Journal of Hydrology 511 (2014) 17–27

Contents lists available at ScienceDirect

Journal of Hydrology

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

Multivariate distributed ensemble generator: A new scheme for ensemble radar precipitation estimation over temperate maritime climate

Qiang Dai *, Dawei Han, Miguel Rico-Ramirez, Prashant K. Srivastava

Water and Environment Management Research Centre, Department of Civil Engineering, University of Bristol, Bristol BS8 1TR, United Kingdom

ARTICLE INFO

Article history:
Received 18 July 2013
Received in revised form 8 December 2013
Accepted 10 January 2014
Available online 18 January 2014
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Emmanouil N. Anagnostou, Associate Editor

Keywords:
Ensemble generator
Radar-rainfall estimates
Copula
Hydrological model uncertainty

SUMMARY

It is broadly recognized that large uncertainties are associated with radar rainfall (RR) estimates, which could propagate in the hydrologic forecast system and contaminate its final outcomes. Ensemble generation of probable true rainfall is an elegant and practical solution to characterize the uncertainty of RR estimates and behavior in the hydrologic forecast system. In this study, we have proposed a fully formulated uncertainty model that can statistically quantify the characteristics of the RR errors and their spatial and temporal structure, which is a novel method of its kind in the radar data uncertainty field. The error model is established based on the distribution of gauge rainfall conditioned on radar rainfall (GR|RR). It's spatial and temporal dependencies are simulated based on the t-copula function. With this proposed error model, a Multivariate Distributed Ensemble Generator (MDEG) driven by the copula and autoregressive filter is designed and applied in the Brue catchment (135 km^2), an extensively gauged site in the United Kingdom. The products from MDEG include a time series of ensemble rainfall fields with each of them representing a probable true rainfall. A series of tests show that the ensemble fields generated by MDEG have realistically maintained the spatial and temporal structure of the random error in RR as they have relatively low mean absolute errors (MAEs) of spatio-temporal correlation towards the observed ones. In addition, the results show that the simulated uncertainty bands derived by the 500 realizations of ensemble rainfall encompass most of the reference rain gauge measurements, indicating that the proposed scheme is statistically reliable.

1. Introduction

Weather radar has been widely used in hydrologic forecasting and decision making; nevertheless, there is an increasing attention on its uncertainty that propagates to hydrologic models. It has been acknowledged that the uncertainty of real-time hydrologic models comes from the model structure, model parameters and the rainfall input (Carpenter, 2003; Carpenter and Georgakakos, 2004; Krajewski et al., 1991). Many studies have been explored to solve the former two factors, such as the generalized likelihood uncertainty estimation (GLUE) approach (Beven and Freer, 2001; Beven, 2001; Montanari, 2005; Srivastava et al., 2013b), Bayesian concept (Ajami et al., 2007; Jin et al., 2010; Krzysztofowicz, 1999) and Markov Chain Monte Carlo (MCMC) (Benke et al., 2008; Vrugt et al., 2003). However, the study that accounts for the uncertainty of radar rainfall (RR) input in hydrologic forecasting is still in an early age (Ciach et al., 2007; Krzysztofowicz, 2001) though there are some studies attempt to solve this problem recently (Cunha et al., 2012; Fu et al., 2011; Schröter et al., 2011). In fact, the known uncertainty of the rainfall input should help in addressing other uncertainties of hydrological models.

Weather radars are subjected to several sources of uncertainties such as attenuation, ground clutter and occultation, partial beam blocking, the sampling and averaging method, and the variation of the vertical reflectivity (Dai et al., 2013a; Rico-Ramirez and Cluckie, 2007; Villarini et al., 2008). In reality, it is vastly difficult to solve these separate error sources step by step and study their interdependencies (Ciach et al., 2007). Therefore, statistically modeling the error of RR with the help of rain gauge information is a practical method to describe the overall feature of RR uncertainty. A number of error models related to this idea are enumerated in the literature. One of the simplest error models for RR is using the differences or ratio between radar and gauge rainfall (Habib et al., 2008). It is the fundamental conceptual framework of many studies. Barnston (1991) proposed a so-called error variance separation (EVS) method, which was formulated by Ciach and Krajewski (1999a) and Ciach (2003). The EVS is capable of estimating the error variances of radar rainfall and recognizing the rain gauge inaccuracy at the same moment. There are numerous

* Corresponding author. Tel.: +44 (0117) 331 5724, mobile: +44 7598957138.
E-mail address: q.dai@bristol.ac.uk (Q. Dai).

0022-1694/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2014.01.016

successive studies striving to apply or improve this scheme (Anagnostou et al., 1999; Bringi et al., 2011; Habib et al., 2004; Kirstetter et al., 2010). One significant improvement of this method, the conditional distribution transformation (CDT), is proposed by Habib et al. (2004). The CDT was designed to filter out rain gauge errors to retrieve the bivariate probability distributions of RR estimates and the corresponding true rainfall (Morrissey, 1991; Morrissey and Greene, 1993). Ciach et al. (2007) proposed a product-error driven (PED) method to quantitatively model the RR error in a probabilistic form by separating the RR uncertainties into deterministic and random errors. Both of them are estimated by a nonparametric functional estimation approach. The essence of this model is to formulate an empirically based error model for probabilistic precipitation estimation through comparing the RR estimates with the corresponding assumed true rain gauge measurement. Many applications in various catchments have been undertaken based on the model (Habib and Qin, 2011; Villarini et al., 2009, 2010).

However, these error models focus on the signal radar pixel error based on a single point (rain gauge), without containing the spatial and temporal correlation for the whole study area. Though some studies have noticed and analyzed the correlation (Ciach, 2003; Ciach et al., 2007), it's possible but hard to incorporate the spatio-temporal features into their error models by strict and realistic equations. Nowadays, ensemble generation of a massive number of probable true surface rainfall fields with their natural spatio-temporal structure to represent the RR uncertainty is considered to be an effective and practical approach. In addition, an ensemble generator of estimated rainfall can be easily integrated with radar rainfall forecast model or numerical weather prediction to generate ensemble forecasts of rainfall (Liechti et al., 2013). Therefore, in this study, we propose a new scheme for generating ensemble rainfall fields to model the RR uncertainty and its spatio-temporal structure.

The study on ensemble radar rainfall generation can be dated back to 1985, when Krajewski and Georgakakos (1985) proposed a scheme to generate synthetic radar-rainfall by imposing random noise on the known radar-rainfall field. The noise field is determined based on the statistical conditions such as of the mean, variance and correlation to the resultant field. The idea of modeling the stochastic random noise and generating a number of noise fields is considered to be the basic conceptual framework in the foregoing studies. Germann et al. (2009) proposed a so-called REAL solution to generate an ensemble of precipitation fields by characterizing the residual errors in radar-rainfall estimates in a mountainous region. It is based on singular value decomposition of the error covariance matrix using the LU decomposition method (Cressie, 1992; Goovaerts, 1997) and autoregressive filtering (Priestley, 1981). However, there is no strict error model used in this study. Instead, the error of radar-rainfall is simply regarded as the differences between radar and gauge rainfall for every time step, and the variance was computed using the simplified mathematical definition with the radar intensity as the weight. Villarini et al. (2009) also used the LU decomposition method but with more consideration on the error model. The product-error-driven error model proposed by Ciach et al. (2007) was introduced to compute the systematic bias and build the covariance matrix. In the same manner to Germann et al. (2009), this study required that the residual error of radar-rainfall was Gaussian distributed due to the constraint of the LU decomposition method. Nevertheless, this assumption is not always satisfied in all catchments (Anagnostou et al., 1999; Ciach and Krajewski, 1999b). More importantly, the temporal dependence of residual error is neglected as the covariance matrix cannot satisfy the presumed condition when the spatial and temporal correlation is considered at the same time. In a recent study, AghaKouchak et al. (2010a,b) developed an ensemble generation method for remotely sensed rainfall estimates using copulas. The simulated rainfall fields consist of the original radar rainfall and the uncertainty component multiplying the radar rainfall under the assumption that the radar rainfall is proportional to the magnitude of rain rate. The spatial dependence of the random errors is simulated through a presumed copula, while the random errors are assumed to be temporally uncorrelated. The focus of this research is based on the selection of appropriate copula functions, with a further consideration of the marginal distribution of the RR errors.

The copula is defined as a standardized multivariate distribution function whose one-dimensional marginal distributions are uniform on $[0,1]$ (Nelsen, 1999). It is mainly applied in the financial sector. In hydrology, copulas by far have been used in analyzing extremes, geo-statistics, etc. (Bárdossy and Li, 2008; Zhang et al., 2012, 2013). As is stated earlier, AghaKouchak et al. (2010a,b) attempted to apply copulas in RR uncertainty, but only spatial dependence was considered. However, the essential integration of marginal distribution and spatio-temporal dependence were not given. In fact, the main advantage provided by the copula is that the marginal distribution for each radar pixel can come from different distribution families or different parameters for a family, without affecting the spatio-temporal structure. Owing to this benefit, this study presents the first fully formulated error model for radar precipitation estimation based on a multivariate distribution. With this error model, a Multivariate Distributed Ensemble Generator (MDEG) is implemented and applied in the Brue catchment. The generated ensemble rainfall can maintain the marginal distribution constraint, spatial dependence and temporal dependence.

The ensemble generator consists of two components: the error model established component and the real-time running component. The error model is built for each radar pixel and the spatial and temporal correlations among these radar pixels are also calculated. With these outcomes, we can generate the ensemble rainfall with given radar-rainfall uncertainty and spatio-temporal correlation in a real-time radar-rainfall estimated system. Specifically, this paper is organized as follows. After the introduction, Section 2 describes the algorithm of error model for radar precipitation estimation. The sections on the ensemble generator cover the concept and processing of MDEG. The data section describes the radar–gauge data used. Then Section 5 presents the outcomes of the simulation, evaluation, and potential applications of the MDEG. Finally, the conclusion section summarizes the key findings and the future work.

2. Error model and its spatio-temporal structure

In quantifying the uncertainties in a given RR estimate, we assume the true areal averaged pixel-scale rainfall is composed of two components, namely the deterministic distortion component and the random component. Theoretically speaking, the deterministic bias is categorized into overall bias and conditional bias. The overall bias, also known as systematic bias, is the overall discrepancy between RR values and reference rainfall. The conditional bias illustrates the divergence of RR towards the reference rainfall for different rainfall intensities. In this study, the process of removing the overall bias is considered as a fundamental quality control before the ensemble generation. Thereby the deterministic distortion bias only refers to the conditional bias herein. The random component represents the combined effect of all random error sources. In essence, the ensemble generator for RR estimation is to produce a number of probable true rainfall outcomes by adding a series of random fields to the deterministic component of the rainfall. Thus the strictly statistical representation of the ensemble generator is expressed as: