Atmospheric Research xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Atmospheric Research

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

An improved procedure for the validation of satellite-based precipitation estimates

Ling Tang [a,b,*], Yudong Tian [a,b], Fang Yan [c], Emad Habib [c]

[a] Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD, 20740, USA
[b] Code 617, Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA
[c] Department of Civil Engineering, University of Louisiana at Lafayette, P.O. Box 42991, Lafayette, LA, 70504, USA

ARTICLE INFO

Article history:
Received 21 March 2014
Received in revised form 19 December 2014
Accepted 23 December 2014
Available online xxxx

Keywords:
Satellite-based precipitation estimates
Validation procedure
Additive and multiplicative error model
Logarithmic transformation
Uncertainty estimation

ABSTRACT

The objective of this study is to propose and test a new procedure to improve the validation of remote-sensing, high-resolution precipitation estimates. Our recent studies show that many conventional validation measures do not accurately capture the unique error characteristics in precipitation estimates to better inform both data producers and users. The proposed new validation procedure has two steps: 1) an error decomposition approach to separate the total retrieval error into three independent components: hit error, false precipitation and missed precipitation; and 2) the hit error is further analyzed based on a multiplicative error model. In the multiplicative error model, the error features are captured by three model parameters. In this way, the multiplicative error model separates systematic and random errors, leading to more accurate quantification of the uncertainties. The proposed procedure is used to quantitatively evaluate the recent two versions (Version 6 and 7) of TRMM's Multi-sensor Precipitation Analysis (TMPA) real-time and research product suite (3B42 and 3B42RT) for seven years (2005–2011) over the continental United States (CONUS). The gauge-based National Centers for Environmental Prediction (NCEP) Climate Prediction Center (CPC) near-real-time daily precipitation analysis is used as the reference. In addition, the radar-based NCEP Stage IV precipitation data are also model-fitted to verify the effectiveness of the multiplicative error model. The results show that winter total bias is dominated by the missed precipitation over the west coastal areas and the Rocky Mountains, and the false precipitation over large areas in Midwest. The summer total bias is largely coming from the hit bias in Central US. Meanwhile, the new version (V7) tends to produce more rainfall in the higher rain rates, which moderates the significant underestimation exhibited in the previous V6 products. Moreover, the error analysis from the multiplicative error model provides a clear and concise picture of the systematic and random errors, with both versions of 3B42RT have higher errors in varying degrees than their research (post-real-time) counterparts. The new V7 algorithm shows obvious improvements in reducing random errors in both winter and summer seasons, compared to its predecessors V6. Stage IV, as expected, surpasses the satellite-based datasets in all the metrics over CONUS. Based on the results, we recommend the new procedure be adopted for routine validation of satellite-based precipitation datasets, and we expect the procedure will work effectively for higher resolution data to be produced in the Global Precipitation Measurement (GPM) era.

1. Introduction

Space-borne precipitation products, with their global coverage, high resolution, frequent sampling and easy access, have been widely used in various applications (e.g., natural hazards,

* Corresponding author at: Hydrological Sciences Laboratory, Code 617, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

http://dx.doi.org/10.1016/j.atmosres.2014.12.016
0169-8095/

hydrology, agricultural forecasts, and climate studies). However, the errors associated with these satellite precipitation products need quantitative evaluation, because of the highly nonlinear nature of the physical process to measure precipitation from space. The strengths and limitations of those satellite precipitation products need to be understood so they can be interpreted correctly between the data-producing community and data users, especially during the Global Precipitation Measurement (GPM) era with large volume of higher resolution (~10 km, hourly) precipitation data expected to be generated in near future (Huffman et al., 2012).

Quantitative evaluation of satellite precipitation products is critical for both data producers and external users. On one hand, effective error analysis will yield insight into the sources of errors in the precipitation products and possible ways to correct or reduce them. This will lead to the improvement of next generation data algorithms and enhance their data quality. On the other hand, for end users, such evaluation and error characteristics analysis will give better guidance in selecting products for their particular applications, and help them assess the impact of input errors propagated into their applications (Tian et al., 2009).

The Algorithm Inter-comparison Projects (AIP) of the Global Precipitation Climatology Project (GPCP) (e.g., Arkin and Xie, 1994; Ebert et al., 1996), the Precipitation Inter-comparison Projects (PIP) (e.g., Smith et al., 1998; Adler et al., 2001), and a comprehensive validation study at global scale called the Pilot Evaluation of High Resolution Precipitation Products (PEHRPP) (Arkin and Turk, 2006) are examples of major past inter-comparison studies. The International Precipitation Working Group (IPWG, online at www.isac.cnr/it~ipwg/), builds upon the earlier evaluation experiences, has established a validation program to provide both the data producers and external users with up-to-date information on the quality of the precipitation estimates from virtually all the operational satellite algorithms (Ebert et al., 2007). Meanwhile, a number of new multi-sensor precipitation algorithms have been developed to exploit the complementary strengths of three different types of precipitation measuring sensors: Infrared (IR), Passive Microwave (PMW) Radiometers, and Precipitation Radar (PR) (e.g., Sorooshian et al., 2000; Joyce et al., 2004; Huffman et al., 2007). Some multi-sensor precipitation products also incorporate ground-based precipitation measurements, such as rain gauge data. A considerable number of evaluation studies have been devoted to the error analysis and uncertainty quantification for satellite-based precipitation products (e.g., McCollum et al., 2002; Gottschalck et al., 2005; Ebert et al., 2007; Hossain and Huffman, 2008; Lin and Hou, 2008; Tian and Peters-Lidard, 2007; Tian et al., 2007, 2009; Sapiano and Arkin, 2009; Kubota et al., 2009; Habib et al., 2009a; Tian and Peters-Lidard, 2010; Tian et al., 2010; Kirstetter et al., 2012; Tian et al., 2013; Chen et al., 2013a,b; Maggioni et al., 2014; Tang et al., 2014). Most of these researches used conventional error metrics to quantify the uncertainties (e.g. bias, root mean square error). For instance, Ebert et al. (2007) evaluated several operational satellite and numerical weather prediction (NWP) precipitation products, against gauge-based data sets over the continental US, Australia, and Europe, using several conventional error metrics (e.g., correlation, bias ratio, probability of detection and false alarm ratio, etc.). They found that satellite precipitation estimates are more accurate during summer and at lower latitudes. Meanwhile, Hossain and Huffman (2008) proposed a conceptual framework for developing error metrics in three general dimensions: 1) spatial (how does the error vary in space?); 2) retrieval (how "off" is each precipitation estimate from the true value?); and 3) temporal (how does the error vary in time?). They employed formulations for error metrics specific to each dimension, in addition to the conventional error metrics. They applied the error framework on four satellite precipitation products and found that this error framework can identify seasonal and regional differences in uncertainties of data sets more clearly than the conventional error metrics.

Some recent validation studies focused on the performance of the newest version (V7) of Tropical Rainfall Measurement Mission (TRMM) Multi-sensor Precipitation Analysis (TMPA). Yong et al. (2013) compared the performance of TMPA real-time and post-processed products 3B42RT and 3B42 Version 7 (V7) with the previous Version 6 (V6) over two river basins in China, and found that V7 algorithm significantly reduced systematic bias in the low-latitude river basin, while it was ineffective in the high-latitude river basin. Y. Chen et al. (2013) evaluated 3B42 V7 precipitation estimates for tropical cyclone rainfall on two terrain types: low-lying atoll sites (considered as open ocean), and coastal and island sites (land). The results show that 3B42V7 tends to overestimate heavy rain frequency on atoll sites, and underestimate heavy rain frequency on coastal and island sites. Chen et al. (2013a,b) gave comprehensive evaluations of TMPA V7 products over China and continental US against the daily gauge analysis, and found that relative bias and RMSE significantly decreased while correlation increases from V6 to V7. Generally most studies agree that for heavy rainfall, significant underestimation observed in 3B42 V6 is reduced in 3B42 V7 (e.g., Li et al., 2013; Xue et al., 2013; Y. Chen et al., 2013; Chen et al., 2013a,b). The underestimation in heavy rainfall is most severe over higher terrain (e.g., Y. Chen et al., 2013).

The errors in satellite-based precipitation estimates are from two major sources: sampling error and retrieval errors. The former results from estimation the precipitation amount for a continuous spatial and temporal domain with measurements at discrete space and time intervals, such as estimating the daily or monthly total precipitation from instantaneous observations at 3-hour intervals. The sampling error has been studied extensively, and its relationship with rain-rate and spatial/temporal resolution has been well established both empirically and theoretically (e.g., Laughlin, 1981; Huffman, 1997; Bell and Kundu, 1996, 2000, 2003; Bell et al., 2001; Steiner et al., 2003; Nijssen and Lettenmaier, 2004). As well, this part of the errors is beyond the scope of this paper.

The retrieval error arises from the remote-sensing procedures involved to convert satellite observations (brightness temperature) to rain rate. This error type is more complex, because of its dependencies on many factors, including sensor type (conical vs. cross-track, active vs. passive microwave), sensor resolution and viewing geometry, precipitation type, surface type, atmospheric condition, cloud microphysics, and retrieval algorithm itself (e.g., Arkin and Xie, 1994; Sorooshian et al., 2000; Adler et al., 2001; McCollum et al., 2002; McCollum and Ferraro, 2003; Gottschalck et al., 2005; Hossain and Anagnostou, 2006; Ebert et al., 2007; Tian et al., 2007; Tian and Peters-Lidard, 2010; Tian et al., 2010; Kirstetter et al., 2012;