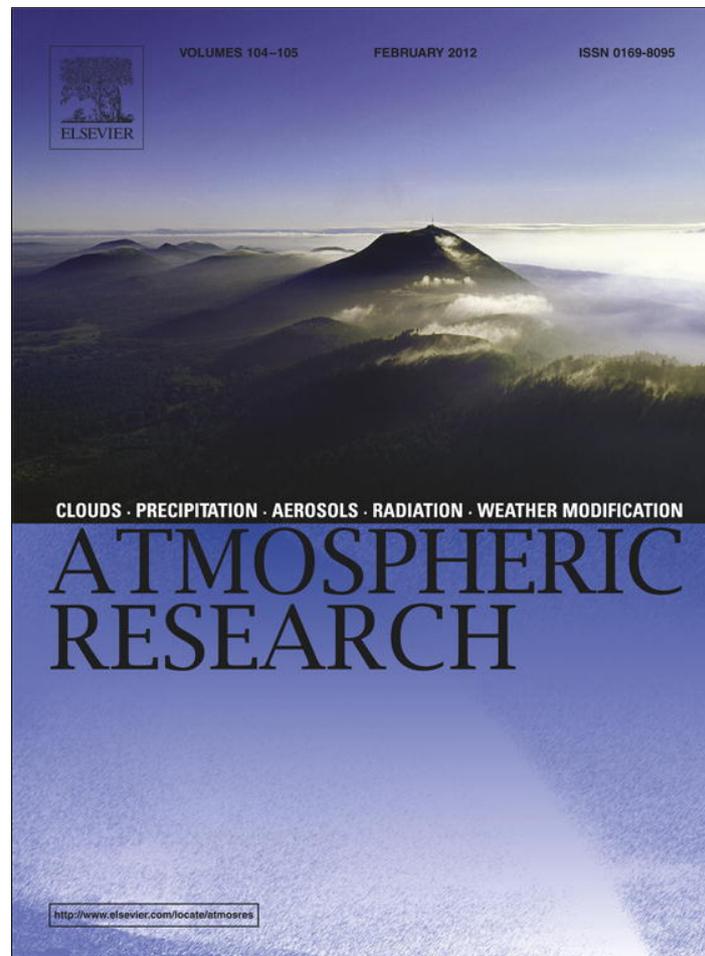


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmos

Investigating the similarity of satellite rainfall error metrics as a function of Köppen climate classification

Ling Tang, Faisal Hossain*

Department of Civil and Environmental Engineering, Tennessee Technological University, Cookeville, TN 38505-0001, USA

ARTICLE INFO

Article history:

Received 30 November 2010

Received in revised form 10 October 2011

Accepted 11 October 2011

Keywords:

Precipitation

Satellite

Uncertainty

Climatology

Similarity TRMM and GPM

ABSTRACT

This study addressed the question: *How much similarity exists in uncertainty of space-borne precipitation products for similar Köppen climate zones in different and distant landmasses?* Various metrics of satellite rainfall uncertainty were identified using a six year (2002–2007) archive of NASA's TRMM Multi-satellite Precipitation Analysis (TMPA) data product called 3B42V6 for two large distant landmasses that share many similar Köppen climate zones: 1) United States and 2) Australia. The level of quantitative similarity in error metrics for the same Köppen climate zones was then investigated. It was found that the bias and root mean squared error exhibited very close levels of similarity for similar Köppen climate zones in the US and Australia. However, similar inferences could not be drawn for other (higher-ordered) error metrics such as Probability of Detection (POD). The contrasting nature of the ground validation (GV) data (i.e., NEXRAD-radar in US and point gauge in Australia) for characterizing uncertainty may be one of the reasons for this observed lack of similarity. Using a dense gauge network of 42 gauges over a standard 3B42V6 grid box ($\sim 0.25^\circ$) as a ground validation benchmark, the dependence of uncertainty as a function of gauge density was quantified. These relationships were then cast in the context of our Köppen climate similarity experiment to identify the minimum level of gauge density that would be needed to resolve more accurately the actual level of similarity of error metrics for distant landmasses.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Satellite remote sensing of rainfall has witnessed tremendous progress in the last three decades (Gebremichael and Hossain, 2010). The first 'global' rainfall products were developed using satellite infrared sensors on geostationary orbits. In recent times, satellite rainfall estimation has seen significant improvement in resolution both spatially and temporally (Scofield and Kuligowski, 2003). From typical resolutions of degree-daily in the 1980s (such as the Global Precipitation Climatology Project-GPCP, Huffman et al., 2001), the current suite of high resolution precipitation products (HRPP) now provide rainfall estimates using various sensors and orbiting

platforms at typical scales of 25×25 km every 3 h or less across the globe. A few examples of such products are Climate Prediction Center MORPHing Technique (CMORPH; Joyce et al., 2004), Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; Huffman et al., 2007) and Global Satellite Mapping of Precipitation GSMaP (Ushio et al., 2009). Some satellite rainfall products also provide operational rainfall data at much smaller resolutions, such as 30 min 4×4 km (Behrangi et al., 2009). Hereafter, rainfall is used as shorthand for precipitation.

NASA's planned Global Precipitation Measurement (GPM) mission, in collaboration with other international partners, now represents a unique tool for observing precipitation from measuring satellites comprising a high-resolution, multi-channel PMW rain radiometer known as the GPM Microwave Imager (GMI), augmented by TRMM-like Dual-frequency Precipitation Radar (DPR) (Hou et al., 2008). GPM is currently scheduled for launch in 2013. GPM will seek to achieve

* Corresponding author at: Department of Civil and Environmental Engineering, Tennessee Technological University, 1020 Stadium Drive, Box 5015, Cookeville, TN 38505, USA. Tel.: +1 9313723257; fax: +1 9313726239.

E-mail address: fhossain@tntech.edu (F. Hossain).

measurements with a 3-hour average revisit time over 80% of the globe. It is also expected to provide high resolution (global) precipitation products with temporal sampling rates ranging from three to six hours and spatial resolution of 5 to 10 km (Hou et al., 2008). Considering that *in-situ* rainfall measuring networks are on the decline (Shiklomanov et al., 2002; Stokstad, 1999), the potential value of GPM to the global society cannot be underestimated (Hong et al., 2007).

GPM, together with the current range of HRPPs, will represent an interesting challenge in advancing the application of satellite HRPPs over ungauged regions (i.e., those lacking in ground measurement of rainfall). Most rainfall products employ their own distinct estimation methodology (i.e., algorithm) even though the primary raw data input is essentially the same (radiance from IR and PMW sensors and/or TRMM radar reflectivity for calibration). Because of the difference in estimation methodology, the accuracy of each product naturally does not have similar characteristics as a function of location, landform and climate features (Tian et al., 2010; Bitew and Gebremichael, 2010; Zeweldi and Gebremichael, 2009). Yet, many users need to know the errors of the satellite rainfall datasets across the range of time/space scales over the whole domain of the dataset. This uncertainty can provide insightful information to the user on the realistic limits of predictability that can be achieved with satellite rainfall HRPPs when used for applications such as hydrologic prediction (Hossain and Huffman, 2008). On the other hand, satellite rainfall datasets are most useful over the vast ungauged regions of the developing world lacking in ground validation (GV) data. Hence, the desire to provide uncertainty information and the need to promote global applications of GPM rainfall data or HRPPs by the data producers is currently rather difficult to reconcile.

This is where a 'climate classification' approach may be worthwhile for investigation. Generally, rainfall climatology refers to precipitation related weather conditions averaged over period of time. The potential value of elucidating the climatologic features of satellite rainfall estimation uncertainty can be better appreciated by the Köppen climate classification system. The Köppen climate classification is one of the most widely used climate classification systems developed by Wladimir Köppen, a Russian climatologist, around 1900. It combines the information on native vegetation, temperature, and precipitation to express the climate of the world into 29 patterns (McKnight and Hess, 2000). In Köppen climate classification, the world climates are divided into five main groups and several types and subtypes (Fig. 1). These five groups are: 1) Group A: Tropical/megathermal climates, 2) Group B: Dry (arid and semiarid) climates, 3) Group C: Temperate/mesothermal climates, 4) Group D: Continental/microthermal climates, and 5) Group E: Polar climates. Each Köppen climate group has a distinguishable mean temperature and precipitation patterns in space and time (seasons).

If the world climate could be amenable to such a Köppen 'climate classification map' where similar climate patterns are observed in various and distant parts of the world, then is it possible to achieve a similar error classification map for satellite HRPPs? And, if it was indeed possible to classify the spatial structure of uncertainty as a function of climate across the globe, then, can we 'transfer' the known uncertainty information from a GV site to a distant non-GV site known *a*

priori to have similar climatology? Because Köppen climate classification system includes the information of precipitation patterns across the space and time, we consider it as our first-cut pathfinder map to investigate a global error climatology classification scheme for satellite HRPPs. Our initial hypothesis is that a statistical similarity of uncertainty should exist in the same Köppen climate classes across the different landmasses of the world. Hereafter, no distinction is made between the terms 'uncertainty' and 'error'.

Recent research has revealed that the associated uncertainty of satellite rainfall does indeed vary distinctly in space and time (Gebremichael et al., 2010; Ebert et al., 2007; Dinku et al., 2007; Tian et al., 2007; 2010; Gottschalck et al., 2005; Huffman, 1997). Hence, if the 'average' and reproducible features of satellite rainfall uncertainty could be defined as a function of space and time scale, then we could potentially resolve the following open question — *How much similarity exists in uncertainty of space-borne precipitation products for similar Köppen climate zones in distant landmasses?*

In this paper, we address this open question. The data used as an example of a representative HRPP is a 6-year (2002–2007) archive of NASA's TRMM Multi-satellite Precipitation Analysis (TMPA) rainfall data over the United States (US) of America and Australia. TMPA rainfall data is currently available at a quasi-global scale, and may be considered a potential pathfinder to GPM-era design of HRPPs. Nine Köppen similar climate classes that occur in two distant landmasses (US and Australia) were investigated for resolving the open question posed above.

This paper is organized as follows: Section 2 discusses the study region and data; Section 3 presents the methodology and the definition of error metrics for HRPPs. Section 4 presents the results and discussion of our investigation. Finally, conclusions are presented in Section 5.

2. Study region and data

2.1. Study regions

We selected two distant landmasses for our investigation: 1) US and 2) Australia (Fig. 1). Both regions are large enough to incorporate various Köppen climate zones and more importantly, they share many between them. Furthermore, the two regions are sufficiently distant geographically that spatial correlation can be ruled out if there is any similarity in uncertainty. A previous study by Tang and Hossain (2009) has revealed that most error metrics for a typical HRPP product (such as 3B42RT) decorrelate to the 'noise' level after 5 grid boxes at the native resolution. The nine Köppen climate classes common to both US and Australia are summarized in Table 1.

2.2. Study data

In order to minimize the error of the GV data, we used the National Center for Environmental Prediction's (NCEP) 4 km Stage IV NEXRAD rainfall data for the US. This data is adjusted to gauges and conveniently available as a quality-controlled data mosaic over the US (Lin and Mitchell, 2005; Fulton et al., 1998). Because NEXRAD Stage IV-type data was unavailable in Australia, we used gauge data gridded at 0.25° and daily resolution as the GV data (obtained from the Bureau of

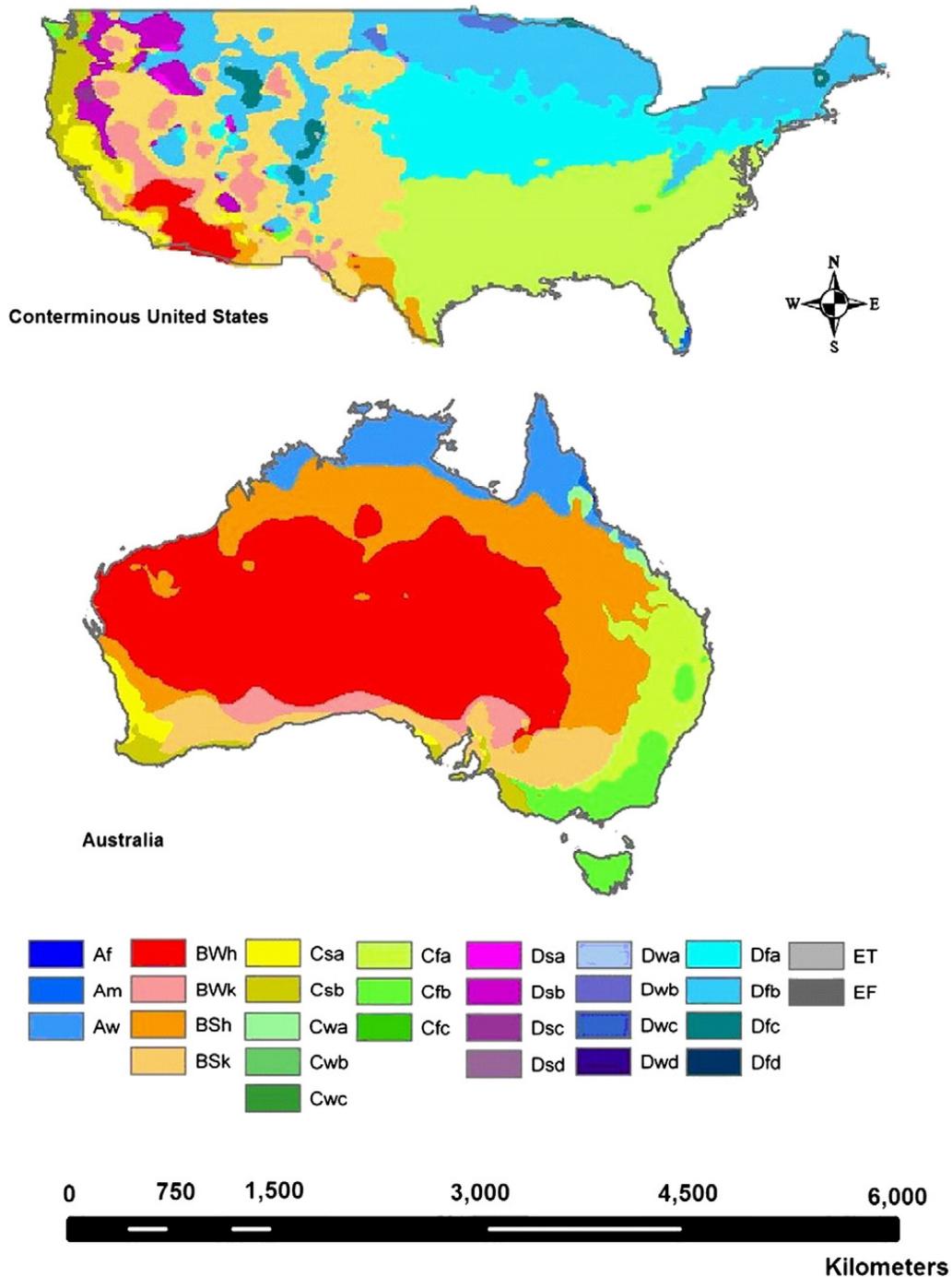


Fig. 1. Study region: Köppen map for US and Australia (lower panel) (After Peel et al., 2007).

Table 1

Köppen climate classes common to US and Australia.

Group	Types	Köppen climate features
A	Aw	Tropical
B	Bwk	Desert (cold)
	Bwh	Desert (hot)
	Bsk	Semi-arid (cold)
C	Bsh	Semi-arid (hot)
	Cfa	Humid subtropical
	Cfb	Oceanic
	Csa	Mediterranean (hot)
	Csb	Mediterranean (warm)

Meteorology, Australia; <http://www.bom.gov.au/>). Naturally therefore, our common temporal resolution of error metrics pertained to the daily time step to allow comparison between US and Australia, while spatial resolution was 0.25°. As a prelude to GPM, TMPA satellite rainfall data-product labeled as 3B42V6 was used as the satellite rainfall data (Huffman et al., 2007). This is a quasi-globally available as a research-grade product at 0.25° and 3 hourly resolution from the worldwide web (see <ftp://trmmopen.gsfc.nasa.gov> or <http://precip.gsfc.nasa.gov>). This data was aggregated and averaged at daily time step. The 3B42V6 data product is particularly suited for our investigation because it is better quality controlled (in

terms of data discontinuities and data homogeneity) than its real-time counterparts such as the IR-based 3B41RT or PMW-IR merged 3B42RT. The data for GV and satellite rainfall data spanned the period of 2002–2007 (6 years). Table 2 lists the details of these datasets.

For the US, the NEXRAD Stage IV GV rainfall data was first remapped to 0.25° 3 hourly resolution for consistency with the native scale of the satellite rainfall product. Remapping from 0.04° to 0.25° was a spatial aggregation process where values at 0.04° located within the 0.25 degree grid box were aggregated and averaged. In addition, hourly NEXRAD data was accumulated in 3 hour time steps, and averaged to generate the 3 hourly data.

The Australian rain gauge network consisted of up to 6000 sites that measured 24-hour accumulated rainfall at approximately 9 a.m. local time (approximately 00 UTC). Within these sites, about 1000 to 1500 reported daily rainfall in near-real time for the Australian Bureau of Meteorology's operational daily rainfall analysis. These additional data were combined with the original gauge data in the Bureau's climate database to produce a more accurate reanalysis of daily rainfall, which was used in this study. The data was already available gridded at 0.25°. This gridding was derived from a multi-pass inverse distance-weighting scheme to map the rainfall observations onto a 0.25-degree grid over Australia (Weymouth et al., 1999).

3. Methodology

3.1. The error metrics

Three widely used error metrics were computed for 3B42V6 (see Appendix A). These metrics are: Bias (BIAS), Root Mean Squared Error (RMSE) and Probability of Detection for rain (POD-rain). Readers should note that POD is also alternatively referred to as 'Percent Correct' or 'Hit Rate' in relevant literature. These metrics were computed for each grid box over the 6-year period. The assumption in this study is that a 6-year 'average' on the error would yield a relatively stationary spatial field of 'climatologic' error metrics for a given region or climate zone. The reader is referred to Ebert et al. (2007), Ebert (2008) and Hossain and Huffman (2008) for more details on the significance of these error metrics. The reader is also referred to the website: <http://www.cawcr.gov.au/projects/verification/> where useful guidance is provided on measures for validation of satellite precipitation products.

3.2. Testing the similarity of error metrics over for similar Köppen climate classes

The three error metrics were calculated for the two geographically distant landmasses for testing the hypothesis on

similarity of satellite rainfall uncertainty over the same Köppen climate class. As mentioned in Section 2, conterminous US and Australia contained nine common Köppen climate classes (Table 1). The grid boxes located within the same climate class (Fig. 1) were aggregated and the 6 year average error metrics calculated for both US and Australia. The joint probability of rain (i.e., detection) was also calculated. Our assumption was that joint probabilities will not be as sensitive to the GV data type and thus help us elucidate patterns more clearly (Villarini et al., 2009). The joint probability used in this paper was taken from Hossain and Anagnostou (2006). This is a POD of rain that calculates the probability that a satellite detects rains (>0), given that GV rain is also greater than a given threshold. The POD-rain is plotted as a function of increasing thresholds of GV rain rate. According to Hossain and Anagnostou (2006), this plot can be reasonably modeled as a sigmoidal function having the form:

$$\text{Joint probability}(R) = \frac{1}{A + \exp(-BR)} \quad (1)$$

where R is the threshold for GV data. In this study, comparison between the coefficients A and B in Eq. (1) for US and Australia Köppen classes revealed the extent to which hypothesis on error similarity was valid for joint probability of rain.

4. Results and discussion

Fig. 2 provides the comparison of the 6 year average BIAS and RMSE over United States and Australia for the same Köppen climate zone. For the same Köppen climate classes (x axis), we observe that the level of similarity is quite striking. Barring the negligible scalar difference (<10%) in the error metrics, our study indicates that the maximum and minimum 'climatologic' (6 year average) BIAS and RMSE occur in the same Köppen climate zones for both US and Australia. A point to note is that the error metrics for BIAS and RMSE are for unconditional cases (covering both rain and no-rain events). For climates with higher frequency of rain (e.g., Cfa-Humid Subtropical), we clearly see much closer agreement in BIAS and RMSE values for US and Australia (solid lines in Figs. 2a and b). This provides a strong indication that a world map of average satellite rainfall uncertainty for BIAS and RMSE may be possible for 3B42V6 using a classification system similar to that of the Köppen climate.

A potential drawback of unconditional metrics (BIAS and RMSE) may be that the large percentage of no-rain events are falsely contributing to this observed similarity. Hence, the conditional BIAS and RMSE were also computed. The conditional case refers to the events when GV data detected rain (see dashed lines in Figs. 2a and b). In general, a climate-based similarity still persists in the uncertainty for BIAS and

Table 2
Summary of satellite rainfall data and ground validation rainfall data.

DATA	Products	Native spatial scale (degree)	Native temporal resolution (hours)	Period (years)
Satellite rainfall data: NASA's TMPA	3B42RT	0.25	3	2002–2007
Ground validation: NEXRAD (for US)	Stage IV	0.04	1	2002–2007
Ground validation gridded gauge (for Australia)	Gauge	0.25	24	2002–2007

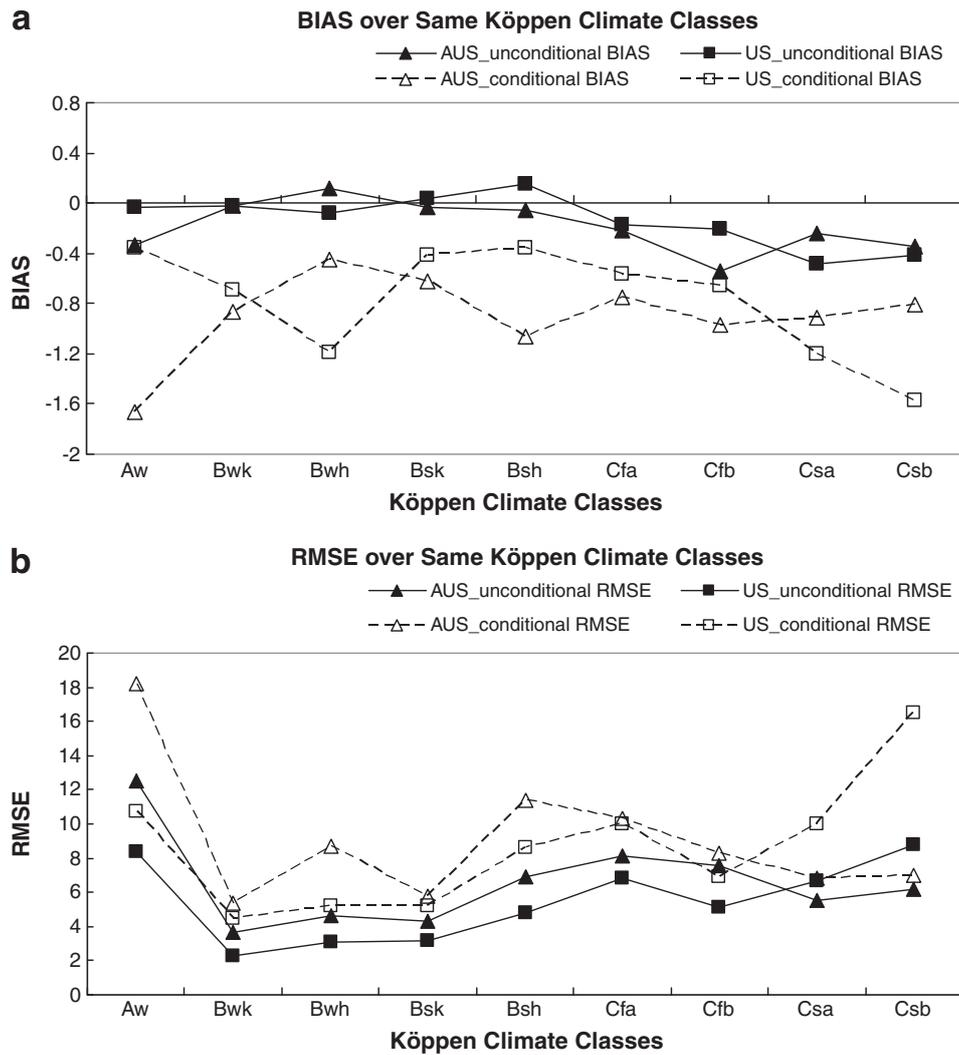


Fig. 2. a. Comparison of BIAS in United States and Australia over same Köppen climate classes for 3B42V6 (based on 6 years of data; units on y-axis: mm/day). b. Comparison of RMSE in United States and Australia over same Köppen climate classes for 3B42V6 (based on 6 years of data; units on y-axis: mm/day).

RMSE even when the rainy events (as defined by GV data) are exclusively considered. There are however cases where the climate regime-dependent values for conditional BIAS and RMSE are not quite similar. For example, the climate regimes Aw and Bwh, which have greatest differences in both conditional BIAS and RMSE, contain much less sampling points (83 and 773) in U.S. compared to Australia (1125 and 4777 points). This may indicate that the sampling error has inadvertently crept into the results, even with a 3-year dataset. A general implication of this finding is that the BIAS and RMSE (conditional or unconditional) for a GV site can be potentially ‘transferred’ to far away non-GV sites having the same Köppen climate provided some care is provided to sampling and scale.

Even though there appears to be a detectable level of similarity for metrics BIAS and RMSE for various climate classes, it is obvious that the terrain features (which can influence satellite rainfall estimation uncertainty) are not necessarily uniform. For example, for the Köppen class Cfa, that represents the humid sub-tropical climate, there may be mountainous as well as flat terrain regions distributed in both US and Australia. In order to investigate the similarity of metrics for a similar

terrain, we extracted the mountainous grid boxes for the regions of US and Australia belonging to the Cfa climate. For the US, 977 grid boxes were identified as ‘mountainous’ (near the Appalachian mountain range) from a total of 4025 Cfa grid boxes. For Australia, a similar sample size of 786 Cfa grid boxes were identified as mountainous. Table 3 summarizes the extent of similarity of metrics for BIAS and RMSE for the mountainous regions of the Cfa climate. It seems that the Köppen climate similarity hypothesis still holds even for similar terrain features within the same climate.

Fig. 3 provides the comparison of joint probability (of detection) of rain over US and Australia for the same Köppen

Table 3
Similarity of error metrics (BIAS and RMSE) for mountainous regions within the Cfa (humid subtropical) climate zone for US and Australia.

Error (mm/day)	USA (Cfa)		Australia (Cfa)	
	Unconditional	Conditional	Unconditional	Conditional
BIAS	-0.247023	-0.634599	-0.211505	-0.749554
RMSE	6.686498	9.302722	8.083081	10.269574

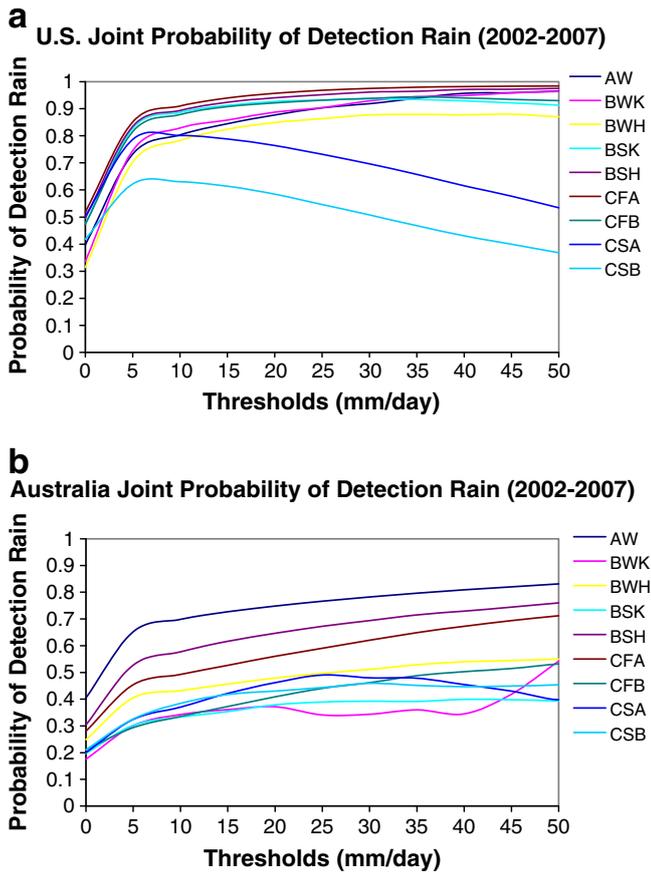


Fig. 3. a. Joint Probability of Detection for rain (POD-rain) over United States. X-axis represents the GV rain threshold for defining the probability of detection as satellite rain > 0 given GV rain > threshold. b. Joint Probability of Detection for rain (POD-rain) over Australia. X-axis represents the GV rain threshold for defining the probability of detection as satellite rain > 0 given GV rain > threshold.

climate classes. A different picture is revealed herein. The extent of similarity for the joint probability of detection rain (y axis) in United States and Australia is found to be quite low and unlike that for BIAS and RMSE. The maximum values for the POD-rain are larger in United States than in Australia for all the climate classes. This is not unexpected because Stage IV radar (GV data for United States) data, being a spatial average, is more sensitive at detection of rain events than the point-based gauge data (GV data for Australia).

To further quantify the statistics on the trend of the joint probability, we fitted the joint probability using a sigmoidal model function (in Section 3.2, Eq. (1)). There are two parameters in the model: A and B; A is inversely related to the maximum joint probability, while B represents the sensitivity to detecting the rain. Higher B means that the joint probability rises fast with increasing threshold (of GV data) and misses much less lighter rain than one with lower B. Table 4 shows how quantitatively different the parameters A and B are for US and Australia, while in Fig. 4, we show the POD-rain at a given rainfall threshold (5 mm/day and 20 mm/day) for various climate zones. Herein (Table 4 and Fig. 4), we find it hard to establish the hypothesis of similarity of error metrics in Köppen climate classes.

Because it is evident that the point-based nature of gauge data (even if it is gridded afterwards) led to an underestimation

Table 4

Joint probability (POD-rain) model parameters (A and B of Eq. (1)) over United States and Australia for similar Köppen climate classes.

Joint probability model (A and B)	United States		Australia	
	A	B	A	B
AW	1.1	1.1	1.4	0.5
BWK	1.1	1.1	3.75	1
BWH	1.2	0.5	2.1	0.25
BSK	1.05	0.5	4.2	0.5
BSH	1.05	0.5	1.5	0.1
CFA	1.05	0.5	1.75	0.1
CFB	1.05	0.5	2.25	0.1
CSA	1.25	1.5	3.75	0.1
CSB	1.5	1.5	3.6	0.1

of POD-rain compared to a spatially-averaged GV (like NEXRAD), we also investigated the possible role played by gauge density using a small-scale highly dense rain gauge network (Micronet: <http://grl.ars.usda.gov/micronet/>) The Micronet network is located in the Little Washita River Experimental Watershed in southwestern Oklahoma of the US with 5 minute meteorological data from 42 stations covering an area of 610 km² (approximately one 0.25 degree 3B42V6 grid box; see Fig. 5). Micronet data is quality controlled and flagged for bad quality data, which is very important because a significant component for a successful research study is the requirement of high quality data (Anagnostou et al., 2010). The Micronet data was available for the period of 2002–2004 (3 years).

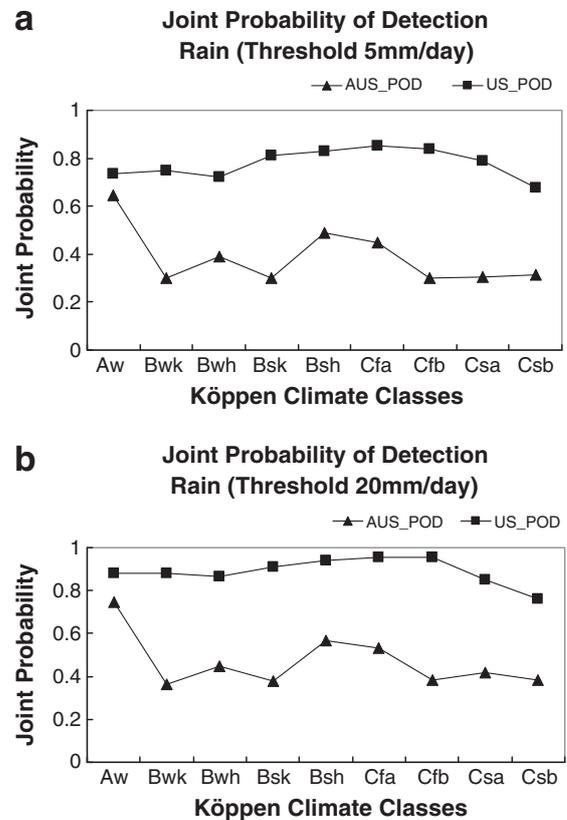


Fig. 4. a. Joint Probability of Detection for rain (POD-rain) over Australia and US, for GV rain > threshold 5 mm/day. b. Joint Probability of Detection for rain (POD-rain) over Australia and US, for GV rain > threshold 20 mm/day.

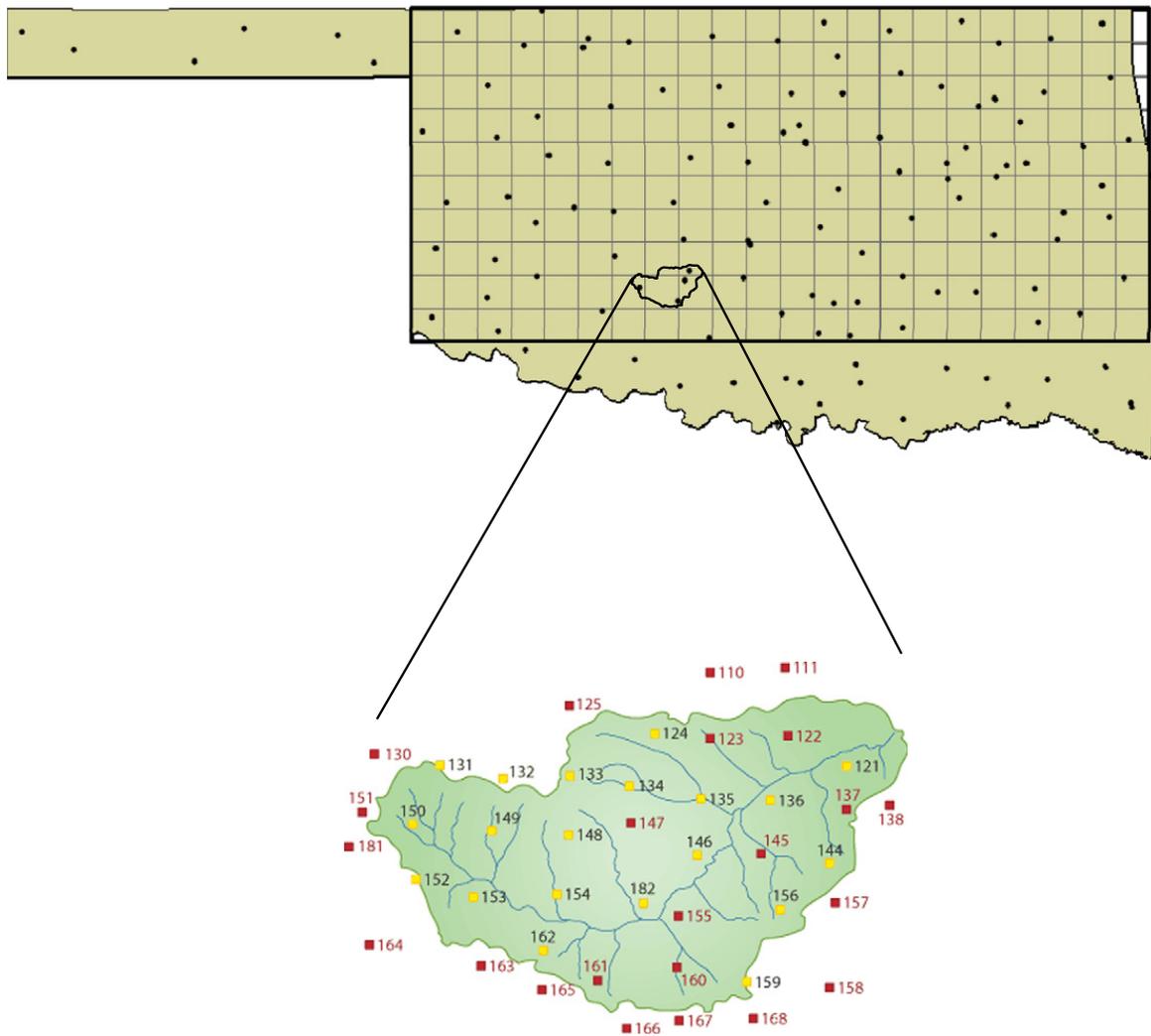


Fig. 5. The Oklahoma Mesonet (black dots on the Oklahoma map) and Micronet (location shown on the Little Washita River watershed shown above in lower panel) networks. The stations marked with red color for Micronet are retired since 2005. The small grid boxes shown (in grey) inside of the big black box for Mesonet is of 0.25 degree size. There are 42 gauges within the 0.25 degree grid boxes comprising the Micronet. Each grid box on the upper panel is equivalent to a 0.25 degree grid box [Figure from Anagnostou et al. (2010)].

Fig. 6 shows the dependence of the 3 year average values of BIAS, RMSE and POD-rain on Micronet gauge density. The POD-rain shown here refers to the marginal probability of detection (see Appendix A). This dependence is shown for NEXRAD, 3B42V6 and Gauge. Here, 'Gauge' is analogous to the Australian scenario where only one Micronet gauge (randomly selected) was used. As for the sensitivity analysis, a set of gauges numbering 1, 5, 10, 20, 30 and 42 were randomly sampled 10 times from the total set of 42 gauges. This means that once the Micronet stations were sampled, the mean rainfall value was computed as the average of the sampled stations. This average yielded the GV data time series realization (for a given combination of gauges), which was then used for the calculation of error metrics. Finally, the BIAS, RMSE and POD-rain computed for each realization (of random selection of Micronet gauges) were averaged over the 10 realizations to produce the expected value (mean) of uncertainty for a given combination of Micronet gauges.

The dependence of error metrics on gauge density may explain the reported lack of similarity for POD-rain between

US and Australian Köppen climate zones. For BIAS and RMSE, it appears that gauge density is not a critical factor. A maximum number of 5 gauges per 0.25 degree grid box appear sufficient to obtain stable estimates of error over a 3 year average. The Australian GV data had typically 1 gauge within a 0.25 degree grid box, which may explain the observed similarity observed for BIAS and RMSE between US and Australia for similar climate zones. On the other hand, POD-rain (lowermost panel) appears highly sensitive to the gauge density, barely stabilizing even after 20 gauges in the 0.25 degree grid box to yield a somewhat stable error estimate for different data types (Fig. 6).

Herein it is important to raise the issue of rainfall threshold to define an event as rainy or non-rainy. On the surface, the requirement of a very high density (more than 20 per 0.25 degree box) of gauges for POD-rain estimates to stabilize may seem somewhat counter-intuitive as one would expect the Micronet rainfall hyetograph to converge to the 'true' series with the use of fewer gauges. However, the Fig. 6 POD plots were derived for a threshold of 0.0 mm/day, which perhaps contributed to the noisy nature of the plot. On the other hand, if a systematically

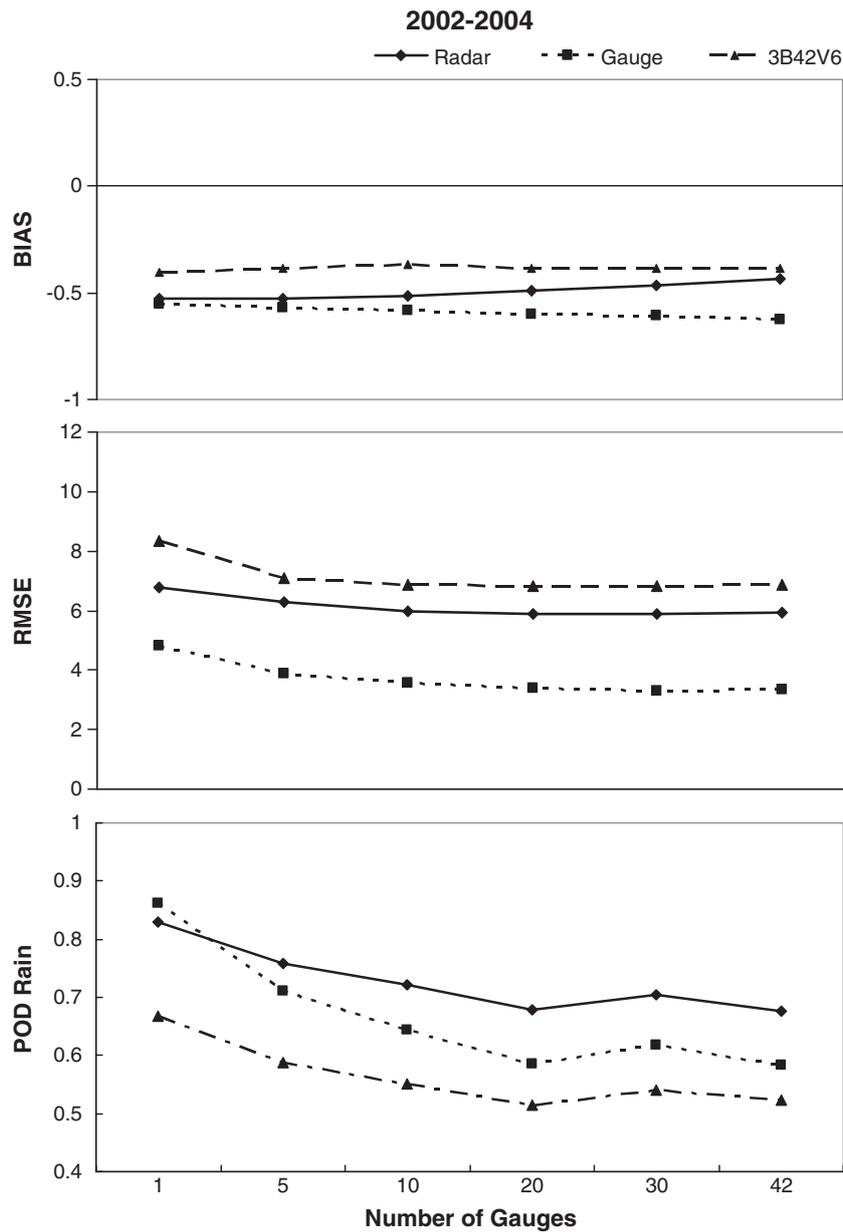


Fig. 6. Estimation of error metrics as a function of Micronet gauge density in a typical 0.25 degree 3B42V6 grid box for three years from 2002 to 2004. Here, 'Radar' (solid line) refers to NEXRAD, while Gauge (dotted line) refers to one gauge selected randomly from the 42 Micronet gauges (BIAS and RMSE unit: mm/day). POD (lowermost pane) refers to detection for rain with a small threshold of 0.01 mm/day.

increasing non-zero threshold is used, the POD is seen to stabilize after a finite number of gauges. For example, a threshold of 0.01 mm/day reveals a 20 gauge/0.25 degree density is adequate to achieve stable estimates of POD (Fig. 7). A threshold of 0.05 mm/day reveals the stabilization number of gauges is considerably less and approximately around 5 gauges (see Fig. 8).

The issue of threshold raises an interesting question that is beyond the scope of this study. *What should be an acceptable threshold for POD analysis?* To a hydrologist engaged in hourly simulation of hydrologic fluxes, 0.05 mm/day (equivalent to 0.0021 mm/h) might be considered measurement 'noise' for the sensor or rain gauge (radar, tipping bucket or satellite), or it might even be obscured by model noise and precision issues. This issue certainly needs more investigation as methods are

prototyped for transfer of error metrics during the GPM era. At the moment, it appears that a Köppen-type climatologic classification map for POD (whether rain, no-rain or combined) cannot be resolved yet without access to highly dense gauge network at GV sites.

5. Conclusion

Our investigation revealed that only the first and second ordered moments of error (BIAS and RMSE) are most amenable to a Köppen-type climate type classification in different landmasses. These two error metrics exhibited striking similarities for maximum and minimum values for a 6 year average for distant landmasses. Use of a high gauge density network (Micronet) revealed further that contingency based metrics such as

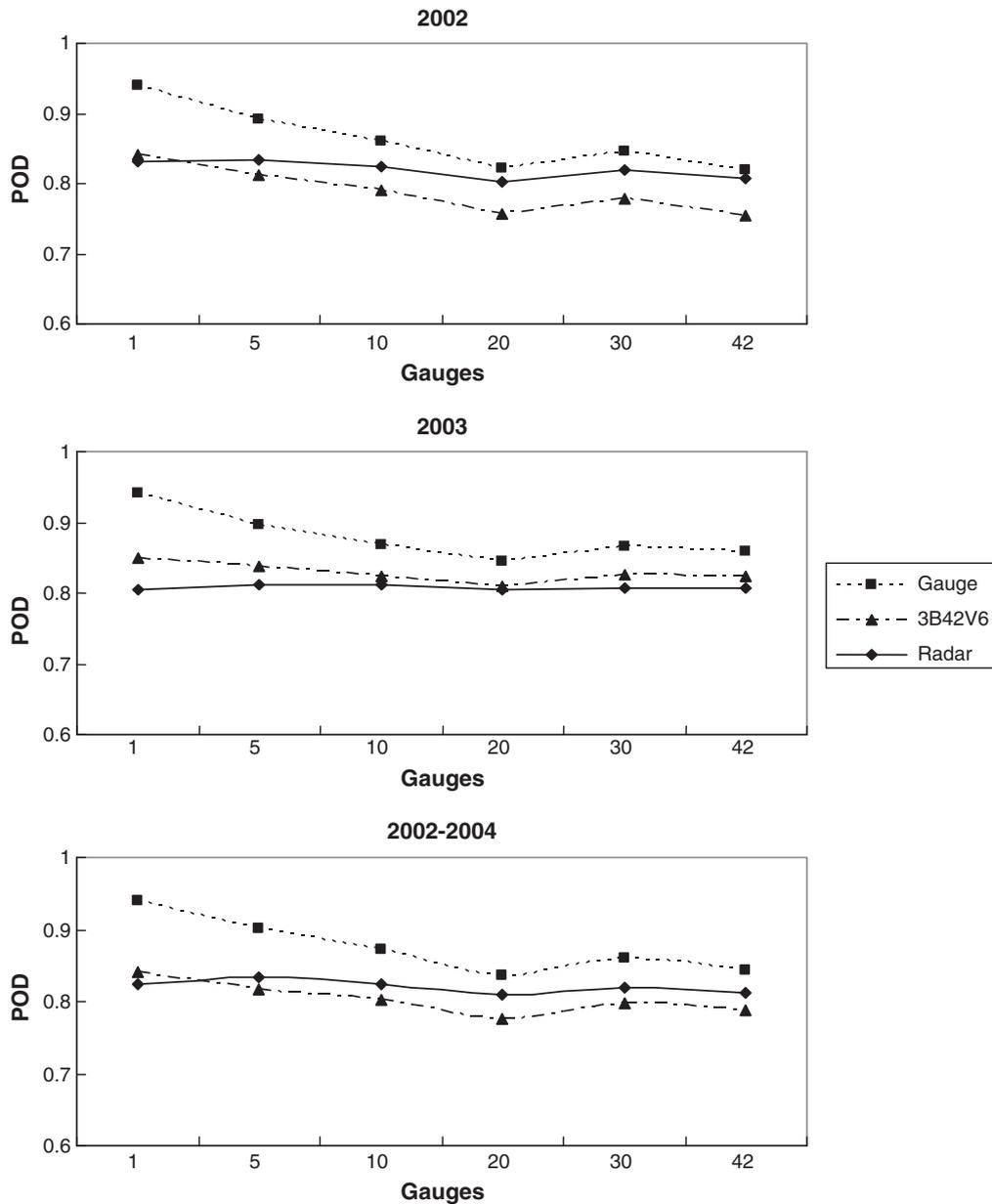


Fig. 7. Probability of Detection for rain and no-rain (POD) as a function of Micronet gauge density in a typical 0.25 degree grid box of 3B42V6, with a threshold of 0.01 mm/day.

POD are strongly sensitive to the gauge density to yield stable estimates.

The results of this climate similarity investigation indicate that the long-term average BIAS and RMSE for a GV site can be used as a proxy for a far away non-GV (ungauged) site provided the Köppen climates are similar. On the other hand, such a method does not seem likely to succeed for higher-ordered metrics as of yet. As mentioned earlier, most satellite HRPPs use essentially the same raw data input. It is basically the nature of the estimation methodology combined with the rain system being applied on that dictates the space-time structure of error metrics in the data domain. In this regard, the work of Berg et al. (2006) may be insightful in furthering the investigation of similarity of error metrics. By breaking up uncertainty into difference and detection (as our study) and connecting it to physical variables such as column water

vapor, rain system (convective and stratiform), aerosol concentrations etc. Berg et al. (2006) provides a foundation for exploring how the similarity (or lack of) in uncertainty can be explained.

In the lead up to GPM, it may now be appropriate to investigate the various estimation methodologies used in HRPPs as a function of rain systems, location and climate in order to understand the similarity of error according to an easily 'classifiable' and 'mappable' underlying pattern. Such patterns can help estimate uncertainty of GPM products over ungauged regions where satellite rainfall data will be most useful. Errors are also known to have a strong dependency on the magnitude of rainfall. So, the study of joint distribution of error with rainfall estimates would be worthwhile as an extension of the climate similarity work. It is our belief that an underlying pattern will most likely be revealed according to some

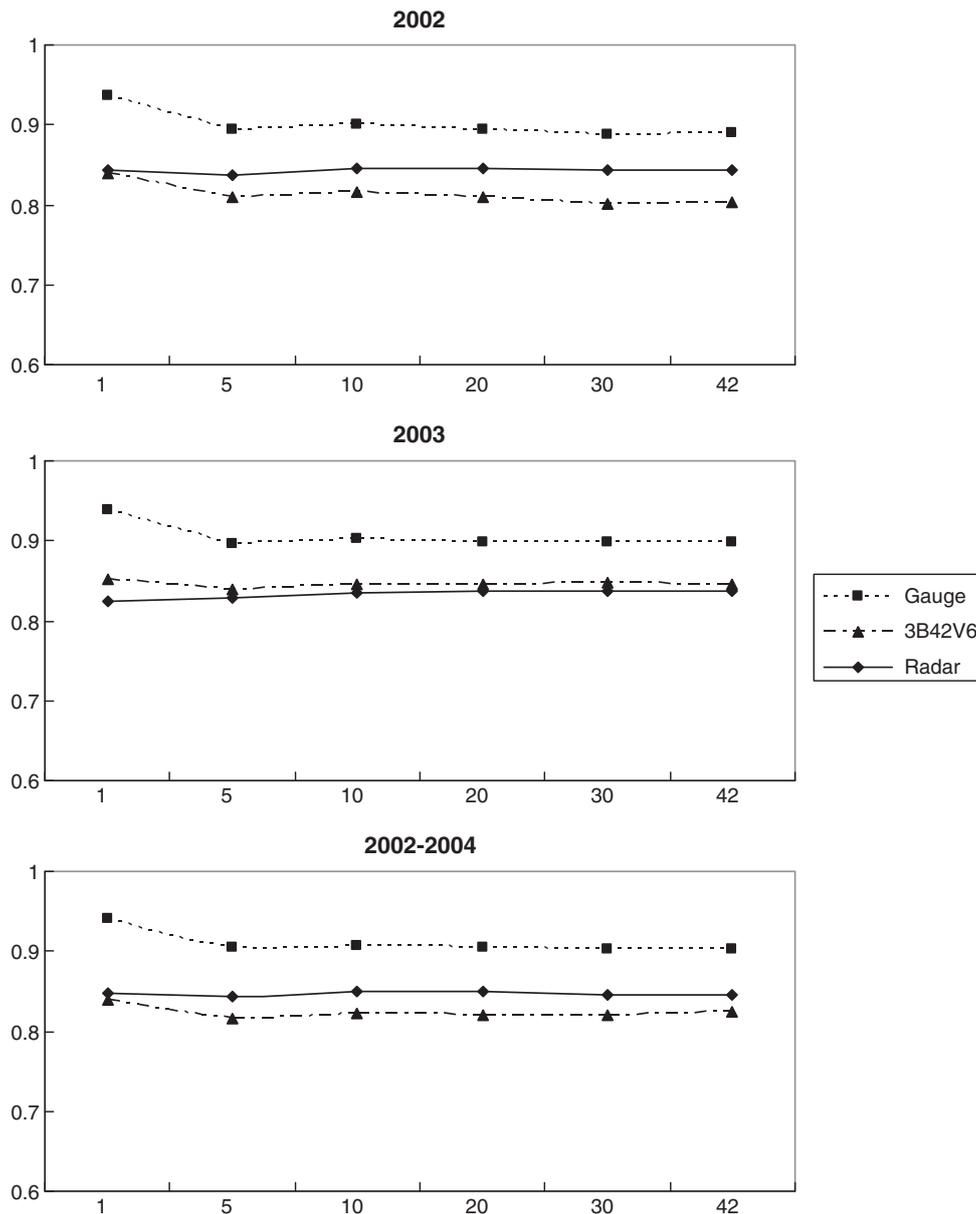


Fig. 8. Probability of Detection for rain and no-rain (POD) as a function of Micronet gauge density in a typical 0.25 degree grid box of 3B42V6, with a threshold of 0.05 mm/day.

simple relationship that could be leveraged for the creation of a global classification map on satellite rainfall estimation errors.

Acknowledgment

The first author (Tang) was supported by the NASA Earth System Science Fellowship (2008–2011). The second author (Hossain) was supported by the NASA New Investigator Program Award (NNX08AR32G). Authors acknowledge guidance received from Dr. Beth Ebert of Bureau of Meteorology-Australia for collection of ground validation data for the Australian region. Authors also acknowledge Dr. George Huffman of SSAI/NASA Laboratory of Atmospheres for guidance on the use of 3B42RT. The paper significantly improved in content after the reviews received from two anonymous reviewers.

Appendix A. Formulation of error metrics

For consistency in comparison between US and Australia, the temporal resolution at which the metrics was computed was daily. The spatial resolution was 0.25°.

BIAS

If i indicates one daily data in a grid box (derived from eight 3 hourly values of the day), Y_i is the satellite estimate, and X_i is the corresponding NEXRAD observation. N is the number of data in the time period, then BIAS is computed as the average of errors as:

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^N (Y_i - X_i).$$

RMSE

RMSE (Root Mean Square Error) is calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - X_i)^2}$$

Probability of Detection (POD)

		GV detection	
		Yes	No
Satellite detection	Yes	Hits	False alarms
	No	Misses	Correct rejection

POD measures the fraction of rain or no-rain events that were correctly detected by satellite.

$$\text{POD-rain} = \frac{\text{hits}}{\text{hits} + \text{misses}}$$

$$\text{POD-norain} = \frac{\text{correct.rejection}}{\text{correct.rejection} + \text{false.alarms}}$$

In this study, the focus of POD was mainly on the detection of rain events (POD-rain), while POD for detection of both rain and no rain is the sum of POD-rain and POD-norain.

References

- Anagnostou, E.N., Maggioni, V., Nikolopoulos, E., Taye, T., Hossain, F., 2010. Benchmarking high-resolution global satellite rain products to radar and rain gauge rainfall estimates. *IEEE Trans. Geosci. Remote Sens.* 48 (4), 1667–1683. doi:10.1109/TGRS.2009.2034736.
- Behrangi, A., Hsu, K., Imam, B., Sorooshian, S., Huffman, G.J., Kuligowski, R.J., 2009. PERSIANN-MSA: a precipitation estimation method from satellite-based multispectral analysis. *J. Hydrometeorol.* 10, 1414–1429.
- Berg, W., L'Ecuyer, T., Kummerow, C., 2006. Rainfall climate regimes: the relationship of regional TRMM rainfall biases to the environment. *J. Appl. Meteorol. Climatol.* 45, 434–454. doi:10.1175/JAM2331.1.
- Bitew, M.M., Gebremichael, M., 2010. Spatial variability of daily summer rainfall at a local-scale in a mountainous terrain and humid tropical region. *Atmos. Res.* 98 (2–4), 347–352.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S.J., Ropelewski, C.F., 2007. Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* 28 (7), 1503–1526. doi:10.1080/01431160600954688.
- Ebert, E.E., 2008. Fuzzy verification of high resolution gridded forecasts: a review and proposed framework. *Meteorol. Appl.* 15, 51–64.
- Ebert, E.E., Jonowiak, J.E., Kidd, C., 2007. Comparison of near real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteorol. Soc.* 88 (1), 47–64.
- Fulton, R.A., Breidenbach, J.P., Seo, D.-J., Miller, D.A., O'Bannon, T., 1998. The WSR-88D rainfall algorithm. *Weather. Forecast.* 13 (2), 377–395.
- Gebremichael, M., Hossain, F. (Eds.), 2010. *Satellite Rainfall Applications for Surface Hydrology*. Springer Publications, Germany, 327 pp.
- Gebremichael, M., Wang, J., Sammis, T.W., 2010. Dependence of remote sensing evapotranspiration algorithms on spatial resolution. *Atmos. Res.* 96 (4), 489–495. doi:10.1016/j.atmosres.2009.12.003.
- Gottschalk, J., Meng, J., Rodell, M., Houser, P., 2005. Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. *J. Hydrometeorol.* 6, 573–598. doi:10.1175/JHM437.1.
- Hong, Y., Adler, R.F., Hossain, F., Curtis, S., 2007. Global runoff simulation using satellite rainfall estimation and SCS-CN method. *Water Resour. Res.* 43 (W08502). doi:10.1029/2006WR005739.
- Hossain, F., Anagnostou, E.N., 2006. A two-dimensional satellite rainfall error model. *IEEE Trans. Geosci. Remote Sens.* 44 (6), 1511–1522. doi:10.1109/TGRS.2005.863866.
- Hossain, F., Huffman, G.J., 2008. Investigating error metrics for satellite rainfall at hydrologically relevant scales. *J. Hydrometeorol.* 9 (3), 563–575.
- Hou, A., Jackson, G.S., Kummerow, C., Shepherd, C.M., 2008. Global precipitation measurement. In: Silas, Michaelides (Ed.), *Precipitation: Advances in Measurement, Estimation, and Prediction*. Springer Publishers, pp. 131–164.
- Huffman, G.J., 1997. Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteorol.* 36, 1191–1201. doi:10.1175/1520-0450.
- Huffman, G.J., Adler, R.F., Morrissey, M.M., et al., 2001. Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.* 2, 36–50.
- Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM multi-satellite precipitation analysis: quasi-global, multi-year, combined sensor precipitation estimates at fine scales. *J. Hydrometeorol.* 8, 28–55.
- Joyce, R.L., Janowiak, J.E., Arkin, P.A., 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.
- Lin, Y., Mitchell, K., 2005. The NCEP Stage II/IV hourly precipitation analyses: 1999 development and applications. 19th AMS Conference on Hydrology.
- McKnight, T.L., Hess, D., 2000. *Climate Zones and Types: The Köppen System*. Physical Geography: a Landscape Appreciation. Prentice Hall, Upper Saddle River, NJ. ISBN: 0-13-020263-0, pp. 200–201.
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* 11, 1633–1644.
- Scofield, R.A., Kuligowski, R.J., 2003. Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Weather. Forecast.* 18, 1037–1051.
- Shiklomanov, A.I., Lammers, R.B., Vörösmarty, C.J., 2002. Widespread decline in hydrological monitoring threatens pan-arctic research. *EOS Trans.* 83 (2), 16–17.
- Stokstad, E., 1999. Scarcity of rain, stream gages threatens forecasts. *Science* 285, 1199.
- Tang, L., Hossain, F., 2009. Transfer of satellite rainfall error from gauged to ungauged locations: how realistic will it be for the global precipitation mission? *Geophys. Res. Lett.* 36. doi:10.1029/2009GL037965.
- Tian, Y., Peters-Lidard, C.D., Choudhury, B., Garcia, M., 2007. Multitemporal analysis of TRMM based satellite precipitation products for land data assimilation applications. *J. Hydrometeorol.* 8, 1165–1183. doi:10.1175/2007JHM859.1.
- Tian, Y., Peters-Lidard, C.D., Adler, R.F., Kubota, T., Ushio, T., 2010. Evaluation of GSMaP precipitation estimates over the contiguous United States. *J. Hydrometeorol.* 11, 566–574. doi:10.1175/2009JHM1190.1.
- Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K., Aonashi, K., Inoue, T., Takahashi, N., Iguichi, T., Kachi, M., Oki, R., Morimoto, T., Kawasaki, Z.-I., 2009. A Kalman filter approach to the global satellite mapping of precipitation (gsmap) from combined passive microwave and infrared radiometric data. *J. Meteorol. Soc. Jpn.* 87A, 137–151.
- Villarini, G., Krajewski, W.F., Smith, J.A., 2009. New paradigm for statistical validation of satellite precipitation estimates: application to a large sample of the TMPA 0.25 3-hourly estimates over Oklahoma. *J. Geophys. Res.* 114, D12106. doi:10.1029/2008JD011475.
- Weymouth, G., Mills, G.A., Jones, D., Ebert, E.E., Manton, M.J., 1999. A continental-scale daily rainfall analysis system. *Aust. Meteorol. Mag.* 48, 169–179.
- Zeweldi, D., Gebremichael, M., 2009. Sub-daily scale validation of satellite-based high-resolution rainfall products. *Atmos. Res.* 92 (4), 427–433. doi:10.1016/j.atmosres.2009.01.001.