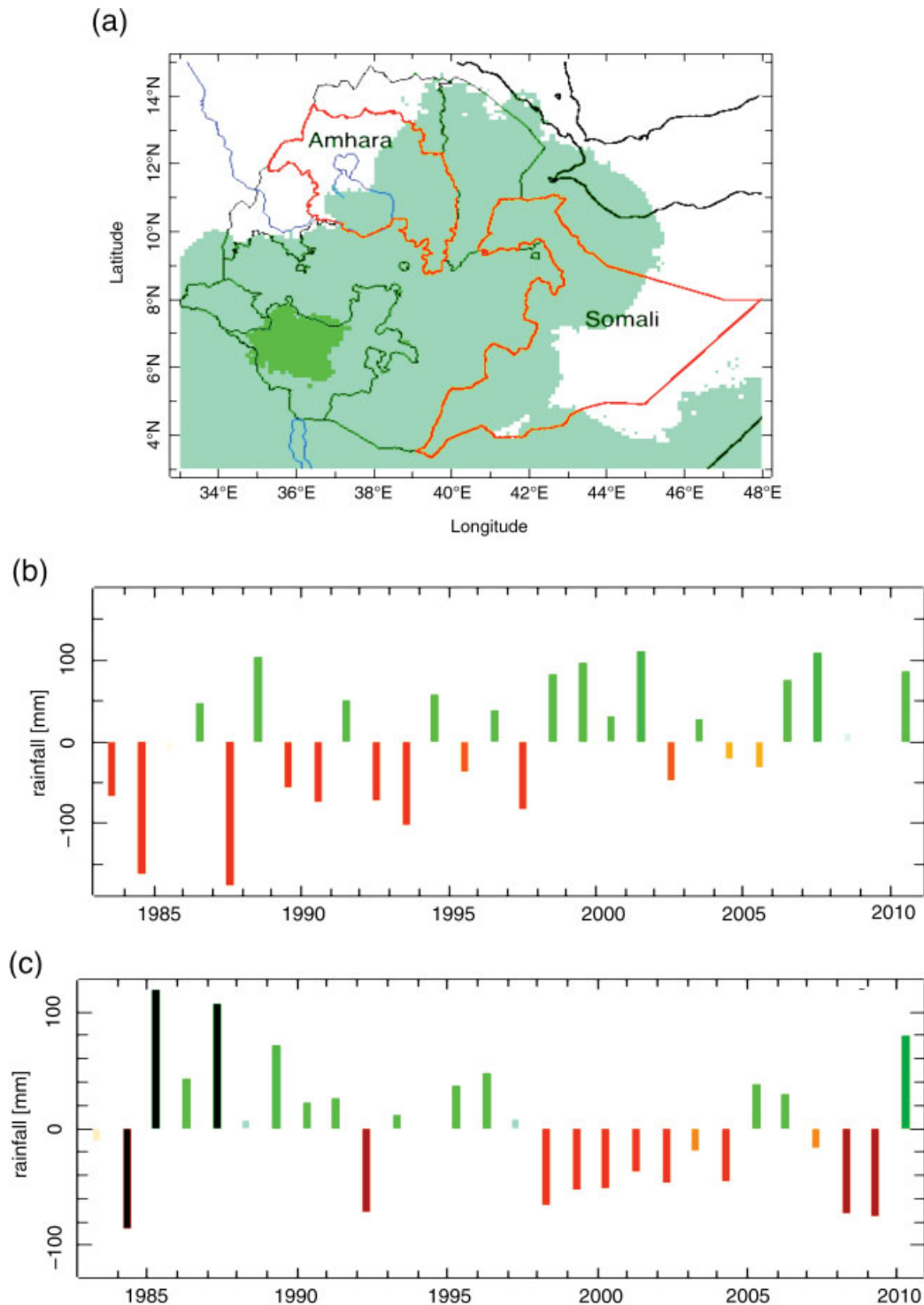
Figure 7. Time series of seasonal rainfall anomalies from 1983 to 2010 extracted for the Amhara and Somali regions of Ethiopia (a). The time series over Amhara (b) is for the June–July–August season while the one over the Somali region (c) is for the March–April–May season. The background of the map (a) is monthly mean rainfall for the month of January.

of the country. Thus, the main advantage of the merged product is over data-sparse parts of the country.

The increase in the number of stations from about 80 (Figure 6(a)) to over 500 (Figure 6(b)) has improved both the gridded and combined products. The combined product (Figure 6(a)) has spatial structure that looks more like the gauge data. However, the improvement in quality is not proportional to the number of stations. The

spatial structure of the merged product with 80 stations is close to that with 500 stations. This shows that there is an optimal number of stations, which would be much less than 500 but also more than 80, to produce a reasonably good merged product. The spatial distribution of the gauges would be as important as the number of stations.

Figure 7 shows the use of the merged data set in NMA's climate analysis maproom. Figure 7(a) is the

mean rainfall field for the month of January, while Figure 7(b) and c are seasonal rainfall anomalies extracted for the Amhara and Somali regions of Ethiopia (Figure 7(a)), respectively. The first time series (Figure 7(b)) is for the June–July–August season, which is the main rainy season for the Amhara region. The second time series (Figure 7(c)) is for March–April–May, which is the main rainy season for the Somali region. The two figures also represent the general trends of the two seasons over the country. The most noticeable features in Figure 7 are the drought of the 1980s (Figure 7(b)) and the recent decline in rainfall over the Somali region (Figure 7(c)). The recent decline depicted in Figure 7(c) is similar to what has been observed over the Horn of Africa (Lyon and DeWitt, 2012). It is worth noting that the distribution of stations over the Amhara region is significantly better than over the Somali region. However, the merged data enables similar analysis over both regions. All the plots in Figure 7 are obtained from NMA's virtual 'map room' at http://www.ethiomet.gov.et/.

4.2. Evaluation of the products

Figure 6(a) and (b) offers a qualitative comparison of the different rainfall products. Quantitative evaluations are presented in this section. The satellite rainfall estimate is compared to other similar satellite products. The gridded and combined products are evaluated using cross-validation, in which gridded and merged products are compared with corresponding gauge measurement.

4.2.1. Evaluation of satellite rainfall estimate

The satellite rainfall estimates generated and used for the merging were compared with three widely used satellite products. Two of these products are generated by the CPC at the National Oceanic and Atmospheric Administration (NOAA): the CPC rainfall estimate (CPC-RFE, Herman *et al.* (1997); Xie *et al.* (2002)) and the CPC morphing algorithm (CMORPH, Joyce *et al.* (2004)). The third product, from the National Aeronautics and Space Administration (NASA), is the TRMM (Tropical Rainfall Measuring Mission) Multi-satellite Precipitation Analysis (TMPA) (Huffman *et al.*, 2007). Version 7 of the TMPA-3B42 product was used here. All the three satellite products combine PMW and TIR data in different ways. The TMPA and CPC-RFE algorithms also employ some gauge adjustments. The reference data used for the comparison is the gridded rain gauge product generated using all rain gauges. The CMORPH and TMPA-3B42 products are available at a spatial resolution of $0.25°$ latitude–longitude. The other products are thus averaged to the same resolution. To ensure the quality of the reference data, only gridded pixels containing at least one rain gauge for that specific dekad were used for the validation of the satellite products. The TAMSAT algorithm was calibrated using rain gauge data from 1995 to 2005. Then these calibration parameters are used to generate rainfall estimates going back and also into the future. Thus, data from 2006 to 2008 were used for validation.

Table 2. Error statistics comparing satellite rainfall estimate calibrated over Ethiopia (ETHTAM) with some other widely used satellite products at spatial resolution of $0.25°$ and dekadal time scale.

	CC	Eff	Bias	MAE
ETHTAM	0.84	0.62	0.70	14.6
CMORPH	0.80	0.61	0.84	15.0
CPC-RFE	0.75	0.54	0.83	16.4
TMPA-3B42	0.83	0.66	1.02	14.2

The results are presented in Table 2 and Figure 8. Table 2 compares correlation coefficients (CC), efficiency (Eff), bias, and mean absolute error (MAE). The efficiency statistic, also known as coefficient of efficiency (Nash and Sutcliffe, 1970; Legates and McCabe, 1999), measures the skill of the satellite estimates relative to a reference value (in this case the mean of the rain gauge data). A negative value means that climatological values are better than the satellite estimates. A zero Eff value indicates that climatology is as good as the satellite estimates, and positive values show good skill for the satellite estimates.

The ETHTAM shows better performance than CPC-RFE in all the statistics presented in Table 2 except that it has higher bias than CPC-RFE. The performance of ETHTAM is similar to that of CMORPH and TMPA-3B42 except that TMPA-3B43 has very small bias. It should be noted that the ETHTAM product evaluated here is not bias-adjusted. The only advantage of this product is that the algorithm was calibrated over the study region. Figure 8 also shows the good performance of ETHTAM except for underestimation of high rainfall values. These results are consistent with those reported by Dinku *et al.* (2007). The good performance of ETHTAM is very significant considering that it is a TIR-only product. It is interesting to note that it is as good as or better than CMORPH and TMPA-3B42, which use what is considered to be state of the art operational algorithms with PMW inputs.

4.2.2. Cross-validation of satellite, gridded, and combined products

Cross-validation involves removing a station at a time and then using the neighbouring stations to estimate the value at the location of the removed station. In this case, estimation means either interpolating at station location or merging satellite pixel at station location with measurements from the neighbouring stations. In case of the satellite product, the pixel within which the station is located is taken. Cross-validation was done using four different data sets. The first set consists of all available stations while the second set consists of just 50 stations covering the whole country (Figure 9). The third and fourth data sets are from the two selected regions shown in Figure. 1. The 50-station subset (Subset50) was used to compare merging with all available stations *versus* merging using just about 10% of the available stations. These 50 stations were selected semi-randomly.
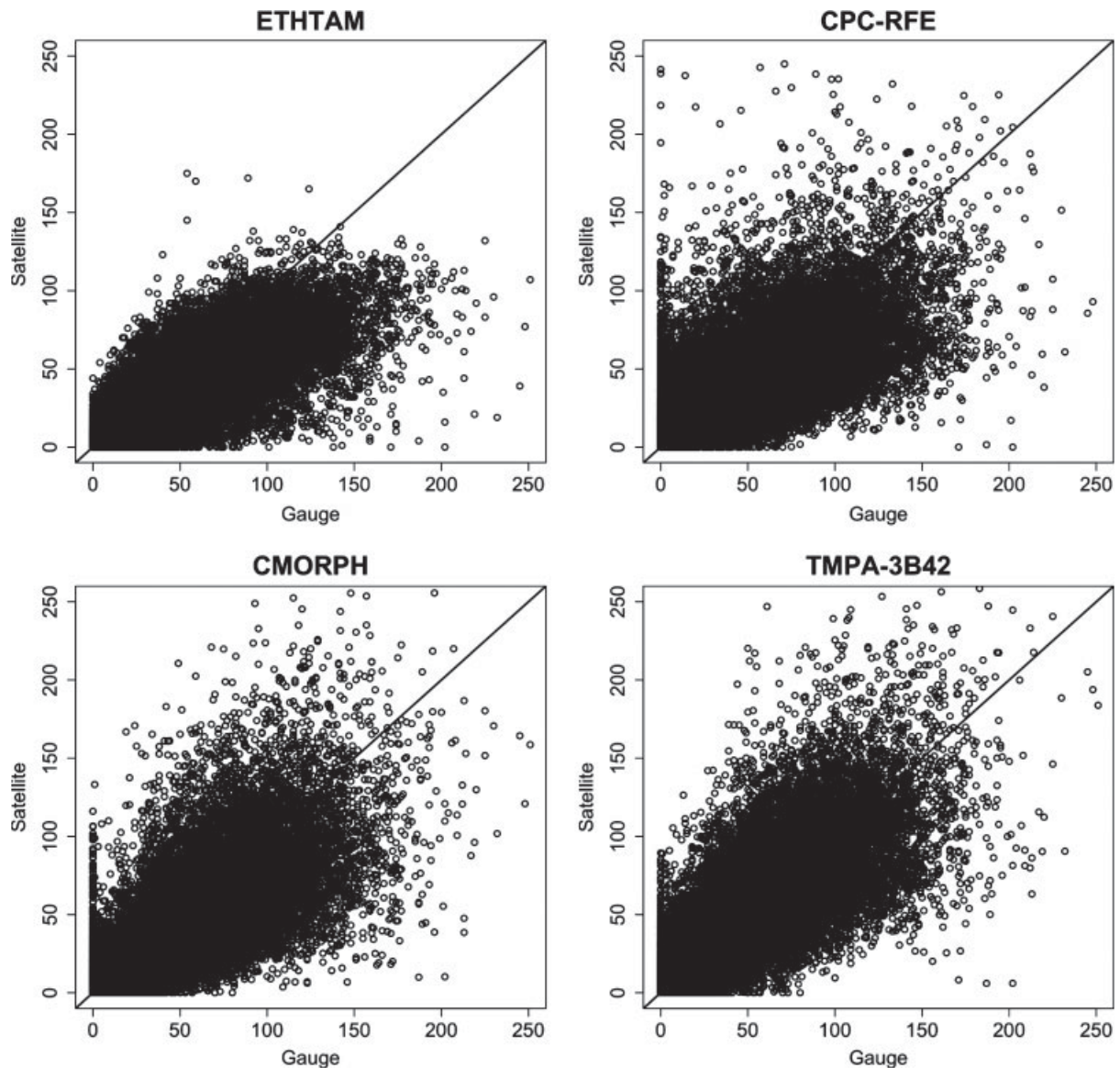
Figure 8. Comparison of the current satellite rainfall estimate (ETHTAM) with other widely used satellite products at spatial resolution of 0.25°
and dekadal aggregation.

A station was first picked randomly, then it is checked if it has at least 15 years of data and is at least 75 km away from all other stations in the same set. The 75-km is a distance at which it is possible to retain about 50 stations. The third and fourth data sets are used to compare merging over different parts of the country with different climate and different number of stations. The first region is located over northern highlands and has a good station network (Figure 1). The second region covers the parts of southern and southeastern lowlands where station distribution is close to non-existent. The results are presented in Table 3 and Figure 10. Table 3 presents the error statistics used earlier for comparing the different satellite products for, Region 1 (R1), Region 2 (R2), Subset50, and all stations. Figure 10 compares the scatter plots of the three products over the two regions. The merged product shows better performance in all cases, while the gauge-only (gridded) product is better than the satellite estimate except over R2. The satellite estimate is better than the gridded gauge over R2 because of the sparse distribution of rain gauges over this region. The main weakness of the satellite estimate is severe underestimation (bias values of 0.59–0.70). The merged product removes these biases effectively. The performance of the merged product is very similar to gridded gauge over R1 because of the relatively dense station network. The merged product is much better than the gridded product over R2. The gridded product overestimates most of the low rainfall amounts while underestimating high values. The overestimation of low rainfall amounts balances the underestimation of the high values. That is why the bias values of 1.6 in Table 3 seem to contradict the scatter plot in Figure 10. However, the performance of the satellite estimate and the merged product is somewhat similar over R2 except that the merged product overcomes the severe
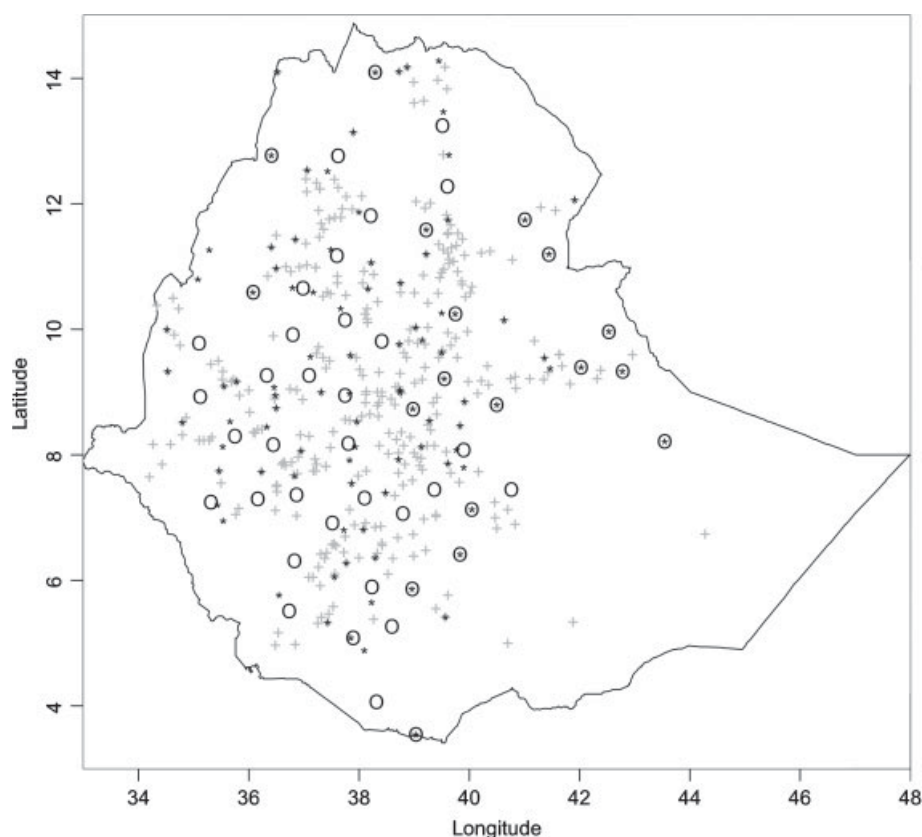
Figure 9. Distribution of the 50-station sample used for validation set (O) operational stations, (*) and the rest of the stations (+).

underestimation by the satellite. This is because the contribution of the satellite estimate is relatively higher over this region. The merged product with the 50 stations selected from across the country performs better than both the satellite and gauge-only products, particularly in terms of skill (Eff) and random error (MAE). Comparing merged products with all stations with gridded gauge, the merged products show better skill and less random error; however, the difference is not very significant. This is because the results are biased towards areas with good station coverage. These comparisons show that the merged product is close to gridded gauge over areas with good station distribution and closer to the satellite estimate over areas with sparse station network. However, the merged product effectively removes the underestimation by the satellite product in all cases. This is the main purpose of the bias adjustment method employed here.

It should be noted here that the above evaluation involves point-to-point and point-to-area average (in case of the satellite) comparisons. As a result, random errors are introduced from uncertainties in rain gauge locations and values, uncertainty in the location of the satellite pixels, and difference between the location of the reference gauges and grid/pixels centres, as well as differences between point and area-average values. These factors affect the results in Table 3 and it is also reflected in the scatter plots of Figure 10. This may not be the best validation approach, but it is the only

Table 3. Validation statistics for the different products computed for point values using cross-validation.

		CC	Eff	Bias	MAE
R1	Satellite	0.77	0.48	0.59	16.8
	Gridded	0.85	0.71	0.99	13.6
	Combined	0.88	0.76	0.98	12.2
R2	Satellite	0.72	0.47	0.69	8.2
	Gridded	0.63	0.23	1.6	13.8
	Combined	0.78	0.58	1.2	7.7
Subset50	Satellite	0.78	0.54	0.69	15.7
	Gridded	0.72	0.51	1.03	18.9
	Combined	0.80	0.63	1.01	15.2
All	Satellite	0.75	0.49	0.70	17.7
	Gridded	0.78	0.59	1.01	16.5
	Combined	0.84	0.70	1.0	14.3

R1 and R2 are parts of the country shown in Figure 1, while Subset50 represents randomly selected 50 validation stations shown in Figure 9.

approach to compare the derived products with rain gauge measurements. However, these comparisons do show the relative performance of the different products.

The other point about the above comparisons is that it may not show the difference between the gridded and merged products very well. Cross-validation could only be done at station locations. However, the advantage of the merged product is over areas with few or no stations. However, it is not possible to show that using cross-validation as removing a station would also affect the quality of the merged product particularly over
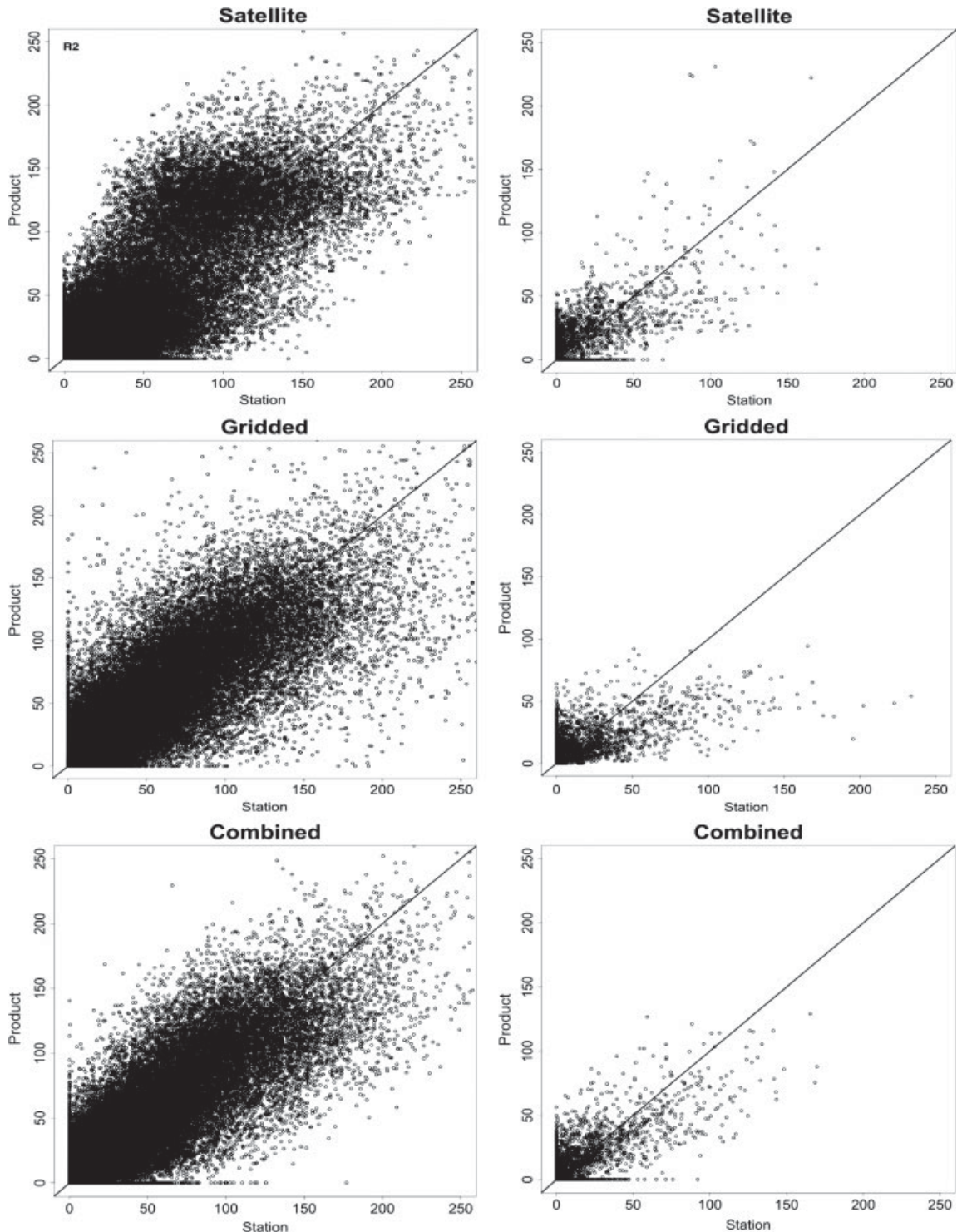
Figure 10. Comparison of the different rainfall products using cross-validation with all available stations. The comparisons were done for point values and two selected parts of the country (R1 and R2 in Figure 1). Panels on the left are for R1 while those on the right are for R2.

areas with sparse station density. Thus, the qualitative comparisons of Figure 6(a) and (b) may show a more realistic difference between the merged and gridded products.

Figure 11 compares the time series of area-averaged gauge data to that of satellite and the combined products. Cross-validation was performed using only the operational stations. Then the point values at gauge locations

were averaged over the whole country. The objective here is to show an overall improvement from the satellite only product to satellite combined with measurements from the operational stations. The selected time series (1991–1995) includes the period where there was a decline in the number of stations during 1991 and 1992. Figure 11 shows the overall underestimation by the satellite product, particularly during the peaks of the rainy
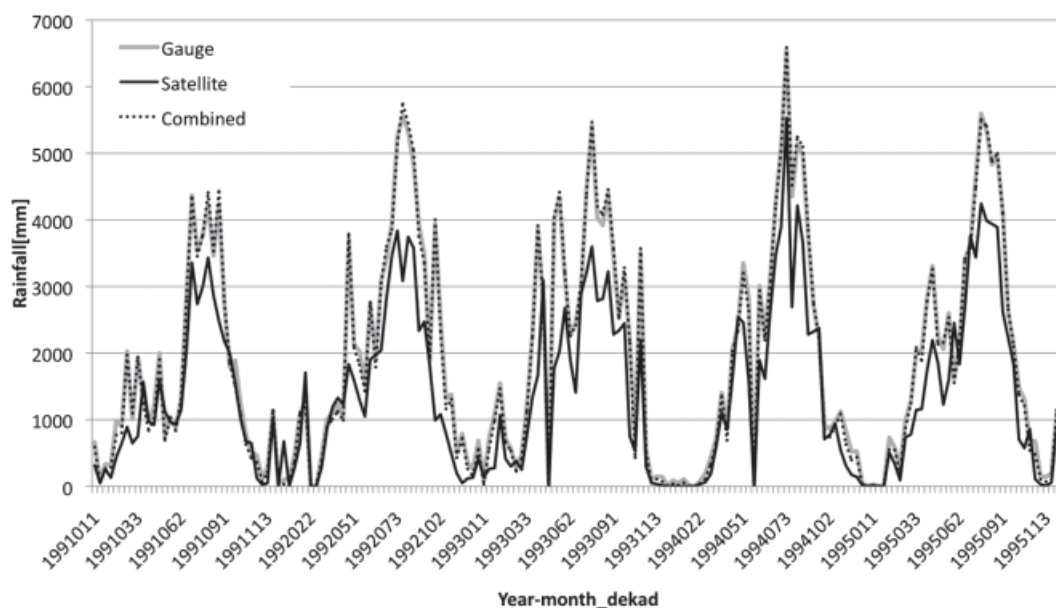
Figure 11. Time series comparing the area averages of the operational stations (about 80 of them) with the satellite and combined product.

seasons. It also shows that the combined product does remove these biases.

4.2.3. Seasonal effects

Seasonal effects have also been investigated. The performance of the satellite and combined products during the two main rainy seasons (March to May and June to August) were evaluated using cross-validations with the operational stations. This is again a point comparison; i.e. no spatial averaging is involved. The results are presented in Table 4 and Figure 12. The values in Table 4 are very similar for both seasons except for bias and MAE values. The March−April−May (MAM) season has more bias (more severe underestimation) while the June−July−August (JJA) season exhibits higher random errors (higher MAE values). The higher random error for the JJA season is also reflected in the scatter plots of Figure 12. Figure 12 also shows that the performance of the satellite product is better during the MAM season. This is mainly due to the fact that rainfall during the MAM season is dominated by convective systems. Both convective and stratiform systems contribute to the rainfall during the JJA season. Table 4 and Figure 12 also show that the combined product created using about 80 operational stations performs reasonably well.

Table 4. Error statistics comparing satellite and combined products for the March−April−May (MAM) and June−July−August (JJA) seasons.

	Satellite				Combined			
	Corr	Eff	Bias	MAE	CC	Eff	Bias	MAE
MAM	0.73	0.46	0.70	14.2	0.80	0.64	0.96	12.4
JJA	0.72	0.46	0.80	23.7	0.79	0.62	1.01	20.5

These are results of cross-validation using about 80 operational stations.

5. Summary

Efforts are underway to improve the availability of climate data over Africa by combining station observations with satellite and other proxies. An example has been provided where historical time series of rainfall data have been generated over Ethiopia by combining satellite rainfall estimates and available rain gauge data. Meteosat TIR data were obtained by the TAMSAT group and calibrated over Ethiopia to generate dekadal (ten-daily) satellite-based rainfall time series. Although this satellite rainfall product was found to underestimate high rainfall values, it has been shown that this product performs better than or as good as other satellite products that use algorithms that are more sophisticated and include PMW inputs. This better performance has been attributed to the use of national rain gauge data for calibrating the retrieval algorithm. The satellite rainfall estimates were then combined with station measurements. The combined product using RK has been compared to simple bias removal. There was no substantial difference between the two approaches. As a result, simple bias adjustment was used. The merged product was shown to be significantly better than satellite estimate. There is no significant difference between gridded gauge and the merged product over regions with relatively dense station network. However, the merged product performs much better over data-sparse parts of the country. This is a desired result as the main objective of this work has been to improve data availability over regions with few or no meteorological observations. The advantage of the gridded data is that it is not limited by availability of satellite data; thus, it could be used to generate a time series over a much longer period. The disadvantage is that the gridded data time series is temporally inhomogeneous in that the gauges used vary from year to year.
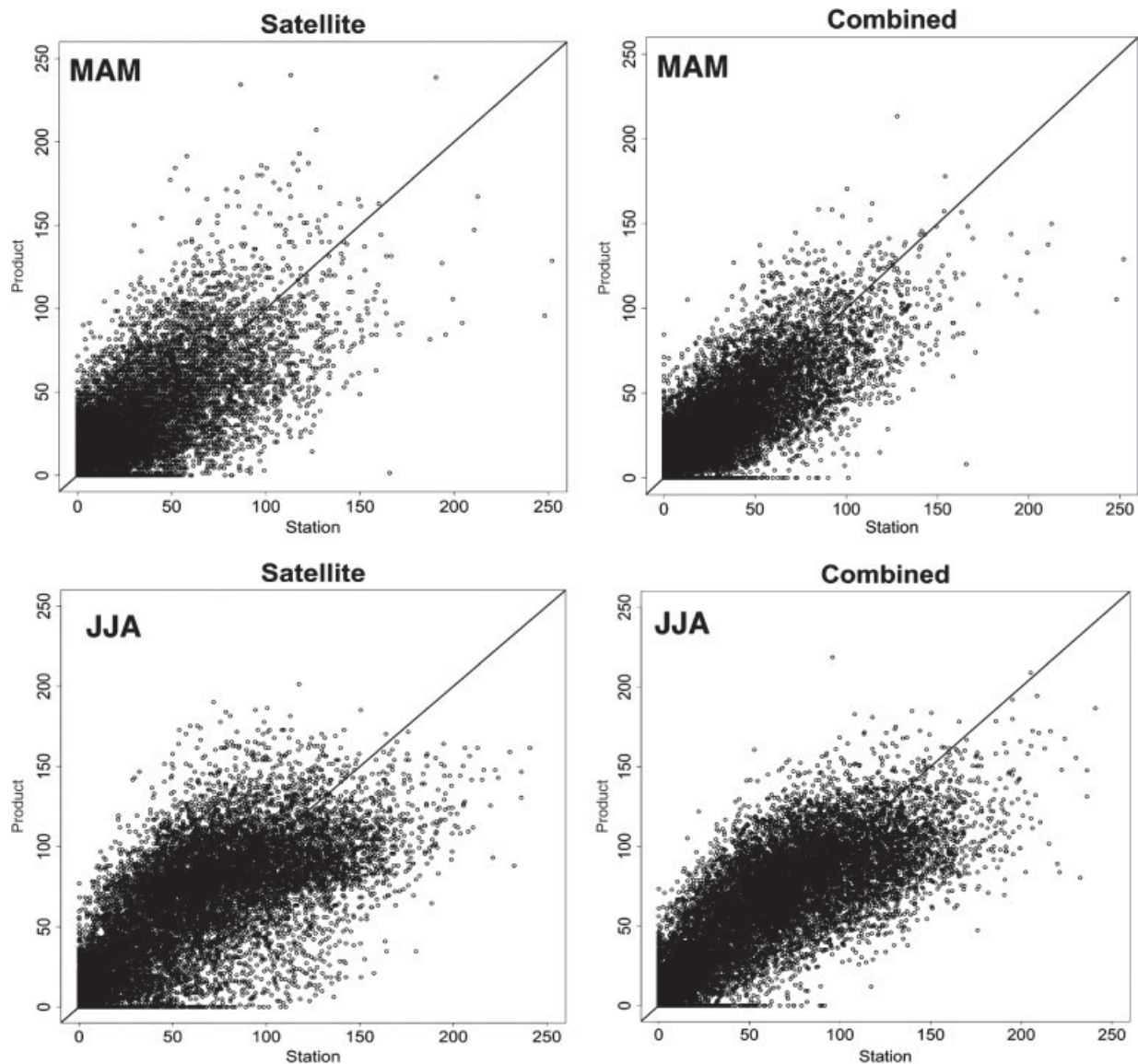
Figure 12. Comparison of the different rainfall products for the main rainy seasons – March to May (MAM) and June to August (JJA).

The current work needs to be extended to daily rainfall data. Extending this to a daily time scale would be more challenging; however, it is very important as many applications, such as hydrological and agricultural modelling, need daily rainfall inputs. The Ethiopian experience may serve as a template to improve the availability of climate data in other countries in Africa. The raw satellite data obtained and processed for this work covers the whole of Africa. Methodologies and computer codes developed for current work could easily be adopted for another country. Thus, expanding this work to other countries in Africa would be faster and cheaper. Similar project has already been completed in Tanzania and Madagascar in collaboration with the National Meteorological Agencies. It is also being implemented in West Africa, at a regional level, in collaboration with the AGRHYMET Centre. Planning is also underway to implement the project in some other countries in Africa.

References

Adler RF, Huffman GH, Chang A, Ferraro F, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P, Nelkin E. 2003. The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **4**: 1147–1167.

Dinku T, Ceccato P, Grover-Kopec E, Lemma M, Connor SJ, Ropelewski CF. 2007. Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Rem. Sensing* **28**: 1503–1526.

Dinku T, Chidzambwa S, Ceccato P, Connor SJ, Ropelewski CF. 2008a. Validation of high-resolution satellite rainfall products over complex terrain in Africa. *Int. J. Rem. Sensing* **29**: 4097–4110.

Dinku T, Connor SJ, Ceccato P, Ropelewski CF. 2008b. Intercomparison of global gridded rainfall products over complex terrain in Africa. *Int. J. Climatol.* **28**: 1627–1638.

Dinku T, Asefa K, Hailemariam K, Connor SJ. 2011. Improving Availability, Access and Use of Climate Information. *WMO Bull.* **60**(2): 80–86.

Eischeid JK, Baker CB, Karl TR, Diaz HF. 1995. The quality control of long-term climatological data using objective data analysis. *J. Appl. Meteorol.* **34**: 2787–2795.

Grimes DIF, Pardo-Iguzquiza E, Bonifacio R. 1999. Optimal areal rainfall estimation using raingauges and satellite data. *J. Hydrol.* **222**: 93–108.

Hengl T, Heuvelink GMB, Stein A. 2004. A generic framework for spatial prediction of soil variables based on regression kriging. *Geoderma* **122**: 75–93.

Hengl T, Heuvelink GBM, Rossiter DG. 2007. About regression-kriging: From equations to case studies. *Comput. Geosci.* **33**: 1301–1315.

Herman A, Kumar VB, Arkin PA, Kousky JV. 1997. Objectively determined 10-day African rainfall estimates created for famine early warning. *Int. J. Rem. Sensing* **18**: 2147–2159.

Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolf DB. 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8**: 38–55.

IRI, GCOS, DfID, ECA. 2006. A gap analysis for the implementation of the global climate observing system program in Africa. International Research Institute for Climate and Society, Technical report Number IRI-TR/06/1, 52pp. http://iri.columbia.edu/outreach/publication/report/06-01/report06-01.pdf.

Joyce RJ, Janowiak JE, Arkin PA, Xie P. 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* **5**: 487–503.

Kidd C, Huffman G. 2011. Review - global precipitation measurement. *Meteorol. Appl.* **18**: 334–353, DOI: 10.1002/met.284.

Kidd C, Levizzani V. 2011. Status of satellite precipitation retrievals. *Hydrol. Earth Sys. Sci. Dis.* **7**: 8157–8177, DOI: 10.5194/hessd-8-8157-2010.

Legates DR, McCabe GJ. 1999. Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **35**: 233–241.

Lyon B, DeWitt D. 2012. A recent and abrupt decline in the East African long rains. *Geophys. Res. Lett.* **39**: L02702, DOI: 10.1029/2011GL050337.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models, I, A discussion of principles. *J. Hydrol.* **10**: 282–290.

Okamoto K, Iguchi T, Takahashi N, Ushio T, Awaka J, Shige S, Kubota T. 2007. High precision and high-resolution global precipitation map from satellite data. *Proceedings, ISAP* **2007**: 506–509.

Odeh I, McBratney A, Chittleborough D. 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma* **67**: 215–226.

Sperling F (ed). 2003. *Poverty and climate change-reducing the vulnerability of the poor through adaptation*. World Bank: Washington, DC; 43.

Thorne V, Coakeley P, Grimes D, Dugdale G. 2001. Comparison of TAMSAT and CPC rainfall estimates with raingauges, for southern Africa. *Int. J. Rem. Sensing* **22**: 1951–1974.

Washington R, Harrison M, Conway D, Black E, Challinor A, Grimes D, Jones R, Morse A, Kay G, Todd M. 2006. Africa climate change: taking the short route. *Bull. Am. Meteorol. Soc.* **87**: 1355–1366.

Xie P, Arkin PA. 1997. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Am. Meteorol. Soc.* **78**: 2539–2558.

Xie P, Yarosh Y, Love T, Janowiak J, Arkin PA. 2002. A real-time daily precipitation analysis over southern Asia. Preprints, *16th Conf. of the Hydro., Orlando, FL, Am. Meteor. Soc*.