Atmospheric Research 154 (2015) 39–59

Contents lists available at ScienceDirect

Atmospheric Research

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

How well can we estimate error variance of satellite precipitation data around the world?

Abebe S. Gebregiorgis *, Faisal Hossain

Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, United States

ARTICLE INFO

Article history:
Received 25 June 2014
Received in revised form 24 October 2014
Accepted 5 November 2014
Available online 12 November 2014

Keywords:
Satellite precipitation error variance
Topography
Climate
Season
Precipitation rate
Ungauged regions

ABSTRACT

Providing error information associated with existing satellite precipitation estimates is crucial to advancing applications in hydrologic modeling. In this study, we present a method of estimating the square difference prediction of satellite precipitation (hereafter used synonymously with “error variance”) using regression model for three satellite precipitation products (3B42RT, CMORPH, and PERSIANN-CCS) using easily available geophysical features and satellite precipitation rate. Building on a suite of recent studies that have developed the error variance models, the goal of this work is to explore how well the method works around the world in diverse geophysical settings. Topography, climate, and seasons are considered as the governing factors to segregate the satellite precipitation uncertainty and fit a nonlinear regression equation as a function of satellite precipitation rate. The error variance models were tested on USA, Asia, Middle East, and Mediterranean region. Rain-gauge based precipitation product was used to validate the error variance of satellite precipitation products. The regression approach yielded good performance skill with high correlation between simulated and observed error variances. The correlation ranged from 0.46 to 0.98 during the independent validation period. In most cases (~85% of the scenarios), the correlation was higher than 0.72. The error variance models also captured the spatial distribution of observed error variance adequately for all study regions while producing unbiased residual error. The approach is promising for regions where missed precipitation is not a common occurrence in satellite precipitation estimation. Our study attests that transferability of model estimators (which help to estimate the error variance) from one region to another is practically possible by leveraging the similarity in geophysical features. Therefore, the quantitative picture of satellite precipitation error over ungauged regions can be discerned even in the absence of ground truth data.

1. Introduction

Over the past two and half decades, remote sensing based precipitation estimates have experienced tremendous progress in providing the world a cost-effective and reliable ways of measuring precipitation from space (Adler et al., 2003; Huffman et al., 2001; Joyce et al., 2004; Kuligowski, 2002; Kidd et al., 2003; Miller et al., 2001; Sorooshian et al., 2000; Xie et al., 2007). As compared to ground observation system, satellite precipitation measurement technique, by far, is more effective to address the spatial and temporal variability of precipitation over the vast ungauged regions of the earth surface. It avoids the hurdle of geo-political boundaries issues; it covers both the terrestrial and water bodies of the earth; it provides a continuous (uninterrupted) observation irrespective of time (day/night), terrain and weather condition on the ground; it evades high operational cost of in-situ networks; and it delivers information on a near real-time basis.

Despite the obvious, the presence of non-negligible error (hereafter used synonymously with ‘uncertainty’) in satellite precipitation estimation presents a hurdle to fully implement

---

* Corresponding author.
E-mail address: abesine2002@gmail.com (A.S. Gebregiorgis).

http://dx.doi.org/10.1016/j.atmosres.2014.11.005

the product for wide ranges of hydrologic applications (Pan et al., 2010). Since it is an essential prerequisite for hydrologic applications, assessing the uncertainty of satellite precipitation estimate has become important over the last few years. It is important to understand the nature and quantify the magnitude of this uncertainty in order for users to apply the a priori knowledge to scientific research and practical applications. There are many publicly available high resolution satellite precipitation products (Huffman et al., 2007; Joyce et al., 2004; Sorooshian et al., 2000) available at a global scale which are potentially helpful for many scientific investigations and applications (Wu et al., 2012; Su et al., 2011; Pan et al., 2010; Shrestha et al., 2008; Su et al., 2008; Artan et al., 2007; Hong et al., 2006; among others). However, the question that remains is: '*to what level do the end users have the knowledge about the error information associated to these satellite precipitation products?*'

The advantage of knowing error information can be valuable from two perspectives: 1) from data producers or algorithm developers to improve data quality; and 2) from data users to improve data application. Investigating the components of error and enumerating each error individually can help algorithm developers (data producers) comprehend the strengths and weaknesses of their algorithms in a variety of settings, understand the aspects that are in greatest need of improvement, evaluate and monitor the performance of existing algorithms, and finally, assist with evaluating algorithm upgrades. On the other hand, data (end) users need to verify the accuracy of satellite precipitation products before using them for a particular application. A thorough verification of satellite-based precipitation products can provide users with information on the expected errors in a wide range of hydrologic application, so that they can quantify the expected accuracy in the prediction. With the existence of various satellite precipitation products, the users need to know the level of uncertainty in each product and its implication for a given surface hydrologic prediction. Therefore, data producers and end users can work together and allow information to flow both ways for communal advantage.

The source of satellite-derived precipitation uncertainty could arise from retrieval errors such as instrument, measurement and algorithmic biases, and sampling error (Nijssen and Lettenmaier, 2004; Huffman, 1997). It is a very challenging task to quantify the uncertainty of satellite precipitation estimate for many reasons. First, precipitation by itself exhibits random variation to represent the uncertainty with simple mathematical formulations (Wilks, 2011; North et al., 1993). Second, in case of high spatial and temporal resolution, there is a problem of assigning the rain field precisely for true location on the ground (Bellerby and Sun, 2005). Third, due to the nature of indirect measurement of precipitation processes such as by observing cloud-top properties in case of infrared (IR) sensor, and from thermal emission and backscatter signals in case of microwave (MW) sensor (Huffman, et al., 2010). In general, the accuracy of satellite based precipitation estimates depends on several factors: method of retrieval (type of algorithm), the nature of sensor used, the surface condition (ocean or land), and precipitation type and so on. The collected effect of all these factors makes the satellite precipitation estimates inescapably uncertain.

Moreover, to validate satellite precipitation estimates, ground truth data from rain gauge and/or radar observations are indispensable. The main problem, however, is that most parts of the globe are ungauged or has limited in-situ precipitation observation network. On the other hand, the existing observation networks continue to decline worldwide (Stokstad, 1999; Shiklomanov et al., 2002). The absence of in-situ measurement in most parts of the world represents a 'paradoxical' situation for evaluating satellite precipitation estimation uncertainty. Under such a circumstance, conventional validation of satellite precipitation products over these regions is quite impracticable and unrealistic. There is now a need for us to think outside the box for global applications. Therefore, a question we ask in our previous studies (Gebregiorgis and Hossain, 2013a, 2013b; Gebregiorgis et al., 2012) is, '*how can the uncertainties of satellite precipitation be estimated without having ground reference data?*' In fact, this question needs a novel approach to predict satellite precipitation uncertainty around the world.

The aforementioned challenges have already prompted the scientific community to design research strategies and recommendations for future investigations. For example, in a recent workshop conducted from 15 to 17 March 2010 at University of California, Irvine on Advanced Concepts on Remote Sensing of Precipitation at Multiple Scales (http://chrs.web.uci.edu/events/Workshop_Report.pdf), various research priorities and recommendations have recently emerged. Quantification of satellite precipitation uncertainty for different climate regions, storm regimes, surface conditions, seasons, and elevations was one of the major recommendations made by the community for advancing satellite remote sensing of precipitation (Sorooshian et al., 2011).

In line with these strategies, Gebregiorgis and Hossain (2013a) have demonstrated a method of estimating satellite precipitation error variance using readily available geophysical features and satellite precipitation rate for 3B42RT, CMOPRH, and PERSIANN-CCS products over Mississippi and Northwest basins. First, the basins were grouped into different regions based on topography and Köppen climate. Then, the nonlinear regression models were fitted for each region by taking into consideration the season type as shown in Eq. (1).

$$Ev_{i,j,t} = \alpha_{i,j,k}\left(RR_{i,j,t}\right)^{\beta_{i,j,k}} \tag{1}$$

where, *i* represents the topographic region, *j* the climate type, *t* the time at daily scale, and *k* the season type; *Ev* denotes pixel's error variance of the same spatial and temporal resolutions with satellite precipitation product; *RR* means the satellite precipitation estimates; α symbolizes the scaling factor for each topographic, climate region and season; β the same as α, except that it designates the power or exponent estimator of the regression model. In general, the model estimators α and β control the behaviour, shape and growth or decay of the fitted curve of the regression function. In this study, error variance refers to the square of the difference between the estimated and true value (square error).

The study of Gebregiorgis and Hossain (2013a) showed that the total error variance is directly proportional to the satellite precipitation rate, i.e. the value of β is always positive but has a different rate of growth for various topographic and climate regimes. Fundamentally, the seasonal precipitation variability and its type are strongly dictated by topographic and climate. It is, therefore, safe to claim that the use of geophysical features