Journal of Hydrology 522 (2015) 544–557

Contents lists available at ScienceDirect

Journal of Hydrology

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

Applying bias correction for merging rain gauge and radar data

E. Rabiei *, U. Haberlandt

Institute of Water Resources Management, Hydrology and Agricultural Hydraulic Engineering, Leibniz Universität Hannover, Appelstr. 9A, 30167 Hannover, Germany

ARTICLE INFO

Article history:
Received 14 October 2014
Received in revised form 15 December 2014
Accepted 6 January 2015
Available online 14 January 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Ashish Sharma, Associate Editor

Keywords:
Rainfall
Radar
Bias correction
Merging
Geostatistics

SUMMARY

Weather radar provides areal rainfall information with very high temporal and spatial resolution. Radar data has been implemented in several hydrological applications despite the fact that the data suffers from varying sources of error. Several studies have attempted to propose methods for solving these problems. Additionally, weather radar usually underestimates or overestimates the rainfall amount. In this study, a new method is proposed for correcting radar data by implementing the quantile mapping bias correction method. Then, the radar data is merged with observed rainfall by *conditional merging* and *kriging with external drift* interpolation techniques. The merging product is analysed regarding the sensitivity of the two investigated methods to the radar data quality. After implementing bias correction, not only did the quality of the radar data improve, but also the performance of the interpolation techniques using radar data as additional information. In general, conditional merging showed greater sensitivity to radar data quality, but performed better than all the other interpolation techniques when using bias corrected radar data. Furthermore, a seasonal variation of interpolation performances has in general been observed. A practical example of using radar data for disaggregating stations from daily to hourly temporal resolution is also proposed in this study.

1. Introduction

Advanced technologies like weather radar help to increase our knowledge regarding the spatial structure of rainfall events. Although weather radar provides rainfall data with relatively high spatial and temporal resolution, the data is subject to several sources of error. Beside common problems associated with weather radar, e.g. existence of clutters and attenuation, the data suffers from the fact that weather radar usually either overestimates or underestimates rainfall. There are several physical factors affecting the accuracy of rainfall measurement which are not all recognised quantitatively, but rather qualitatively. Errors related to weather radar data have been investigated by several studies. Austin (1987) studied the complexity of the relationship between rain intensity derived from radar reflectivity and surface rainfall. She discussed the influence of precipitation type, the existence of frozen particles and several other influential factors on the relationship between radar reflectivity and rain intensity, or the *Z–R* relationship. Others proposed methods trying to compensate common problems like detecting ground clutters by analysing radar pixels, implementing sophisticated algorithms for transforming reflectivity to intensity (Alfieri et al., 2010), attenuation calibration (Rahimi et al., 2006), etc. Alfieri et al. (2010) studied a simple procedure for using continuously updated *Z–R* relationships in time to produce real time rainfall estimation.

Despite the difficulties that radar data has, several studies (e.g. Quirmbach and Schultz, 2002) tried to use radar data directly as an input for water management purposes. In such circumstances, the radar data quality plays a significant role considering the above mentioned problems. On the other hand, merging radar data and rain gauge data is a traditional way to describe rainfall fields when considering the rain gauge network as providing true information. In order to combine the rainfall estimation from radar and the accurate point information from stationary rain gauges, a variety of methods including co-kriging (Krajewski, 1987), kriging with external drift (Haberlandt, 2007; Verworn and Haberlandt, 2011), conditional merging (Ehret, 2002), have been proposed. Most of the methods consider the radar data as secondary information to estimate the rainfall field. In kriging with external drift, it is assumed that the expected value of the primary variable is linearly related to the additional variable. This assumption is not always fulfilled. Although Ehret (2002) did not assume linearity of radar data to the primary variable in conditional merging, the quality of radar data is still an important factor in this method. Berndt et al. (2014) excluded time steps with poor radar quality in order to take into account the influence of radar data quality for merging. They used two criteria: (1) maximum radar rainfall values and (2)

* Corresponding author.
E-mail address: rabiei@iww.uni-hannover.de (E. Rabiei).

http://dx.doi.org/10.1016/j.jhydrol.2015.01.020
0022-1694/

standard errors (between the gauge rainfall values and the corresponding radar-pixel values) for detecting time steps with poor radar data. The poor radar data are detected when exceeding either the 99th percentile of the empirical distribution of the maximum radar rainfall values or the 98th percentile of the empirical distribution of the standard errors.

In addition to merging radar and station data, several studies attempt to adjust the radar image according to rain gauge information. Erdin et al. (2012) implemented a Box–Cox transformation of radar and station data to improve the compliance with model assumptions. However, they recommend attention in implementing this method to avoid excessive transformation which can introduce positive bias. Chumchean et al. (2006) corrected radar data for the mean field bias which resulted in improving radar data quality. Besides, they used different parameters in Z–R relationship for different types of rainfall which also improved the radar data quality. Vogl et al. (2012) assimilated radar and gauge information to derive bias-corrected precipitation fields implementing copulas. This method requires calibration and fitting of the marginal distribution functions. Thorndahl et al. (2014) investigated the use of mean field bias adjustment for correcting radar data. They found that a larger bias exists during summer periods compared to winter. This seasonal variation of error was justified by rainfall type, where a larger bias belongs to convective storms and a smaller to stratiform events.

Quantile–quantile (Q–Q) transformation is usually employed in climate impact studies for scaling and bias correction purposes. Ines and Hansen (2006) corrected the daily General Circulation Models (GCM) rainfall for crop simulation studies. They fitted the data into the gamma distribution function and corrected the daily GCM rainfall accordingly. Jakob Themeßl et al. (2011) found quantile mapping to have the best performance, especially at high quantiles, compared to seven other methods they implemented for reducing regional climate model error characteristics. Chen et al. (2013) compared the performance of six bias correction methods for hydrological modelling over 10 North American river basins. They conducted bias correction on a monthly basis and applied two quantile mapping methods based on (a) an empirical distribution, and (b) a gamma distribution. Bárdossy and Pegram (2011) implemented this method for downscaling regional climate model precipitation to observed values. Additionally, they used double Q–Q transformation for future scenarios. To our knowledge, all of these studies consider a long time period of the observation and target data, which is here radar data, for estimating the bias. The length of this considered time period accordingly plays an important role. For points where no observation data is available, one may use interpolation techniques which introduce uncertainty into the work. This means that the final result depends not only upon the length of the time period, but also the performance of the interpolation techniques. Teegavarapu (2014) implemented two different quantile-based bias-correction methods as well as an optimal single best estimator (SBE) method for corrections of spatially interpolated missing precipitation data. They figured out that using bias-correction methods overcomes the over and underestimation of low and high extremes. Among them, the equi-distance quantile-matching performed the best. Gyasi-Agyei and Pegram (2014) used Q–Q transform to normalise the daily rainfall data for later determination of marginal frequency distribution of rainfall at all sites on the day.

Correcting radar data by applying a quantile mapping transformation and considering the observation network data as the reference is the main objective in this study. In this paper, the bias is defined as the difference between the radar-pixel values and the rain gauge corresponding values.

This paper is organized as follows. After Section 1 the methodologies implemented in this study, are described. Section 3 is then provided. Section four discusses the results. A short summary of the work, comparison of different scenarios and possible use of the method in practice is provider thereafter.

2. Methodology

Considering the value of each radar pixel representing its average rainfall amount occurring over a certain time and space, a large deviation between radar-pixel data and the accurate point-measurement devices like ordinary rain gauges can be detected. Because of this deviation, merging these two data sources might not be optimal, especially for the time steps where this deviation is highest. In the following, by implementing quantile mapping technique on the radar data, the radar image for each time step is corrected assuming that the spatial bias in the radar data dominates. In Section 3 part, the reasons for taking this assumption are discussed.

The methods, assumptions, and definitions used in this study are explained in this section.

2.1. Q–Q transformation

As described in several studies, the basic idea of this method is to correct one data source considering another data source as true by comparing their probability distribution functions. In this study, first theoretical distribution functions are fitted to the two data sources. Then, the quantile for each radar pixel value is estimated (the data source which will be corrected) from its cumulative distribution function (CDF). Thereafter, by considering the estimated quantile and using the inverse CDF of observed station data, the radar-pixel value is replaced. As mentioned earlier the primary assumption is that the rain gauge network is providing true information. Eq. (1) formulates the correction procedure:

$$Z'_R(x,t) = F^{-1}_{obs,t}(F_{rad,t}(Z_R(x,t))) \tag{1}$$

where $Z_R(x, t)$ is the value of radar cell at position x and time t, $F_{rad,t}$ is the cumulative distribution function estimated from radar data at time t, and $F^{-1}_{obs,t}$ is the inverse cumulative distribution function derived from the rain gauge network at time t which converts the quantiles estimated by $F_{rad,t}$ back to rain intensities, $Z'_R(x,t)$. The inverse cumulative distribution function is estimated from observed rainfall data.

In contrast to conventional implementation of the quantile mapping method where a certain time period from the two data sources is considered, in this study the radar image is corrected for each time step separately. This means that each radar image is corrected independently. There are two general ways to estimate the quantiles for each value in a data source, either implementing an empirical distribution function or fitting a theoretical distribution function to the data and estimating the quantiles accordingly. Using empirical distribution functions introduces uncertainties when too few points from the data source exist. This problem could be solved by applying an interpolation method, e.g. linear interpolation, but estimating quantiles between the points which are located far from each other might be an unrealistic approach. Instead of implementing empirical distribution functions with unknown uncertainties introduced when an interpolation method is applied, it is decided to use a theoretical distribution function. Fig. 1 illustrates the method visually.

The first step is to choose the time steps to correct. For this, different criteria need to be considered. Since there is usually enough data from the radar data source, the time steps chosen for correction depend on: (a) the number of available stations for each time step and (b) the average rainfall recorded at the available stations. In order to increase the sample size from the rain gauge network