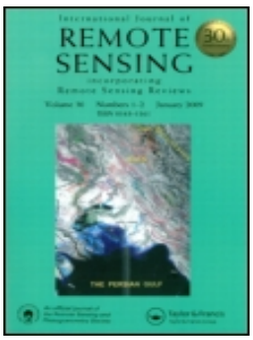

International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

Multitemporal comparative analysis of TRMM-3B42 satellite-estimated rainfall with surface gauge data at basin scales: daily, decadal and monthly evaluations

Yashon O. Ouma , Titus Owiti , Emmanuel Kipkorir , Joel Kibiiy & Ryutaro Tateishi

To cite this article: Yashon O. Ouma , Titus Owiti , Emmanuel Kipkorir , Joel Kibiiy & Ryutaro Tateishi (2012) Multitemporal comparative analysis of TRMM-3B42 satellite-estimated rainfall with surface gauge data at basin scales: daily, decadal and monthly evaluations, International Journal of Remote Sensing, 33:24, 7662-7684, DOI: 10.1080/01431161.2012.701347

To link to this article: https://doi.org/10.1080/01431161.2012.701347

Published online: 04 Jul 2012.

Submit your article to this journal

Article views: 574

View related articles

Citing articles: 32 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tres20

International Journal of Remote Sensing
Vol. 33, No. 24, 20 December 2012, 7662–7684

Taylor & Francis
Taylor & Francis Group

Multitemporal comparative analysis of TRMM-3B42 satellite-estimated rainfall with surface gauge data at basin scales: daily, decadal and monthly evaluations

YASHON O. OUMA*†‡, TITUS OWITI†, EMMANUEL KIPKORIR§, JOEL KIBIIY† and RYUTARO TATEISHI‡

†Department of Civil and Structural Engineering, Moi University, Eldoret, Kenya
‡Centre for Environmental Remote Sensing (CEReS), Chiba University, Chiba, Japan
§Environmental Planning and Management Division, SES, Moi University, Eldoret, Kenya

(*Received 12 April 2011; in final form 28 February 2012*)

In order to examine the reliability and applicability of Tropical Rainfall Measuring Mission (TRMM) and Other Satellites Precipitation Product (3B42) Version 6 (TRMM-3B42) at basin scales, satellite rainfall estimates were compared with geostatistically interpolated reference data from 12 rain gauge stations for three consecutive years: 2005, 2006 and 2007. Gauge–TRMM-3B42 statistical properties for daily, decadal and monthly multitemporal precipitations were compared using the following cross-validation continuous statistical measures: mean bias error (MBE), root mean square difference (RMSD), mean absolute difference (MAD) and coefficient of determination (r^2) metrics. The averaged spatial–temporal comparisons showed that the TRMM-3B42 rainfall estimates were much closer to the geostatistically interpolated gauge data, with minimal biases of -0.40 mm day^{-1}, -1.78 mm decad^{-1} and -6.72 mm month^{-1} being observed in 2006. In the same year, the gauge and TRMM-3B42 rainfall estimates marginally correlated better than in 2005 and 2007, with the daily, decadal and monthly coefficients of determination being 82.2%, 93.9% and 96.5%, respectively. The results showed that the correlations between the gauge-derived precipitation and the TRMM-3B42-derived precipitation increased with increasing temporal intervals for all three considered years. Quantitatively, the TRMM-3B42 observations slightly overestimated the precipitations during the wet seasons and underestimated the observed rainfall during the dry seasons. The results of the study show that the estimates from TRMM-3B42 precipitation retrievals can effectively be applied in the interpolation of missing gauge data, and in the verification of precipitation uncertainties at the basin scales with minor adjustments, depending on the timescales considered.

1. Introduction

Information on precipitation rates, amounts and distribution is indispensable for a wide range of applications, including agronomy, hydrology, meteorology and climatology. A thorough understanding of the spatial and temporal distribution of

*Corresponding author. Email: yashon_o@hotmail.com

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2012 Taylor & Francis
http://www.tandfonline.com
http://dx.doi.org/10.1080/01431161.2012.701347

precipitation is critical to water resource management, where the sizes of basins considered range from 100 to 100 000 km^2 and the temporal integration of rainfall inputs ranges from hours to days. It therefore follows that sustainable decisions regarding water resource management require accurate information on water availability, which comprises continuous quantification of the spatial and temporal distribution of rainfall at different functional hydrologic scales such as river basins.

In most countries, rain gauges are the main source of rainfall measurements. However, the reliability of rain gauge data is limited by their lack of adequate spatial coverage. This is further complicated by areas with varied climatology and topographical characteristics. Furthermore, rain gauge networks are limited to over the land only and remain sparse over most of the globe, while radar networks are limited to only a few countries (Kummerow *et al.* 1998, Pardo-Iguzqiza 1998, Hong *et al.* 2006).

In poorly gauged regions, rainfall availability and analysis are hampered by the fact that only minimal hydrometeorological information is available. As such, only discrete rain and stream gauge data are available; however, these may be unsuitable for answering specific questions in water resource demands such as in operational irrigation water supply, running hydropower stations and/or flood forecasting for disaster management (Li *et al.* 2009). This implies that other data sources for deriving hydrometeorological information, such as satellite-based precipitation retrievals, are significant for land surface hydrological research and applications, especially for regional or watershed-scale studies. (Tian *et al.* 2007, Li *et al.* 2009, Wagner *et al.* 2009).

Recent advances in the field of remote sensing have led to an increase in available rainfall data at regional and global scales (Adler *et al.* 2003, Hong *et al.* 2004). Such precipitation data sources include, among others: the Tropical Rainfall Measuring Mission (TRMM), Earth Observing System (EOS) Aqua satellite, National Oceanic and Atmospheric Administration (NOAA)/Climate Prediction Center morphing technique (CMORPH), and Global Satellite Mapping of Precipitation (GSMaP), which provide the hydrologic community with an abundance of new precipitation data. While satellite data have the advantage of regional coverage as compared with rain gauge data, they are also subject to uncertainties due to the indirect nature of measurements (Greene and Morrissey 2000, Hong *et al.* 2004). This implies that the reliability of satellite-based precipitation estimates needs to be evaluated and inter-compared for a specific study area and temporal interval.

Satellite-estimated rainfall uncertainties arise from various factors such as cloud-top reflectance, thermal radiance, retrieval algorithm and infrequent satellite overpasses (Petty and Krajewski 1996, Hossain *et al.* 2006). The contribution of each error in the overall uncertainty depends on the retrieval algorithm and the considered spatial–temporal scale (Yan and Gebremichael 2009). Therefore, the quantification of precipitation estimation error is essential as it propagates in hydrologic processes, resulting in significant error in hydrologic predictions (Tetzlaff and Uhlenbrook 2005, Hong *et al.* 2006, Schuurmans and Bierkens 2007, Khan *et al.* 2011). Due to the potential significance of satellite-estimated rainfall data in hydrologic applications, efforts are required to assess the uncertainty of these data for localized applications (Hong *et al.* 2004, Ebert *et al.* 2007, Tian *et al.* 2007, Li *et al.* 2009).

In an effort to quantify satellite precipitation estimation errors, Hong *et al.* (2006) proposed a model in which satellite precipitation error was described as a non-linear function of rainfall space–time integration scale, rain intensity and sampling frequency. Using Monte Carlo simulations, Hong *et al.* (2006) generated an

ensemble of rainfall estimates to evaluate the influence of satellite error propagation into hydrological modelling. Hossain and Anagnostou (2004) introduced a probabilistic approach to describe the uncertainty of the passive microwave (PM)- and infrared (IR)-based satellite rainfall data for hydrologic applications. Bellerby and Sun (2005) developed a methodology to represent the uncertainty in satellite rainfall retrievals based on the covariance structure of the rainfall field and the conditional distribution functions of precipitation on the pixel scale. Most of the satellite error assessment studies are usually limited to investigations at regional space scales of the monthly or longer temporal intervals (Chiu *et al.* 2005, Li *et al.* 2009). For hydrological or agricultural applications, shorter submonthly period rain rates such as daily, five-day (pentad), ten-day (decad) or monthly rain rates are more appropriate.

For localized accuracy and reliability assessments, satellite rainfall estimates should be calibrated and validated using rain gauge networks. While rain gauges measure rain accurately and continuously at a point, they offer little information on the rainfall data between gauges. Rain gauges themselves may not be fully accurate and may be influenced by factors such as the calibration accuracy, wind effects and sampling uncertainty, which subsequently limit the accuracy for sampling intervals smaller than 10 min (Pardo-Iguzqiza 1998, Bowman 2005, van de Beek *et al.* 2011).

Satellites and gauges, respectively, determine and measure precipitation in very different ways. The TRMM instruments, for example, make remote, volume-averaged measurements of hydrometeors in the atmosphere, from which the area-averaged surface rain rate over the catchment is inferred. By scanning across the orbit track, TRMM provides a snapshot of the rain rate over an extended region. The satellite measurements are area-averaged instantaneous observations with relatively poor time sampling as compared with the gauges, which provide good time sampling but poor spatial distribution in terms of area coverage (Bowman 2005).

Despite their differences, comparing gauge with satellite estimates remains an essential tool in the validation of satellite products. Some disagreement between averages of satellite data and rain-gauge data is, however, expected because of the very different sampling patterns of the two systems – the satellite provides only occasional snapshots of large areas, whereas rain gauges provide continuous measurements over very small areas. A quantitative estimate of the expected level of disagreement due to the differences in sampling is needed in order for comparisons of the two kinds of averages to be informative.

Measurements of precipitation from space can complement ground-based observations towards the provision of a more complete picture of rain structures at any given temporal or spatial scale. Satellite precipitation measurements provide the impetus for rain algorithm development and improvement. In this study, a validation approach is proposed for the TRMM and Other Satellites (3B42) Version 6 products (TRMM-3B42) against the ground-based rain gauge measurements within the Nzoia River Basin in Kenya at daily, decadal and monthly temporal scales, for three concurrent years: 2005, 2006 and 2007.

The proposed validation approach is designed to capture random error sources and not physical biases and vagueness of remotely sensed precipitation estimates, using the following continuous cross-validation statistical metrics: (i) mean bias error (MBE), (ii) root mean square difference (RMSD) and (iii) mean absolute difference (MAD). The coefficient of determination was used to compare the statistical fit of the TRMM-3B42 with the surface gauge data. This analytical approach is expected to predict how the spatial rainfall distributions are affected by the sampling resolution in space

(basin) and time (daily, decadal and monthly). In comparing the two data sources, there is expected random error bands associated with each of the data.

In this study, these errors are dependent on the rainfall amounts and timescales, and while they may influence pixel-based analyses, they are considered negligible over large study areas. The study is carried out within the framework of the Kenya Water Resources Management Authority's (WARMA) strategic water resources management plan for basin water information and management system (b-WIMS).

2. Study area and data

2.1 *Study area: the Nzoia River Basin*

The Nzoia River Basin is located in the western part of Kenya and lies between latitudes 0° N–1.5° N and longitudes 34° E–36° E. The basin covers a geographical area of approximately 12 709 km^2 with a topographical elevation ranging between 1300 and 3000 m above mean sea level (figure 1, adapted from Khan *et al.* (2011)). The mean annual rainfall varies from a maximum of 1100–2700 mm to a minimum of 600–1100 mm. There are two rainy seasons and two dry seasons, with the short rains running from October to December and the long rains running from March to May. The dry seasons occur in the months of January–February and June–September. Due to the variable local relief and proximity to Lake Victoria, the climatic patterns within the basin are not usually regular. Characteristically, the low-altitude areas of the basin are prone to flooding. Figure 1(*b*) shows the mean monthly rainfall for three selected stations (Chorlim, Kakamega and Uholo), representing the upper, middle plateau and lower elevation regions of the basin.

The basin is traversed by the River Nzoia, which drains into Lake Victoria and the Nile River Basin. The Nzoia River Basin is densely populated and agriculturally productive with rapidly growing commercial and industrial activities in Eldoret, Kitale, Kakamega, Mumias and Nzoia towns.

2.2 *Data*

The satellite and gauge rainfall data sets were considered for three years: 2005, 2006 and 2007. The year 2006 was used as a benchmark because of the availability of continuous and reliable (calibrated) gauge data. The years 2005 and 2007 were chosen since they represented data before and after the reference year. Twelve collocated rain gauge stations (figure 1(*a*)) were used as reference data. The identifications and locations of the 12 stations are presented in table 1. The 12 stations are the World Meteorological Organization (WMO) coded and recognized stations within the basin.

2.2.1 ***In situ*** **rainfall measurements: surface rain gauge data.** Typically, rainfall data collected from rain gauge stations are likely to contain errors due to missing data entries, data recording and data formatting. In rainfall data analysis, incompleteness reduces the length and information content of the rainfall record. Consistency check analysis was carried out on the gauge observed data to determine any missing gaps in the recorded rainfall data using the double-mass curve (DMC) method. Details on the DMC concept and application are presented in §3.

To compare the continuous satellite data and the discrete gauge data, spatial interpolation as a means of creating surface data from sample points is mandatory. The data used to perform the spatial interpolation of the gauge data were mean monthly

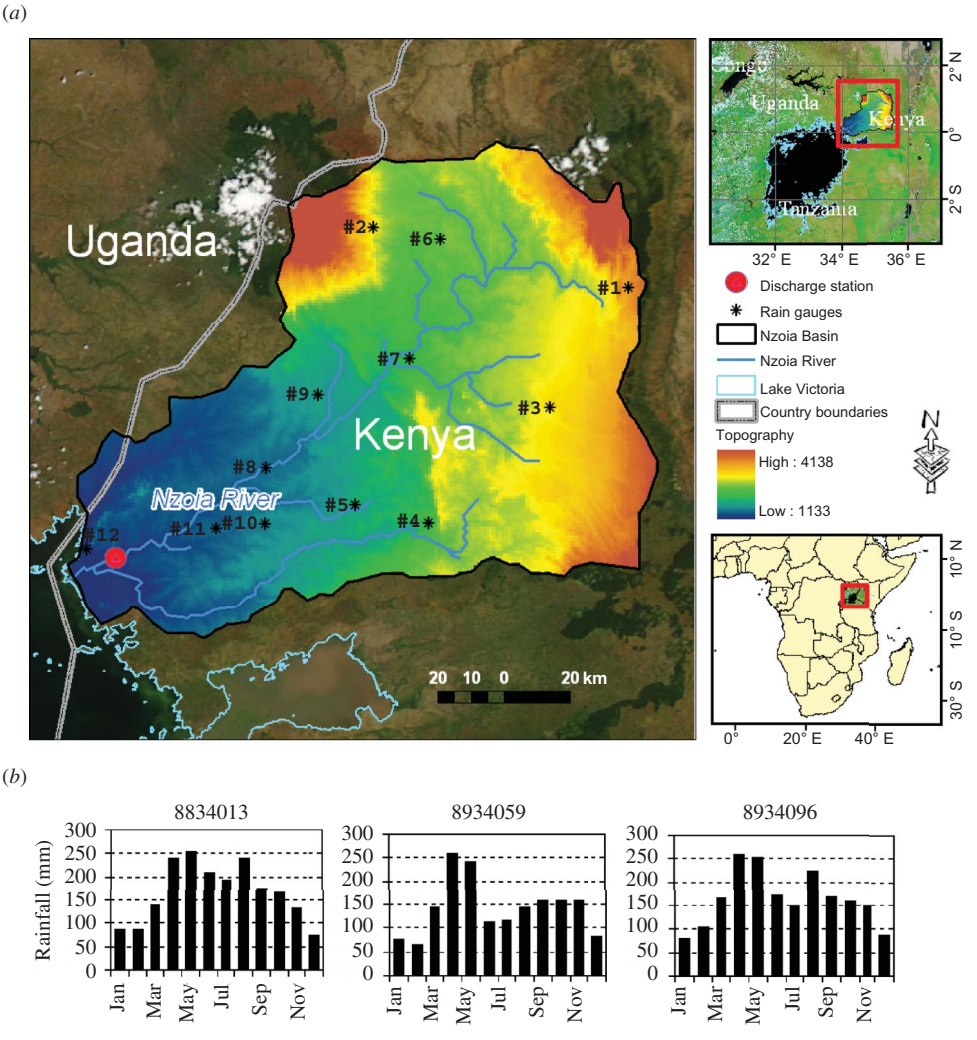

Figure 1. (*a*) Location map of Nzoia River Basin showing the distribution and locations of the 12 rain gauge stations within the watershed (adapted from Khan *et al.* (2011)). (*b*) Three years' (2005, 2006 and 2007) mean monthly rainfall observations for selected (Chorlim, Kakamega and Uholo) rainfall stations. The three stations, respectively, represent the upper, middle plateau and lower elevation regions of the basin.

rainfall measurements from 12 stations in the Nzoia River Basin over a 3 year period, 2005–2007. This implies that there were 36 gauge station samples (12 monthly samples for each of the three years).

2.2.2 Sensor rainfall data – TRMM-3B42. TRMM estimates tropical precipitation from space sensors using a suite of rain retrieval algorithms. Products from the TRMM multisatellite precipitation analysis algorithm include the 'TRMM and Other Satellites' (3B42) and the 'TRMM and Other Sources' (3B43). The 3B42 estimates are produced 3-hourly at a spatial resolution of 0.25°. The major inputs into

Table 1. Geographic location of the World Meteorological Organization (WMO) recognized rain gauge stations for areal rainfall measurements within the Nzoia River Basin.

Station #	Gauge station name	Gauge WMO code	Latitude (°N)	Longitude (°E)
1	CHEBIEMIT AGRIC. OFFICE	8935104	0.8670	35.5000
2	CHORLIM ADC FARM	8834013	1.0330	34.8000
3	ELDORET MET	8935181	0.5330	35.2830
4	KAIMOSI FTC	8934078	0.2170	34.9500
5	KAKAMEGA MET	8934096	0.2670	34.7500
6	KITALE MET	8834028	1.0000	34.9830
7	LUGARI FOREST STN	8934016	0.6670	34.9000
8	MUMIAS SUGAR	8934133	0.3670	34.5000
9	NZOIA SUGAR	8934183	0.5670	34.6500
10	BUTERE HEALTH CENTRE	8934040	0.2140	34.4990
11	PORT VICTORIA FOR STN.	8934191	0.1463	34.0113
12	UHOLO CHIEF'S CAMP	8934059	0.2023	34.3652

Note: The corresponding coordinates of the stations are depicted in figure 1.

the 3B42 algorithm are IR data from geostationary satellites and PM data from the TRMM microwave imager (TMI), special sensor microwave imager (SSM/I), Advanced Microwave Sounding Unit (AMSU) and Advanced Microwave Sounding Radiometer-Earth Observing System (AMSR-E).

The 3B42 estimates are produced in four stages: (a) the PM estimates are calibrated and combined; (b) the IR precipitation estimates are created using the PM estimates for calibration; (c) PM and IR estimates are combined; and (d) the data are rescaled to monthly totals, whereby gauge observations are also used indirectly. The TRMM-3B42 product is available for a few days after the end of each month as monthly data (http://trmm.gsfc.nasa.gov/3b42.html), in addition to the 3-hourly data. There is a near-real-time version, 3B42-real-time (3B42RT), that is available with a time lag of about 6 h. This version is a product at the third step (c) and does not include gauge information. 3B42 estimates are considered to supersede the 3B42RT estimates as each month of 3B42 is computed during the following month. The 3B42 processing is designed to maximize data quality, so 3B42 is strongly recommended for any research work not specifically focused on real-time applications. 3B42 estimates are accumulated and merged with gauge data to produce the monthly product (3B43) at 0.25° spatial resolution. The TRMM-3B42 precipitation (in mm h^{-1}) products are available with a spatial resolution of 0.25° × 0.25° grid and a temporal resolution of 3 h within 50° N –50° S global latitude (Huffman *et al.* 2005, 2007). The TRMM data were downloaded from the NASA/Goddard Space Flight Center (GSFC)/Data and Information Services Center (DISC).

3. Methods

3.1 *Consistency analysis and spatial interpolation of* in situ *gauge data to areal data*

In this study, data verification was carried out on the rain gauge data for consistency analysis before interpolation and comparison with the satellite data. The objective of consistency analysis is to verify and validate the *in situ* measured gauge data.

3.1.1 Gauge data consistency analysis. Before carrying out the areal data estimation from discrete data, it is necessary to perform gauge data consistency checks and apply adjustments if necessary. This process ensures that the trends detected are due to meteorological causes and not attributed to changes in gauge location, exposure and/or observational methods. If the anomalies are not attributed to meteorological causes, adjustments are made using the coefficients determined from the DMC method (McCuen 1989).

Adjusting for gauge consistency is an estimation problem of an effect to minimize the impact on the 'accuracy' of the rain gauge data. In the DMC adjustment method, the variable (rainfall) accumulation at a given station is plotted against the average accumulation for a group of surrounded stations that are having good (high) correlation. In this study, DMC analysis was used to check for inconsistency in the gauge record (P_r). Adjustments for inconsistencies were carried out according to the following equation:

$$P_a = \frac{\tan \alpha\prime}{\tan \alpha} P_r = \frac{\delta_a}{\delta_r} P_r, \tag{1}$$

where P_r is the recorded or observed precipitation in millimetres; P_a is the adjusted precipitation in millimetres; $\tan\alpha\ (= \delta_r)$ is the original DMC slope or slope of the graph at the time P_r is observed; and $\tan\alpha'\ (= \delta_a)$ is the slope of the deviated (inconsistent) section of the DMC or slope of the graph to which the recordings are adjusted (McCuen 1989).

3.1.2 Spatial interpolation of discrete data. Rain gauges provide an estimate of rainfall at a point. Through spatial interpolation, with a suitable geostatistical method, the discrete point data are interpolated to represent continuous surface data (Holawe and Dutter 1999, Grimes and Pardo-Iguzquira 2010). Typically, a network of rain gauges is used to determine rainfall patterns and distributions over a target area. However, a network of rain gauges can only resolve features of the rainfall surface larger than the characteristic distance between gauges in the network. The success of the interpolation process depends on how consistently rain gauge observations represent rain falling in the area of interest.

Mathematically, nearly all spatial interpolation techniques can be represented as weighted averages of sampled data. Spatial interpolators share the same general estimation formulation (equation (2)):

$$\hat{z}\,(\boldsymbol{u}_o) = \sum_{i=1}^{n} \lambda_i z\,(\boldsymbol{u}_i), \tag{2}$$

where $\hat{z}$ is the estimated value of an attribute (primary variable) at the point of interest defined by the vector $\boldsymbol{u}_0$; z is the observed value at the sampled point $\boldsymbol{u}_i$; λ_i is the weight assigned to the sampled point; and n represents the number of sampled points used for the estimation (Laslett *et al.* 1987, Webster and Oliver 2001).

The spatial interpolation techniques used in this study are those that are commonly used in hydrological sciences and geosciences and are classified as either deterministic or stochastic (geostatistical). Deterministic interpolation techniques create surfaces from measured points, based on either the extent of similarity or the degree of smoothing. These techniques do not use a model of random spatial processes. Deterministic interpolation techniques can be divided into two groups, global and local. Global

techniques calculate predictions using the entire data set. Local techniques calculate predictions from the measured points within neighbourhoods, which are smaller spatial areas within the larger study area.

Geostatistics assumes that at least some of the spatial variations of natural phenomena can be modelled by random processes with spatial autocorrelation. Stochastic or geostatistical methods incorporate the concept of randomness, whereby the interpolated surface is conceptualized as one of many that might have been observed, all of which could produce the known data points. Stochastic methods incorporate the concept of randomness and provide both estimations (deterministic) and geostatistical-associated errors (stochastic – uncertainties represented as estimated variances).

In this study, the following deterministic methods were analysed for the rain gauge data interpolation: (a) Thiessen polygons and (b) inverse distance weighting (IDW) (Legates and Willmont 1990, Stallings *et al.* 1992, Collins and Bolstad 1996). The following geostatistical techniques were analysed and compared with the deterministic methods: (a) ordinary Kriging, (b) simple Kriging and (c) universal Kriging (Krige 1951, Matheron 1970, Goovaerts 1997). Kriging methods belong to the linear least-squares estimation algorithms that are based on the Gauss–Markov theorem (Chiles and Delfiner 1999). The Kriging process is divided into two distinct tasks: quantifying the spatial structure of the data and producing a prediction. Quantifying the structure, known as variography is where a spatial dependence model is fitted to the data. To make a prediction for an unknown value for a specific location, Kriging uses the fitted model from variography, the spatial data configuration and the values of the measured sample points around the prediction location. The geostatistical analysis provides many tools to help determine which parameters to use and also provides reliable defaults that can be used to make a surface quickly.

3.2 *Inter-comparison and validation of spatial–temporal precipitation observations*

Statistical comparisons of model estimates or predictions with thought-to-be reliable and pair-wise matched observations remain among the most basic means of assessing model performance in the climatic and environmental sciences. To validate the daily, decadal and monthly TRMM-3B42 data sets against the rain gauge data, the following continuous verification statistical measures were used for the inter-data comparison: (i) MBE – for overall reliability measures, (ii) RMSD and (iii) MAD, where RMSD and MAD are considered as appropriate measures for comparison and overall accuracy (Willmott and Matsuura 2005).

The MBE, also called additive bias, indicates the average direction of the deviation from observed values. MBE, defined by equation (3), measures the average error of a number of observations found by taking the mean value of the positive and negative errors, without regard to sign. A positive bias indicates that the estimated value exceeds the observed value on average, while a negative bias corresponds to underestimation of the observed value on average. The normalized bias $(\mathrm{MBE})_{\mathrm{norm}}$is given in equation (4). Both the MBE and $\mathrm{MBE}_{\mathrm{norm}}$ indices depict the same statistical inference:

$$\mathrm{MBE} = n^{-1} \sum_i (x_i - G_i), \tag{3}$$

$$\mathrm{MBE}_{\mathrm{norm}} = G_i^{-1} \sum_i (x_i - G_i). \tag{4}$$

The RMSD (equation (5)) is a quadratic scoring rule, which measures the average magnitude of the error. Since the errors are squared before they are averaged, the RMSD gives a relatively high weight to large errors. This implies that RMSD is most useful when large errors are particularly undesirable:

$$\mathrm{RMSD} = \left[n^{-1} \sum_i (x_i - G_i)^2 \right]^{\frac{1}{2}}. \tag{5}$$

RMSD has been used as a reliable criterion for comparative studies related to spatial interpolation of rainfall such as high-resolution studies of rainfall; assessing the effect of integrating elevation data into the estimation of monthly precipitation; and comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scales and in spatial distribution of rainfall (Ly *et al.* 2011).

The MAD metric measures the average magnitude of the errors in a set of estimated values, without considering direction and quantifies the accuracy of continuous variables. MAD (equation (6)) is a linear score, implying that all of the individual differences are weighted equally in the averaging process:

$$\mathrm{MAD} = n^{-1} \sum_i |x_i - G_i|. \tag{6}$$

Compared with MAD, RMSD gives greater weight to large errors than to small errors in the average. In equations (3)–(6), n is the total number of gauge stations, $i = 1, 2, 3, \ldots, 12$; x_i is the algorithm rain rate; and G_i is the gauge measured (spatially interpolated) rainfall data.

RMSD and MAD are among the optimal measures of performance, because they summarize the mean difference in the units of observed and predicted values. Comparatively, the bias calculation assesses the average difference between satellite and gauge values, while the RMSD measures the average magnitude of the errors with a focus on extreme values.

To determine the disparity in the compared data sets, the coefficient of determination, which is the square of the correlation coefficient (r) of the best fit linear regression line was used. The correlation between the observed values and predicted values, defined by the coefficient of determination or Pearson's product-moment correlation coefficient, is a performance measurement indicator or a phase error determinant. The coefficient of determination (equation (7)) addresses the question of how well the satellite precipitation retrievals correspond to the measured ground-based gauge values, and is a measure of the degree of linear association between the estimated and observed values of rainfall estimates:

$$r^2 = \left[\frac{n \sum x_i G_i - \left(\sum x_i\right)\left(\sum G_i\right)}{\sqrt{n\left(\sum x_i^2\right) - \left(\sum x_i\right)^2}\sqrt{n\left(\sum G_i^2\right) - \left(\sum G_i\right)^2}} \right]^2 = \left[\frac{\mathrm{cov}\,(x_i, G_i)}{\sigma_{x_i}\sigma_{G_i}} \right]^2, \tag{7}$$

where σ_x and σ_G are the standard deviations of x and G, respectively, and cov (x_i, G_i) is the covariance between the satellite and gauge measurements.

The coefficient of determination (r^2) is significant as it gives the proportion of the variance (fluctuation) of one variable that is predictable from the other variable. It is

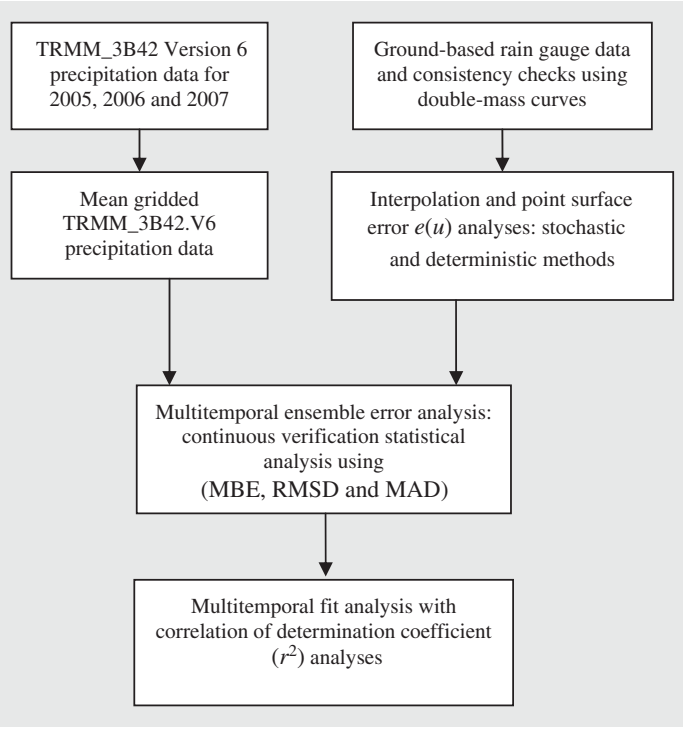

Figure 2. Schematic framework for the spatial–temporal inter-data comparison of TRMM-3B42 precipitation retrievals and rain gauge data.

the ratio of the explained variation to the total variation and measures how well the regression line represents the data. The coefficient of determination is such that $\left[0 \leq r^2 \leq 1\right]$ and denotes the strength of the linear association between satellite rain rates (x) and rain gauge data (G).

To implement the above outlined steps, a schematic conceptual approach, represented in figure 2, was adopted for the inter-comparative evaluation of TRMM precipitation data and ground-based rain gauge data for the study area.

4. Results and discussion

4.1 *Gauge data consistency results and analysis*

The results of the DMC analysis showed very high consistencies for the following stations: Mumias Sugar, Eldoret Metrological, Uhoho Chief's Camp and Nzoia Sugar. Data collected from the consistent stations were used in the determination of consistencies of the remaining network stations. Consistency adjustments based on equation (1) were applied to Chebeimit, Port Victoria, Kakamega, Kaimosi, Chorlim, Kitale, Lugari and Butere stations. The results for the 2005, 2006 and 2007 gauge measured rainfall generally showed high consistencies with minimum and maximum r^2 values of 98.7% and 99.7% being observed in 2006. Therefore, only representative results of 2006 are presented in figure 3. The results in figure 3 show that only less than 2% of the total variation in between gauge and TRMM-3B42 data remains unexplained.

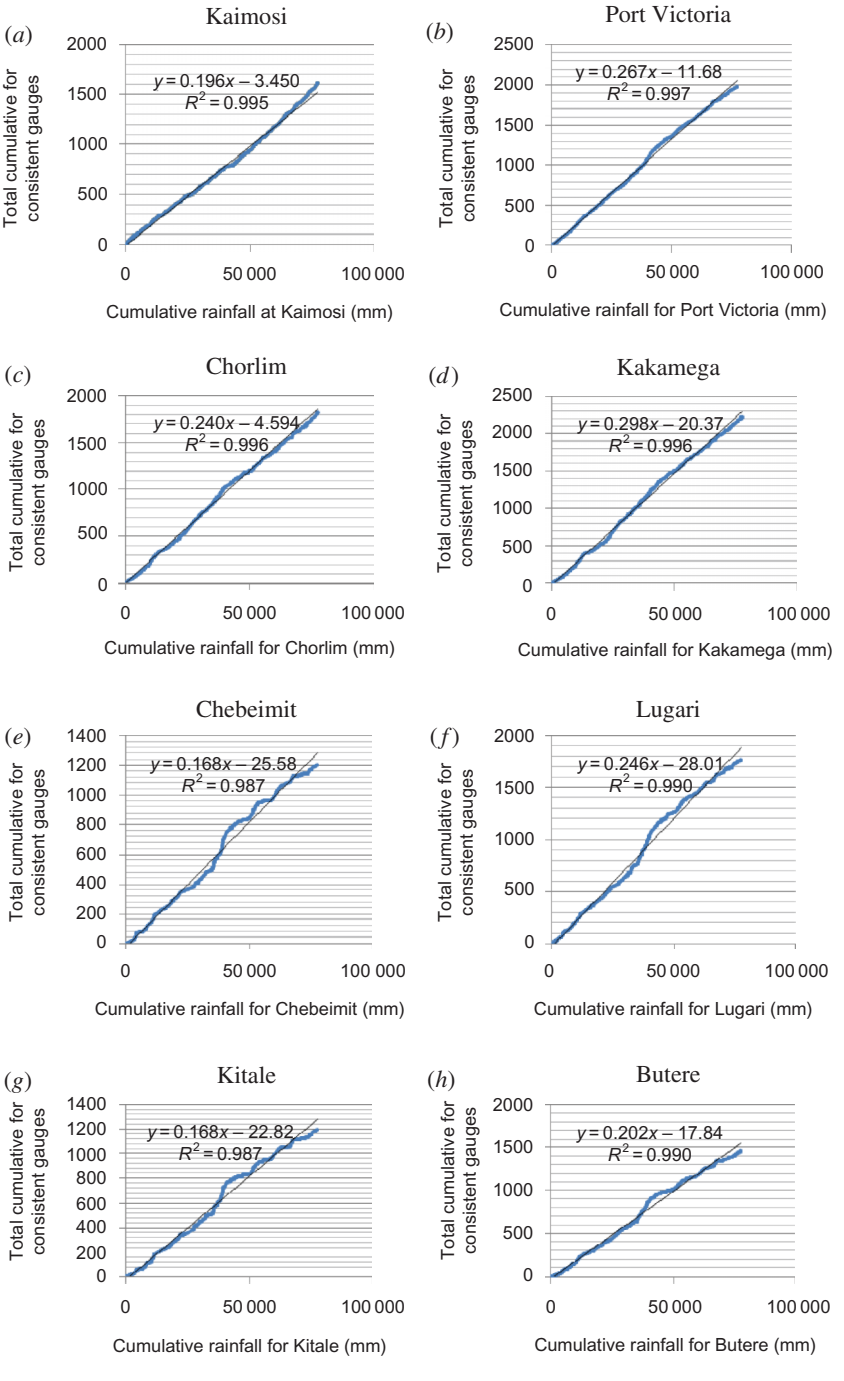

Figure 3. Double-mass curves (DMCs) for rainfall gauging stations: (*a*)–(*d*) represent DMCs for the consistent gauging stations data and (*e*)–(*h*) represent DMCs for the corrected gauging stations data. (The *x*-axis is magnified by ten for better visualization of the correlation lines.).

The results in figure 3 depict over 99% agreement for all of the stations. This implies that the observed marginal differences in the gauge data were not as a result of meteorological influences. Therefore, only very negligible adjustments were applied to the candidate eight gauge stations for the three years.

4.2 *Spatial interpolation of* in situ *gauge data*

To compare and analyse the spatially interpolated data $\left(\hat{z}\left(\boldsymbol{u}\right)\right)$ with the measured gauge data $(z\left(\boldsymbol{u}\right))$, the differences between the average monthly predicted (interpolated) rainfall and the measured values were determined as the error $(e(\boldsymbol{u}))$. $e(\boldsymbol{u})$ was used as a parametric accuracy measure, and the results are presented in table 2 for the three years: 2005, 2006 and 2007. The interpolation results in table 2 show that the magnitudes of the observed errors $(e(\boldsymbol{u})_i)$ were slightly higher in 2005 but lower in 2007. This could be attributed to the reliability of the gauge observations in 2006.

The results indicate that the deterministic Thiessen polygon method is the most inferior with a mean error of +29.3 mm in 2005, while the stochastic-based Kriging (ordinary spherical) presented a mean marginal error of +2.4 mm in 2006. In the 3-year mean monthly observations, all of the deterministic spatial interpolators overestimated the mean monthly rainfall as compared with the Kriging methods. The results showed that Kriging-based interpolation methods describe the best linear unbiased estimator in the sense of least variance. Using deterministic techniques, the weight depends solely on the distance from the measured point to the predicted location, while with the stochastic interpolation techniques, the weights are based not only on the distance between the measured point and the prediction location but also on the overall spatial arrangement of the measured points.

4.2.1 Spatial interpolation of *in situ* gauge data. Comparing Kriging methods with the deterministic interpolators, the major differences are that the former provides

Table 2. The results of compared spatial interpolation methods applied to the mean monthly rainfall recorded from the ith year, where $i = 2005, 2006, 2007$ and $\left[e\left(\boldsymbol{u}\right)_i = \hat{z}\left(\boldsymbol{u}\right)_i - z\left(\boldsymbol{u}\right)_i\right]$.

Interpolation technique	Interpolation method	$e(\boldsymbol{u})_{2005}$	$e(\boldsymbol{u})_{2006}$	$e(\boldsymbol{u})_{2007}$
Deterministic-based models	IDW-Variable search radius	+25.7	+22.2	+21.3
	IDW-Fixed search radius	+22.2	+20.6	+21.9
	Thiessen polygons	+29.3	+26.9	+27.1
Stochastic (Kriging)-based methods	Ordinary spherical	+4.1	+2.4	+3.5
	Ordinary circular	−8.9	−5.0	−6.8
	Ordinary exponential	+10.0	+9.3	+9.4
	Ordinary Gaussian	+15.7	+11.0	+3.6
	Simple spherical	−20.5	−14.2	−15.7
	Simple circular	−19.4	−21.9	−20.5
	Simple exponential	−21.6	−17.9	−19.9
	Simple Gaussian	−6.3	−3.2	−8.5
	Universal circular	−10.9	−10.1	−11.2
	Universal spherical	−14.2	−13.4	−15.0
	Universal exponential	−7.3	−8.8	−8.2
	Universal Gaussian	−11.4	−9.5	−10.1

uncertainty assessment, anisotropy detection or methodology assumptions (Negreiros *et al.* 2010). Kriging describes the best linear unbiased estimator in the sense of least variance, and is the Best Linear Unbiased Estimator (BLUE) and Best Unbiased Estimator (BUE) if data respects the bell curve (Griffith 1993, Negreiros *et al.* 2010).

In the implementation of the Kriging interpolation methods, sample variograms (inverse function of the spatial and temporal covariance) were calculated for all directions (omnidirectional variogram). The best-fitting model was identified by adjusting the nugget, range, sill and anisotropy factor. Model variogram was used to develop Kriged surface to predict spatial continuity. For the 36 samples, spherical, circular, exponential and Gaussina variogram models were generated.

It was observed that the semivariance increased according to the separation distance, explaining that two rainfall data close to each other were more similar, and hence their squared difference was less significant, than those that are farther apart. As shown in table 2, the ordinary spherical model was the most suitable for 2006 and 2007, while for 2005, the simple Gaussian yielded the best results. Comparing the magnitudes, the 2006 results were lower than both the 2005 and 2007 results.

Mathematically, Kriging-based methods can be defined as estimators of the variants of the basic linear regression estimator $z^*(\boldsymbol{u})$ as follows:

$$z^*(\boldsymbol{u}) - m(\boldsymbol{u}) = \sum_{\alpha=1}^{n(\boldsymbol{u})} \lambda_\alpha [z(\boldsymbol{u}_\alpha) - m(\boldsymbol{u}_\alpha)], \tag{8}$$

where $\boldsymbol{u}$, $\boldsymbol{u}_\alpha$ are the location vectors for the estimation point and one of the neighbouring data points, indexed by α; $n(\boldsymbol{u})$ is the number of data points in the local neighbourhood used for estimation of $z^*(\boldsymbol{u})$; $m(\boldsymbol{u})$, $m(\boldsymbol{u}_\alpha)$ are expected values (means) of $z(\boldsymbol{u})$ and $z(\boldsymbol{u}_\alpha)$; and $\lambda_\alpha(\boldsymbol{u})$ is the Kriging weight assigned to datum $z(\boldsymbol{u}_\alpha)$ for estimation location **u**; the same datum will receive a different weight for a different estimation location.

$z(\boldsymbol{u})$ is treated as a random field with a trend component, $m(\boldsymbol{u})$, and a residual component, $R(\boldsymbol{u}) = z(\boldsymbol{u}) - m(\boldsymbol{u})$. Kriging estimates residual at $\boldsymbol{u}$ as the weighted sum of residuals at surrounding data points. Kriging weights, $\lambda_\alpha(\boldsymbol{u})$, are derived from the covariance function or semivariogram, which should characterize the residual component. Distinction between trend and residual arbitrarily varies with scale.

For ordinary Kriging, rather than assuming that the mean is constant over the entire domain, the mean is assumed to be constant in the local neighbourhood of each estimation point, that is, $m(\boldsymbol{u}_\alpha) = m(\boldsymbol{u})$ for each nearby data value $z(\boldsymbol{u}_\alpha)$, which is used to estimate $z(\boldsymbol{u})$. Subsequently, the ordinary Kriging estimator can be expressed as

$$\begin{aligned} z^*(\boldsymbol{u}) &= m(\boldsymbol{u}) + \sum_{\alpha=1}^{n(\boldsymbol{u})} \lambda_\alpha(\boldsymbol{u}) [z(\boldsymbol{u}_\alpha) - m(\boldsymbol{u})], \\ &= \sum_{\alpha=1}^{n(\boldsymbol{u})} \lambda_\alpha(\boldsymbol{u}) z(\boldsymbol{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(\boldsymbol{u})} \lambda_\alpha(\boldsymbol{u})\right] m(\boldsymbol{u}). \end{aligned} \tag{9}$$

From equation (9), the interpolation advantages of the ordinary Kriging method include (i) the ability to compensate for the effects of data clustering, assigning individual points within a cluster less weight than isolated data points or treating clusters more like single points; (ii) giving estimates of the estimation error (Kriging variance)

along with the estimate of the variable, z; and (iii) the availability of estimation error providing the basis for stochastic simulation of possible realizations of $z(\boldsymbol{u})$.

Comparatively, the stochastic-based interpolation techniques are based on the assumption that the parameter being interpolated can be treated as a regionalized variable. A regionalized variable is intermediate between a truly random variable and a completely deterministic variable in that it varies in a continuous manner from one location to the next and therefore points that are near each other have a certain degree of spatial correlation, but points that are widely separated are statistically independent.

4.3 *Retrieved TRMM-3B42 precipitation data*

To convert the satellite rainfall rates to total daily rainfall, each 3-hourly rainfall rate was multiplied by 3 h to get the total rainfall for each 3 h period. Then, for the desired 24 h-day begin and end times, the sum of all the 3-hourly total rainfalls in the defined 24 h period is obtained to get the total daily rainfall (NASA GES-DISC Interactive Online Visualization and Analysis Infrastructure (GIOVANNI) User Manual).

The adjusted gauge data sets were then computed into a 3-year daily areal rainfall data. Figure 4 shows the mean daily gauge data plotted as scatter plots against the daily TRMM-3B42 rainfall retrievals. The mean daily precipitations depict a regular pattern in the wet and dry seasons, with the TRMM marginally overestimating the mean observed rainfall during the wet seasons and slightly underestimating the rainfall amounts during the dry seasons as compared with the gauge recordings. Overall, negligible adjustments were applied.

The 3-year mean monthly rainfall for Chorlim, Kakamega and Uholo representing the upper, middle plateau and lower elevation regions are shown in figure 5. Besides

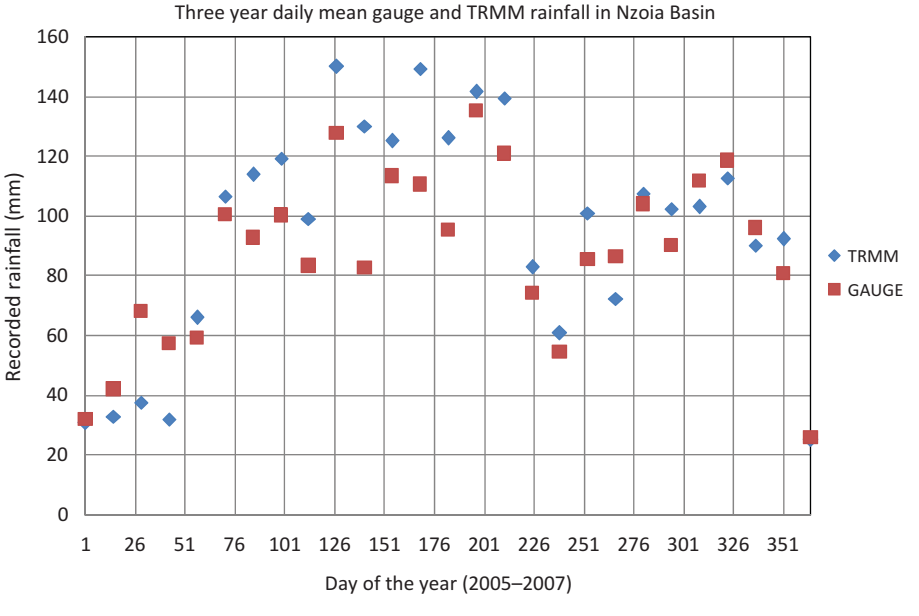

Figure 4. Scatter plot of adjusted cumulated daily gauged and TRMM_3B42.V6 recorded rainfall received in Nzoia River Basin in 2006.

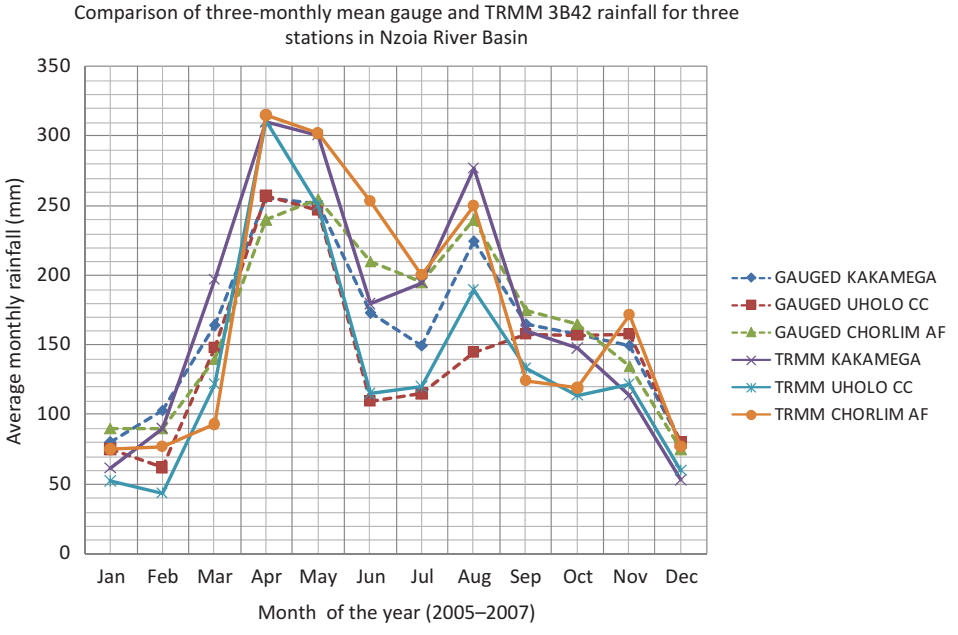

Figure 5. Mean monthly rainfall recorded from 2005 to 2007 for three representative rain gauge stations (Chorlim, Kakamega and Uholo) and TRMM_3B42.V6 satellite.

representing the distinct climate ranges and the three elevation zones within the Nzoia River Basin, the regions also capture the representative grids for the monthly TRMM-3B42 precipitation estimates.

A comparison of the mean daily precipitation (figure 4) and mean monthly (figure 5) precipitation patterns shows a similar pattern in two data sets, except for minor irregularity towards the end of the year. The plot in figure 5 shows that compared with gauge data, TRMM-3B42 overestimated the rainfall amounts during the rainy seasons and underestimated the recorded rainfall during the dry seasons. This observation can be attributed to the lower temperatures and more cloudy conditions during the rainy season, being sensed as contributing to rainfall occurrence, as compared with clear skies and sunny conditions during the dry seasons, being registered as no rainfall occurrences.

4.4 *Temporal analysis and comparison of TRMM-3B42 and rain gauge rainfall data*

4.4.1 Daily TRMM-3B42 and gauged stations rainfall data comparison. The 3-year mean daily rainfall observations are presented in figure 6. The gauge and TRMM-3B42 mean daily rainfall plots show marginal variations.

A regression plot (figure 7) of the TRMM-3B42 versus gauge data for the three consecutive years shows minimal annual disparities in the r^2 factor, with the respective daily coefficients of determination in 2005, 2006 and 2007 being 81.3%, 82.2% and 81.88%. These results show that more than 15% of the total variation between the gauge and TRMM-3B42 mean daily data compared remains uncorrelated.

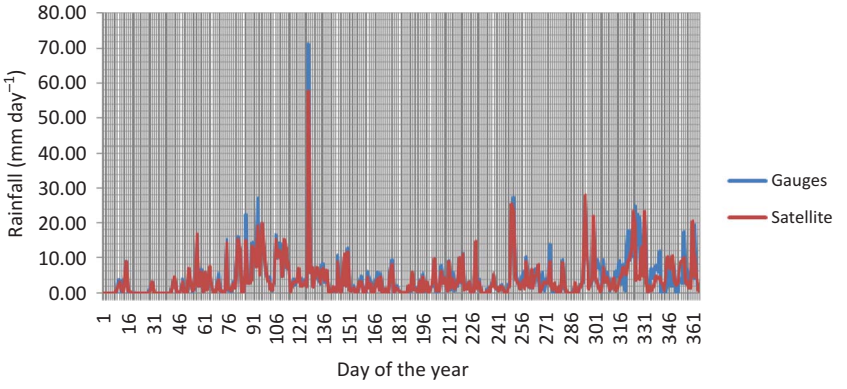

Figure 6. Mean daily rainfall estimates from TRMM and gauge observations for the three years over the Nzoia River Basin.

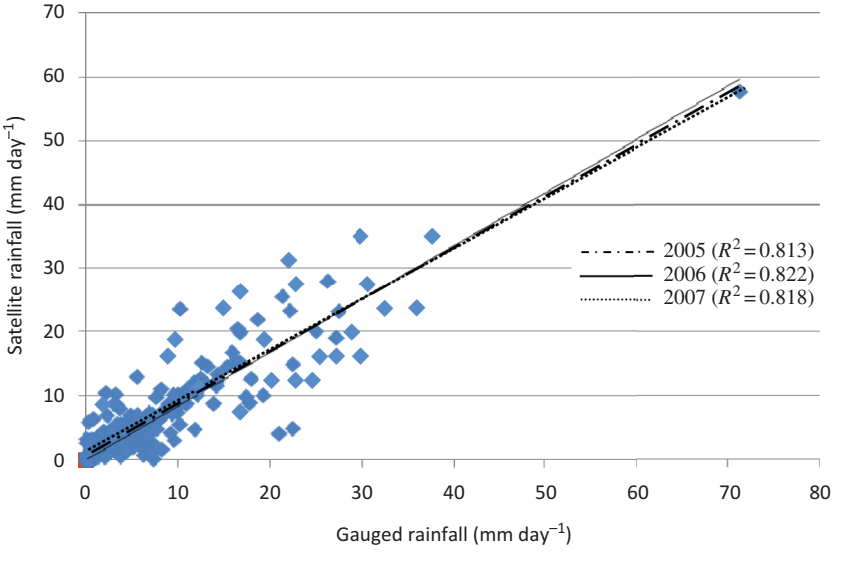

Figure 7. Regression (r^2) plot of TRMM_3B42.V6 versus gauged total daily rainfall in the Nzoia River Basin for the three test years.

4.4.2 Decadal TRMM-3B42 and gauged stations precipitation data comparison. The aggregated mean 10-day rainfall patterns (figure 8) depict dry spells at the onset and high rainfall recordings between the 9th and the 17th decads, followed by a dry spell. The dry period came to an abrupt end in the 26th decad, after which a short dry spell was experienced, recording the lowest rainfall in the 29th decad; thereafter the second high rainfall period was experienced in the month of November.

A decadal regression plot of the two data sets for the three years (figure 9) shows that the coefficient of determination increased from 2005 (r^2 = 0.907) to 2006 (r^2 = 0.939) and slightly dropped in 2007 to (r^2 = 0.915). Compared with the mean daily regression for the same period, it was observed that the mean decadal

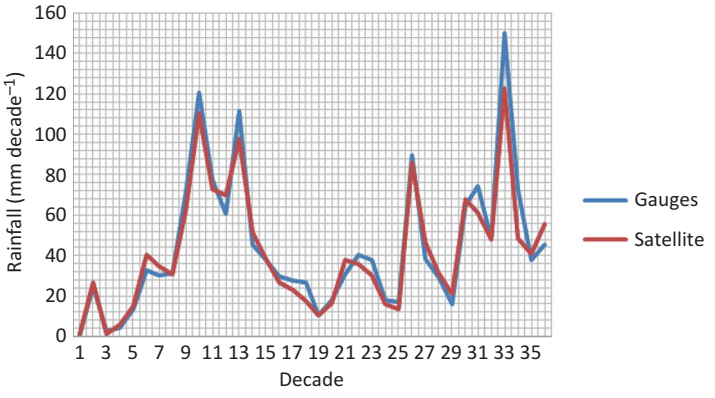

Figure 8. A plot of the mean decadal observed rainfall recorded from 2005 to 2007 in Nzoia River Basin from gauged stations and TRMM_3B42.V6 data product.

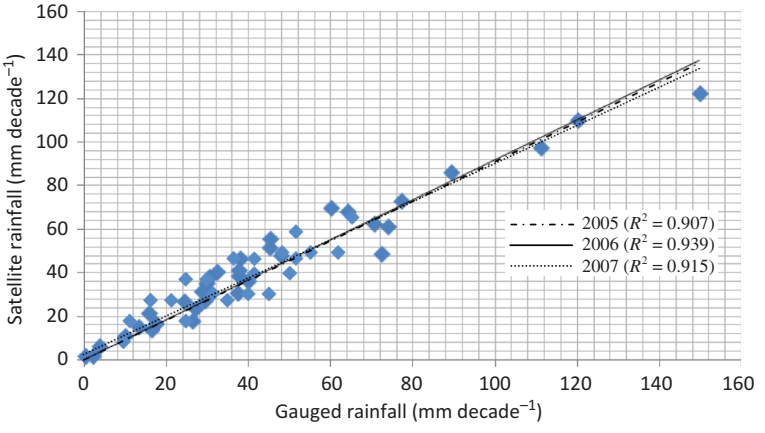

Figure 9. Mean aggregated decadal correlation coefficient (r^2) plots of TRMM_3B42.V6 versus gauge precipitation observations for the three study years.

rainfall correlations increased by more 10%, and only less than 10% of the gauge and TRMM-3B42 datasets were uncorrelated.

4.4.3 Monthly TRMM-3B42 and gauged stations precipitation comparison. Figure 10 shows the aggregated mean monthly precipitation over the Nzoia River Basin for the gauging stations and TRMM-3B42 data sets. The monthly averages were, respectively, observed as 128.4 mm month^{-1} and 131.6 mm month^{-1} for the gauged station and the satellite. The mean monthly difference is observed to be marginal.

The mean monthly coefficients of determination between the two data sets (figure 11) for the three years (2005, 2006 and 2007) were determined, respectively, to be 0.916, 0.965 and 0.944. Compared with the mean daily and decadal observations, the mean monthly rainfall for the three years were observed to be higher than mean

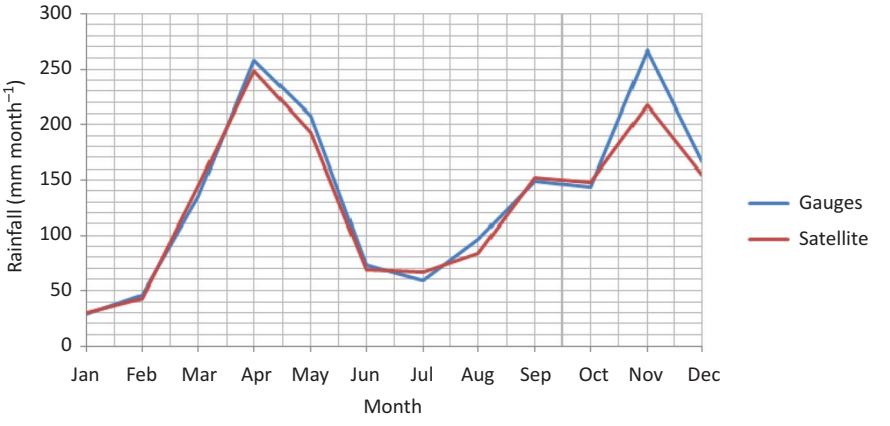

Figure 10. Mean monthly rainfall recorded from 2005 to 2007 in Nzoia River Basin from gauged stations and TRMM_3B42.V6 rainfall estimates.

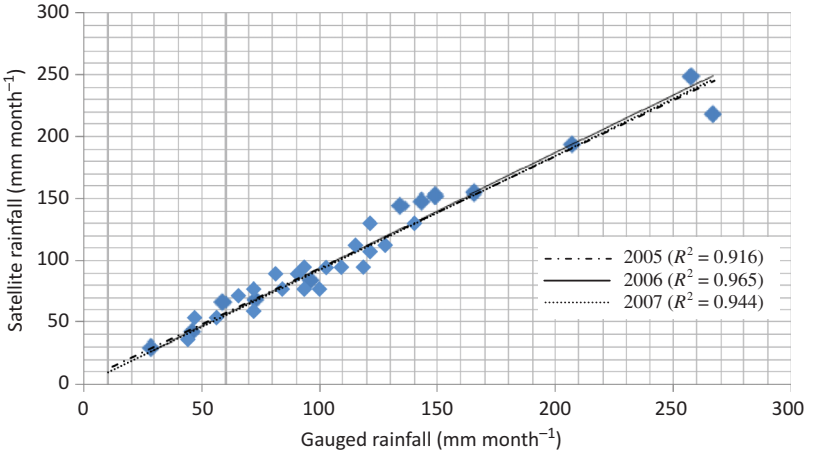

Figure 11. Mean monthly regression (r^2) plots of TRMM_3B42.V6 versus the gauge precipitation averages for the three years.

daily correlations by more than 10% and marginally (<5%) higher than the decadal correlation observations.

Statistical comparisons of the two data sets using the continuous verification measures, MBE, RMSD and MAD, are presented in table 3 for the three years. The results in table 3 indicate that the mean daily, decadal and monthly rainfall observations were overall lower in 2006. Due to the cumulative nature of the observed temporal data, the magnitudes of the compared statistics increased with temporal scales.

When MBE is reported, it is usually intended to indicate average model 'bias'; that is, average overprediction or underprediction. MBE conveys useful information but should be interpreted cautiously since it is inconsistently related to typical-error magnitude, other than being an underestimate. The results in table 3 consistently show a variation in the error magnitudes of the order: MBE $\leq$ MAD $\leq$ RSMD. From the results, RSMD tends to become increasingly larger than MAD (but not necessarily in

Table 3. Summary of the continuous verification statistical comparisons of the mean aggregated daily, decadal and monthly precipitation observations for the TRMM_3B42.V6 and the gauged precipitation data for the years: 2005, 2006 and 2007.

Time series	MBE (mm)			RMSD (mm)			MAD (mm)		
	2005	2006	2007	2005	2006	2007	2005	2006	2007
Daily	+1.29	−0.40	−0.66	+9.65	+6.89	+6.07	+2.87	+1.33	+1.79
Decadal	−2.93	−1.78	−2.04	+9.02	+8.61	+8.29	+9.16	+6.12	+7.76
Monthly	−7.81	−6.72	−7.01	+19.91	+16.22	+17.56	+13.09	+10.81	+11.33

a monotonic fashion) as the distribution of error magnitudes becomes more variable, and it tends to grow larger than the MAD.

4.5 *Discussion*

A comparative evaluation of the gridded averages showed that the gauge data observations were fairly close to the satellite rainfall retrievals in all of the three years. Optimal results were observed in 2006 with a monthly bias of –6.7 mm month^{-1}, decadal bias of –1.8 mm decad^{-1} and daily bias of –0.4 mm day^{-1}. The bias analysis results showed that there was a decrease in precipitation errors with increasing temporal intervals. Similar but not monotonic trends were observed for RMSD and MAD statistics.

While there is evidence of good correlation between the gauge and satellite precipitation estimates, the marginally higher intercept for the monthly observations suggests that there is a tendency for the lower observed rainfalls to be underestimated by the satellite, whereas the gentle slope values suggest that higher rainfall values are overestimated.

The observed increase in averaging times with resulting increase in the correlation distance can be attributed to the fact that there is an increased preservation of the statistical and structural characteristics of the observed rainfall field (Bell and Kundu 2003). The intermittence of rainfall in space and time introduces uncertainty to rainfall estimates based on the limited observations in space and time. This is particularly true for rainfall estimation from space with infrequent observations made at intervals of several hours (Bell *et al.* 1990, Steiner 1996, Bell and Kundu 2000, Bell *et al.* 2001).

Arguably, difficulties for spaceborne precipitation estimates may arise from the geographic characteristics of the investigation area. Nzoia River Basin, as depicted in figure 1, shows a varied topography ranging from the high, middle plateau to low elevation regions. As a consequence, heavy rainfall events can occur in the highland regions while the neighbouring lowland areas experience no rainfall at all, yet the pixel grid representing such zones depicts uniform rainfall amounts. This calls for an improvement in the precipitation retrieval algorithms that can take into consideration the mixed-pixel problem by computing on finer grids or disaggregating the grid through pixel unmixing to finer resolutions. This will also require denser rain gauge networks for comparative studies. Further, in mountainous terrain there is a strong effect of the relief itself on the IR signal, with varying cloud-cover conditions depending on exposure and altitude.

From the foregoing observations and discussions, the anticipated sampling-related uncertainty, say (ϵ), can be said to be a function of rainfall rate (R), domain size

(L), time integration (T) and sampling time interval (ΔT) such that the uncertainty is expected to reduce for higher rainfall rates, larger domain sizes and longer time integration. The uncertainty (ϵ) can be expressed as a function; thus, $\epsilon = f\left(\frac{1}{R}, \frac{1}{L}, \frac{\Delta T}{T}\right)$. This relationship implies that increasing the sampling time interval (i.e. reducing the sampling frequency) will result in a larger uncertainty. Also important in the error function is the averaging area sizes such that as the averaging area increases, the size of the satellite swath sample increases and this subsequently reduces the sampling error. This argument implies that investigations on the satellite-estimated rainfall error, against gauge data, should be modelled on the basis of multiplicative and additive random errors of the satellite precipitation estimates.

The results of the study show that the TRMM-3B42 precipitation estimates can effectively be used in the interpolation of missing rain gauge data with minor adjustments based on the coefficient of determination factor (r^2). The proposed adjustment largely depends on the temporal scales considered, the topographical variation and rainfall season. Generally, the correction index will be higher for the observed daily rainfall data as compared with the cumulative monthly rainfall data.

5. Summary and conclusions

This study presented a multitemporal comparative evaluation of precipitation retrievals from the TRMM-3B42 data with interpolated *in situ* measurements using 12 rain gauge stations located within the Nzoia River Basin. The evaluation was based on the examination of daily, decadal and monthly time series, using continuous verification statistics and scatter plots, for the determination of the measures of agreement and disagreement.

The results of the geostatistical interpolation of gauge data showed that stochastic ordinary Kriging methods gave the best results. The study showed that the spatial interpolation of the gauge data may not adequately and accurately reflect on the spatial–temporal dynamics of rainfall over the Nzoia River Basin, due to the sparse distribution of the gauge stations.

The mean aggregated spatial–temporal comparisons showed that the TRMM-3B42 rainfall estimates were very close to the spatially interpolated rain gauge data, especially at the daily temporal scales. For the three years, the rain gauge and the TRMM-3B42 time-series rainfall correlated well, with the daily coefficient of determination being lowest at 0.813 in 2005 and a maximum mean monthly coefficient of determination of 0.965 in 2006. The results depicted increased correlation with increased temporal interval of precipitation observations. The satellite and surface rain products showed that variances of diurnal cycle were less than 5% of intra-seasonal variances at the basin scale. The results further showed that the satellite observations slightly overestimated the precipitations during the wet seasons and underestimated the observed precipitations in the dry seasons.

The potential outcome of this approach is that it allows for extrapolation of the uncertainties to regions not covered by adequate validation data, but within the same climatic zones. The results of this study point to the importance of using TRMM 3B42 Version 6 (3B42) for rainfall estimation in regions with sparse rain gauge network distribution, especially at monthly temporal scales.

Finally, the quantification of the uncertainty of remotely sensed rainfall estimates is essential in providing guidance for interpretation of rainfall estimates from satellites for local, regional and global applications and in planning of future similar satellite

missions. Further studies using categorical statistical verification of inter-sensor comparisons of TRMM products with other satellite missions such as PERSIANN-Cloud Classification System (CCS) and CMORPH for the same study area and timescale are recommended. For long-term satellite data validation, a number of years with continuous and reliable gauge data need to be compared to further establish the reliability of the results for applications.

Acknowledgements

We thank the Kenyan Meteorological Department for providing the rainfall gauge data in Nzoia River Basin. The suggestions and comments by the editor and anonymous reviewers are highly acknowledged. We further thank the TRMM Science Data and Information System (TSDIS) and NASA's GIOVANNI for data used in this study.

References

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

Bell, T.L., Abdullah, A., Martin, R.L. and North, G.R., 1990, Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model. *Journal of Geophysical Research (Atmospheres)*, **95**, pp. 2195–2205.

Bell, T.L. and Kundu, P.K., 2000, Dependence of satellite sampling error on monthly averaged rain rates: comparison of simple models and recent studies. *Journal of Climate*, **2**, pp. 449–462.

Bell, T.L. and Kundu, P.K., 2003, Comparing satellite rainfall estimates with rain gauge data: optimal strategies suggested by a spectral model. *Journal of Geophysical Research (Atmospheres)*, **108**, 4121, doi:10.1029/2002JD002641.

Bell, T.L., Kundu, P.K. and Kummerow, C.D., 2001, Sampling errors of SSM/I and TRMM rainfall averages: comparison with error estimates from surface data and a simple model. *Journal of Applied Meteorology*, **40**, pp. 938–954.

Bellerby, T.J. and Sun, J., 2005, Probabilistic and ensemble representations of the uncertainty in an IR/microwave satellite precipitation product. *Journal of Hydrometeorology*, **6**, pp. 1032–1044.

Bowman, K.P., 2005, Comparison of TRMM precipitation retrievals with rain gauge data from ocean buoys. *Journal of Climate*, **18**, pp. 178–190.

Chiles, J.-P. and Delfiner, P., 1999, *Geostatistics: Modeling Spatial Uncertainty* (New York: Wiley-Interscience).

Chiu, L., Liu, Z., Rui, H. and Teng, W., 2005, TRMM data and access tools. In *Earth Science Satellite Remote Sensing*, vol. 2, J.J. Qu, W. Gao, M. Kafatos, R.E. Murphy and V.V. Salomonson (Eds.), pp. 202–219 (New York: Springer-Verlag and Beijing: Tsinghua University Press).

Collins, F.C. and Bolstad, P.V., 1996, A comparison of spatial interpolation techniques in temperature estimation. In *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*, 21–25 January 1996, Santa Fe, New Mexico (Santa Barbara, CA: National Center for Geographic Information Analysis (NCGIA)). CD-ROM.

Ebert, E., Janowiak, J.E. and Kidd, C., 2007, Comparison of near real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society*, **88**, pp. 47–64.

Goovaerts, P., 1997, *Geostatistics for Natural Resources Evaluation*, p. 496 (New York: Oxford University Press).

Greene, J.S. and Morrissey, M.L., 2000, Validation and uncertainty analysis of satellite rainfall algorithms. *The Professional Geographer*, **52**, pp. 247–258.

Griffith, D., 1993, Advanced spatial statistics for analyzing and visualizing geo-referenced data. *International Journal of Geographical Information Systems*, **7**, pp. 107–123.

Grimes, D.I.F. and Pardo-Iguzquira, E., 2010, Geostatistical analysis of rainfall. *Geographical Analysis*, **42**, pp. 136–160.

Holawe, F. and Dutter, R., 1999, Geostatistical study of precipitation series in Austria: time and space. *Journal of Hydrology*, **219**, pp. 70–82.

Hong, Y., Hsu, K., Gao, X. and Sorooshian, S., 2004, Precipitation estimation from remotely sensed imagery using artificial neural network cloud classification system. *Journal of Applied Meteorology*, **43**, pp. 1834–1853.

Hong, Y., Hsu, K., Moradkhani, H. and Sorooshian, S., 2006, Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. *Water Resources Research*, **42**, W08421, doi:10.1029/2005WR004398.

Hossain, F. and Anagnostou, E.N., 2004, Assessment of current passive-microwave- and infrared-based satellite rainfall remote sensing for flood prediction. *Journal of Geophysical Research (Atmospheres)*, **109**, D07102, doi:10.1029/2003JD003986.

Hossain, F., Anagnostou, E.N. and Bagtzoglou, A., 2006, On Latin Hypercube sampling for efficient uncertainty estimation of satellite rainfall observations in flood prediction. *Computers & Geosciences*, **32**, pp. 776–792.

Huffman, G.J., Adler, R.F., Curtis, S., Bolvin, D.T. and Nelkin, E.J., 2005, Global rainfall analyses at monthly and 3-hr time scales. In *Measuring Precipitation From Space: EURAINSTAT and the Future*, V. Levizzani, P. Bauer and F.J. Turk (Eds.), pp. 291–306 (Dordrecht: Kluwer Academic Publishers).

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P. and Stocker, E.F., 2007, The TRMM Multi-satellite precipitation analysis: quasi-global, multiyear, combined-sensor precipitation estimates at fine scale. *Journal of Hydrometeorology*, **8**, pp. 38–55.

Khan, S.I., Adhikari, P., Hong, Y., Vergara, H., Fadler, R., Policelli, F., Irwin, D., Korme, T. and Okello, L., 2011, Hydro-climatology of Lake Victoria region using hydrologic model and satellite remote sensing data. *Hydrology and Earth System Sciences*, **15**, pp. 107–117.

Krige, D.G., 1951, A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical Metallurgical and Mining Society of South Africa*, **52**, pp. 119–139.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J. and Simpson, J., 1998, The Tropical Rainfall Measuring Mission (TRMM) sensor package. *Journal of Atmospheric and Ocean Technology*, **15**, pp. 809–817.

Laslett, G.M., McBratney, A.B., Pahl, P.J. and Hutchinson, M.F., 1987, Comparison of several spatial prediction methods for soil pH. *European Journal of Soil Science*, **38**, pp. 325–341.

Legates, D.R. and Willmont, C.J., 1990, Mean seasonal and spatial variability in global surface air temperature. *Theoretical Application in Climatology*, **41**, pp. 11–21.

Li, L., Hong, Y., Wang, J., Adler, R.F., Policelli, R.S., Habib, S., Irwin, D., Korme, K. and Okello, L., 2009, Evaluation of real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Natural Hazards*, **50**, pp. 109–123.

Ly, S., Charles, C. and Degré, A., 2011, Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences*, **15**, pp. 2259–2274.

Matheron, G., 1970, The theory of regionalized variables and its applications. *Les Cahiers du Centre de Morphologie Mathematique*, **5**, 212 p. (ENSMP: Paris).

McCuen, R.H., 1989, *Hydrologic Analysis and Design*, 867 p. (Englewood Cliffs, NJ: Prentice Hall).

Negreiros, J., Painho, M., Aguilar, F. and Aguilar, M., 2010, Geographical information systems principles of ordinary kriging interpolator. *Journal of Applied Sciences*, **10**, pp. 852–867.

Pardo-Igúzquiza, E., 1998, Optimal selection of number and location of rainfall gauges for areal rainfall estimation using geostatistics and simulated annealing. *Journal of Hydrology*, **210**, pp. 206–220.

Petty, G.W. and Krajewski, W.F., 1996, Satellite estimation of precipitation over land. *Hydrological Sciences*, **4**, pp. 433–451.

Schuurmans, J.M. and Bierkens, M.F.P., 2007, Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. *Hydrology and Earth System Sciences*, **11**, pp. 677–693.

Stallings, C., Huffman, R.L., Khorram, S. and Guo, Z., 1992, *Linking Gleams and GIS. ASAE Paper 92- 3613*. (St. Joseph, MI: American Society of Agricultural Engineers).

Steiner, M., 1996, Uncertainty of estimates of monthly areal rainfall for temporally sparse remote observations. *Water Resources Research*, **32**, pp. 373–388.

Tetzlaff, D. and Uhlenbrook, S., 2005, Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data. *Hydrology and Earth System Sciences*, **9**, pp. 29–41.

Tian, Y., Peters-Lidard, C.D., Choudhury, B.J. and Garcia, M., 2007, Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *Journal of Hydrometeorology*, **8**, pp. 1165–1183.

Van De Beek, C.Z., Leijnse, H., Torfs, P.J.J.F. and Uijlenhoet, R., 2011, Climatology of daily rainfall semi-variance in The Netherlands. *Hydrology and Earth System Sciences*, **15**, pp. 171–183.

Wagner, S., Kunstmann, H., Bárdossy, A., Conrad, C. and Colditz, R.R., 2009, Water balance estimation of a poorly gauged catchment in West Africa using dynamically downscaled meteorological fields and remote sensing information. *Physics and Chemistry of the Earth*, **34**, pp. 225–235.

Webster, R. and Oliver, M., 2001, *Geostatistics for Environmental Scientists*, 271 p. (Chichester: John Wiley & Sons).

Willmott, C.J. and Matsuura, K., 2005, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, **30**, pp. 79–82.

Yan, J. and Gebremichael, M., 2009, Estimating actual rainfall from satellite rainfall products. *Atmospheric Research*, **92**, pp. 481–488.